

Publication and Collaboration Patterns in Autonomy Research – A Bibliometric Analysis

Felix Schulte

Journal of Autonomy and Security Studies,
3(2) 2019, 38–65

URL: <http://jass.ax/volume-3-issue-2-Schulte/>

Abstract

Research on territorial autonomy has gained new impetus in recent years. This research note presents a first comprehensive bibliometric analysis of autonomy studies. It introduces the Territorial Autonomy Literature Datasets (TALD), surveys of over 800 peer-reviewed and non peer-reviewed articles published between 1945 and 2018. The study reveals significant imbalances in gender and origin of authors, methodological approaches and studied cases. While the data shows some trend towards greater diversity and team collaboration, we observe that autonomy research is still dominated by male and Western-based scholars, and by single-authored small-n studies on sub-national regions in Europe and post-Soviet Eurasia. Thematically, the analysis shows that researchers almost exclusively study autonomous regions in the context of conflict regulation and minority accommodation.

Keywords

Bibliometric analysis; Social-science methods; Author networks;
Scientific community; Territorial autonomy

About the author: Felix Schulte is a lecturer in the Institute of Political Science at Heidelberg University.

1. Introduction

Antigone buries her brother against the explicit command of the tyrant Kreon, is trapped alive in a cave as punishment, and dies because of the self-imposed law as a consequence of her personal [*autónomos*] self-determination (Sophocles 1672). The praise of the Theban elders in Sophocles' tragedy is considered the oldest use of the word 'autonomy' in literature. Since then, autonomy has become a widely used and much-contested scientific concept that comes up in a variety of fields. A Web of Science search reveals over 20,000 records in disciplines such as medicine, psychology, and robotics (Web of Science 2019). Despite its multidisciplinary use, the etymology of the word draws attention to the fact that autonomy is originally a political concept. Composed of *autos* (gr. self) and *nomos* (gr. law or rule), the term describes the right to make its own laws or to live by one's own rules (Taylor 2019).

In ancient Greek, a polis was considered autonomous if it could govern its internal affairs independently of any superior power. For Herodotus, the counterpart of external autonomy was foreign domination, heteronomy. He used (internal) autonomy as a synonym for self-determination to describe the opposite of tyranny. It was then Thucydides who applied the term only to forms of internal self-determination, and who argued that an autonomous polis does not necessarily include its own jurisdiction or financial sovereignty (see Achouri 2013, Gerolemou 2017). According to this modern interpretation, the concept describes the territorially limited self-determination of a political community within a larger sphere of power. This understanding of autonomy as *territorial* self-governance is still predominant in comparative political science and has gained new impetus with the wave of decentralization that has been affecting most countries since the 1980s.

Nevertheless, (territorial) autonomy remains a highly contested concept, with its precise definition criteria being the subject of ongoing academic discussions. The conceptual debates revolve mainly around the question whether (territorial) autonomy is a concept sui generis, which can be distinguished from other forms of vertical power division, such as federalism, decentralization, or devolution, or whether autonomy should be used as a generic term to describe the overall level of self-rule a sub-state entity enjoys (see Trinn and Schulte 2019). Many political scientists subsume forms of territorial self-governance under the broader concept of political power-sharing. Following the groundbreaking work of Arend Lijphart, regional autonomy is often considered the 'second pillar' of consociationalism next to grand coalitions and veto rights (see Wolff and Yakinthou 2012). As this literature analysis will show, the synonymous use of these concepts can be traced back to the fact that research on autonomies has dealt from the outset, and almost exclusively, with the management of ethnic diversity and severe cultural cleavages, the addressing of which is also the *raison d'être* of consociational theory. However, the use of

umbrella terms such as 'complex power-sharing' (Wolff 2013), 'territorial power-sharing' (Hartzell and Hoddie 2015), or 'dispersive power-sharing' (Strøm et al. 2017) has led to some conceptual stretching and is not unproblematic, as power-sharing and vertical self-governance are institutions based on a different functional logic: while power-sharing arrangements in parliaments or cabinets aim at integrating relevant segments of society together into political decision-making processes, territorial autonomy works in the opposite direction by distributing political power away from the center and creating more or less autonomous entities with special status. Empirical cases in which both functional principles are institutionalized, such as in Nunavut (Canada), Northern Ireland (United Kingdom), or South Tyrol (Italy), are empirically very rare phenomena and therefore are not appropriate conceptual blueprints. The same applies to forms of "non-territorial autonomy", for which we find almost exclusively historical examples, such as the Millets in the Ottoman Empire or the Jewish National Council in Mandatory Palestine (see Malloy, Osipov, and Balázs 2015; Osipov 2018). In addition, most such labelled cases are not really 'non-territorial' but represent situations in which group-based and area-based distinctions are fused. Also, most cases do not constitute instances of political power, that is the ability to rule, but of rather protective minority rights.

The research literature that explicitly deals with territorial autonomy as a dependent or independent variable is comparatively young. While intellectuals such as Karl Renner and Otto Bauer (Springer Rudolf (Renner) 1902) had already discussed autonomy solutions to prevent the break-up of the multi-ethnic Austrian Empire on the eve of the First World War, scientific studies on territorial autonomy only gained momentum in the mid-1980s. Milestones such as Dinstein's "Models of Autonomy" (Dinstein 1981), Hannum's "Autonomy, Sovereignty, and Self-Determination" (Hannum 1996), or Lapidoth's "Autonomy - potentials and limitations" (Lapidoth 1997) paved the way for a rapidly increasing number of empirical publications dealing with autonomous regions. Figure 1 illustrates the growing number of publications since 1945 indexed with the terms 'territorial autonomy' or 'regional autonomy' in the Social Science Citation Index (SSCI) and Google Scholar (Web of Science 2019; Google 2019).

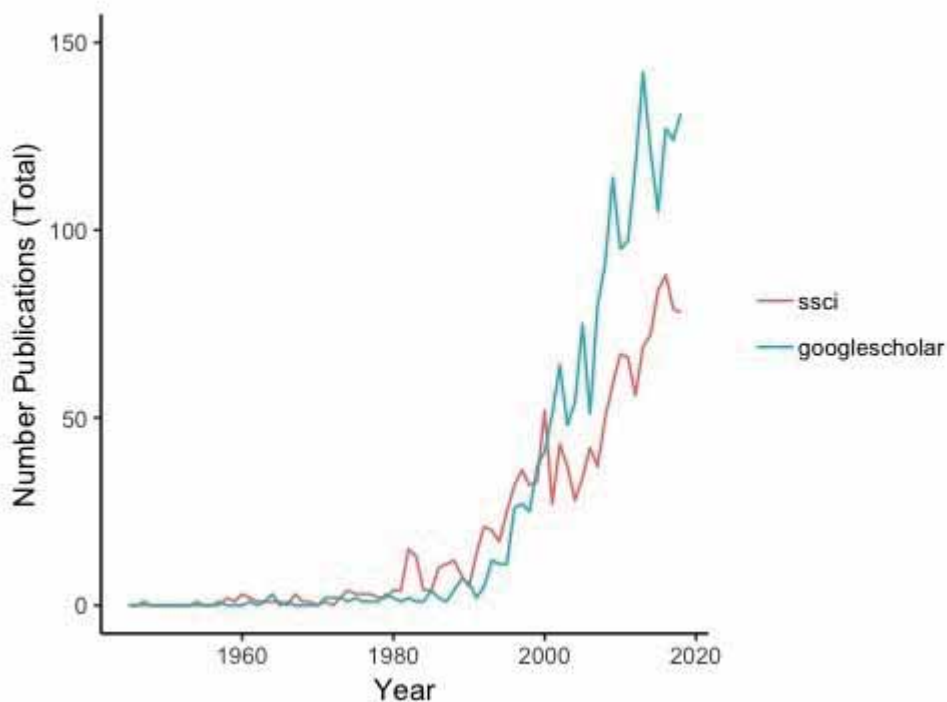


Figure 1: Autonomy publications in SSCI and Google scholar (1945–2018)

In recent decades, autonomy research has undoubtedly developed into a growing and enormously dynamic field. This paper looks back on the past decades of intensive research. Who are the autonomy researchers? Where do they come from, and in which countries do they work? What are the key publications in the field, and which cases are studied with which empirical methods? What is the central research interest of autonomy studies? To answer these questions, this research note conducts a bibliometric analysis of the relevant literature. The aim of the explorative study is to provide a comprehensive overview of the research field and to assess its cooperative and communicative integration. The article proceeds as follows. The next section describes the methodology, particularly the data sets compiled for the bibliometric analysis. Section 3 looks at authors, their gender, and their origin, as well as at the academic stars in the field and important co-citation networks. Section 4 analyzes the empirical interests and research designs used in autonomy studies and shows which empirical cases have been the focus thus far. The fifth section looks at the most important journals and analyses the keywords and abstracts of the publications in order to draw conclusions about their content and thematic focus over time. The final section summarizes the main findings and draws some conclusions for the further development of the research field.

2. Methodology

Bibliometric analyses have been conducted to analyze the influence of scholars and their publications, to explore publication trends, or to study collaboration patterns and citation networks in political science in general and its various sub-fields (Fisher et al. 1998; Pehl 2012; Metz and Jäckle 2017; Goyal and Howlett 2018; Pelke and Friesen 2019). The statistical analysis of publication data enables researchers to systematically examine a large number of publications and to identify strategic and thematic changes in a research field over time. Therefore, it usefully supplements classical literature overviews that usually concentrate on the most influential publications of a literature strand and summarize substantive findings to a particular topic.¹

This research note focuses on journal articles, which represent the primary form of scientific publication. In addition, scientific articles are suitable elements for bibliometric studies as they are usually of comparable length and text structure and provide readily extractable abstracts and keywords that can be easily analyzed. As a first step, all articles in the Web of Science (WoS) that are found with the search terms 'Territorial Autonomy', 'Regional Autonomy', 'Political Autonomy', 'Cultural Autonomy' or 'Power-Sharing'/ 'Power Sharing' are selected. As a second step, all articles which do not fall into social science disciplines are excluded from the sample. All remaining texts are combined into the *Territorial Autonomy Literature Data Set* (TALD1). TALD1 contains 827 articles published between 1945 and 2018. This data set provides full bibliographic information, such as names of authors, co-authors, full references, keywords, and abstract texts, and is thus ready-to-use for all bibliographic analyses, such as Co-Authorship Network Analysis (COA) and Author Co-Citation Analysis (ACA) to analyze cooperation and citation patterns in autonomy studies.² The WoS does not provide for data on gender or origin of authors, empirical methods, or cases. In addition, some journals in the research area, such as 'Ethnopolitics', the 'Journal on Ethnopolitics and Minority Issues in Europe', or the 'Journal of Autonomy and Security Studies' are not listed in the WoS, which possibly excludes relevant articles from the analysis.

For these reasons, a second data set (TALD2) is compiled, which contains the missing information and promises more robust results. For TALD2, all articles published between January 1980 and December 2018 in political science journals and listed with the keywords

-
- 1 In autonomy studies a plethora of thematically more or less specific literature overviews are available (Brancati 2008; Bednar 2011; Gagnon and Keating 2012), which are not discussed in greater detail here.
 - 2 COA takes into account the frequency of author collaborations. ACA, in contrast, analyzes citations and can thus draw conclusions about the relative importance of individual authors for the research field as well as their thematic "distance". If two authors appear in the same bibliography of a publication, they are co-cited. The more co-citations exist, the more likely it is that the research content of the authors is similar (Wang, Bu, and Huang 2018).

'Territorial Autonomy' or 'Regional Autonomy' in the title or free text in the Social Science Citation Index (SCCI), POLLUX (Informationsdienst Politikwissenschaft) or HEIDI (Bibliographic catalogue Heidelberg University) are selected. TALD1 and TALD2 exclude book reviews but include articles in non peer-reviewed journals. This is done to reach the aim of mapping the research field as comprehensively as possible. Assuming an ideal scientific world with a barrier-free exchange of ideas and knowledge between well-connected scholars, we expect authors to be influenced by texts in both peer-reviewed and non-reviewed journals.

TALD2 contains 588 articles in 228 journals and provides information for each publication on the number of authors (*number_aut*), their biological gender (*gen*), their regional origin and current domicile (*origin* and *home*), as well as the ethnic heterogeneity of this country (*heterog*).³ The variable *interest* states the primary research interest of the respective article. A distinction is made between a) primarily theoretical work or work that is interested in legal aspects, b) primarily empirical-analytical work that, however, does not develop original theoretical arguments, c) articles that develop both a theoretical argument and test devolved hypotheses empirically, and d) methodological studies that primarily concentrate on methodological questions and the implementation or application of empirical methods. Furthermore, the data set provides information on whether the article makes a causal argument (*causal_claim*) and which empirical methods authors apply. The variable *method* indicates whether the article uses descriptive or inferential statistical methods, applies set-theoretical methods such as Qualitative Comparative Analysis (QCA) or Necessary Condition Analysis (NCA), is based on comparative or single-case studies, uses a mixed-methods research design, or does not apply any empirical method in the proper sense. The variables *cases_region* and *number_cases* provide information on the regional focus of the study and the number of cases investigated. The codebook in the Appendix gives a detailed overview of the variables included in TALD2.⁴

3 Heterog1 describes the number of relevant ethnic groups in the respective country (home) based on information in the EPR data set (Wimmer, L.-E. Cederman, and Min 2009); Heterog2 describes the occurrence of violent self-determination movements in the respective country during the period under study (Sambanis et al. 2018); Heterog3 describes the number of ethnic autonomies in the respective country (Panov and Semenov 2018).

4 Both datasets can be downloaded here: <https://www.felix-schulte.de/data/tald/>

3. Authors and author networks

3.1 Gender and origin

Testing the ideal of a vibrant and well-connected intellectual community, the empirical evaluation of the data sets shows that jointly written contributions are not the predominant pattern in autonomy research. While the average number of authors per article has slowly yet steadily increased from 1.0 in 1980 to around 1.5 in 2018, a large majority of articles (71.6 percent) is single-authored. Only 8 percent of the articles in TALD2 are written by three or more authors. This is not untypical in political science literature, with similar numbers being observable in various sub-fields and countries (Arzheimer and Schoen 2009; Cancela, Coelho, and Ruivo 2014; Leifeld and Ingold 2016; Pelke and Friesen 2019). The analysis of the author's gender clearly shows that autonomy research is a male-dominated research field, with 68 percent of first authors and 75 percent of second authors being male. However, the percentage of female authors has grown significantly from 17 percent in the 1980s to around 30 percent in the 1990s, and 36 percent in the 2000s (Figure 2). Even though in the 1990s the proportion of women as co-authors was even higher than their proportion as first authors, the ratio has reversed in recent years.⁵ The underrepresentation of women becomes particularly clear when we look at the number of female-only articles, which make up just 3 percent of the sample. Although a slight increase can also be observed here, female author teams, such as Szöecsik/Zuber (2015), Åkermark/Stephan (2013), or Willett/Giovannini (2014), are still rare.

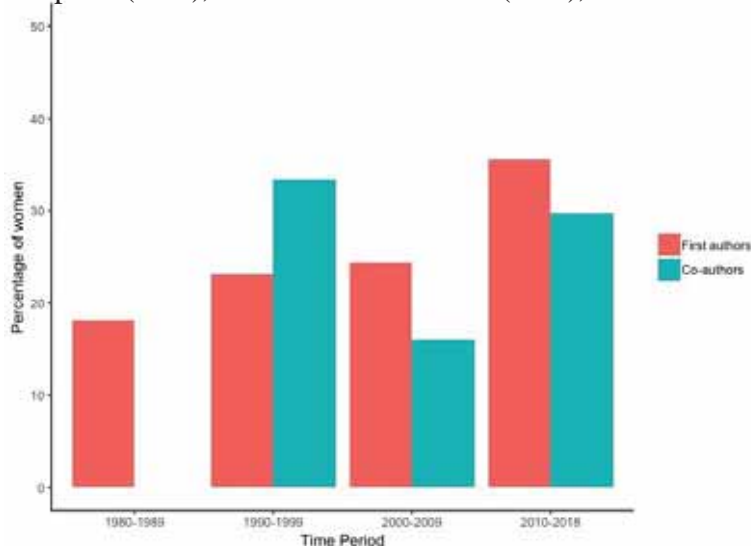


Figure 2: Percentage of women as first authors and co-authors

5 First authorship is determined by which author is named first in an article. A closer look reveals that the order of names is usually not purely for alphabetical reasons. It is much more common for men to be listed as first authors and women as second authors than vice versa.

Also notably underrepresented in the field are non-Western authors. 71 percent of the publications stem from authors from Northern, Western or Southern European countries, North America, and Canada. Their home institutions are usually large, research-strong universities in the United States, the United Kingdom, and Canada. Research institutions in Australia, Germany, and Switzerland are also widely represented. 13 percent of first authors come from post-Soviet countries, including the Balkans and Russia. Authors from other world regions together account for just 16 percent. While authors from the Middle East and North Africa (MENA) region have a share of 4 percent, scholars from Oceania, Sub-Saharan Africa, and Southeast Asia account for only around 3 percent of the sample. Publications by scholars from South Asia, East Asia, Latin America, and the Caribbean are hardly to be found in the research field. They each make up only about 1 percent of the total. While publications by non-Western authors before 1990 are practically non-existent, their share is continuously rising. Nevertheless, the analysis of the data set shows that the regional imbalance remains very much in favor of Western institutions and authors.

3.2 Academic stars and author networks

Co-authorships imply direct communication and intellectual debate between individual scholars and can therefore serve as an indicator for scientific collaboration and an intensive exchange of ideas and knowledge (Leifeld et al. 2017). The above-mentioned average number of authors per article already suggests that autonomy studies are a rather fragmented and less-cooperative field lacking extensive author networks. This assumption is confirmed by a systematic co-authorship (COA) analysis. While a rather low degree of author cooperation is typical for the social sciences compared to the natural sciences, it is remarkable that more or less no distinctive cooperation networks can be found in the data. The research field appears to be a collection of smaller, loosely coupled intellectual islands rather than a cohesive and well-connected scientific world. This may be due to the fact that autonomy research is still a cross-cutting issue that attracts researchers from many other sub-fields, such as comparative politics, legal studies, or conflict research, while only very few researchers focus exclusively on autonomous regions as a research topic. It seems that there is a wide gap in terms of professional backgrounds, theoretical positions, or methodological approaches that prevents autonomy researchers from publishing together. A notable exception is a collaboration network of conflict researchers at the ETH Zurich in Switzerland, including Lars-Erik Cederman, Andreas Wimmer, and Simon Hug, among others, which has published several influential journal articles (Cederman et al. 2015; Cederman, Gleditsch, and Wucherpfennig 2017; Cederman, Skrede, and Wucherpfennig 2018). When writers cooperate, they mostly do it in duos. Some of the two-writer teams

have built remarkably durable cooperations and have contributed substantially to the field. Examples include the highly successful cooperations between Caroline Hartzell and Matthew Hoddie, John McGarry and Brendan O’Leary, or Arjan Schakel and Emanuele Massetti.

Citations indicate not only an intellectual relationship between citing and cited source, but also thematic relationships. The author co-citation analysis (ACA) draws a rather different picture than the co-authorship analysis. It shows a very integrative network with a dominant epicentre consisting of a single academic – Arend Lijphart. A graphical illustration of the ACA illustrates his outstanding position (Figure 3).⁶

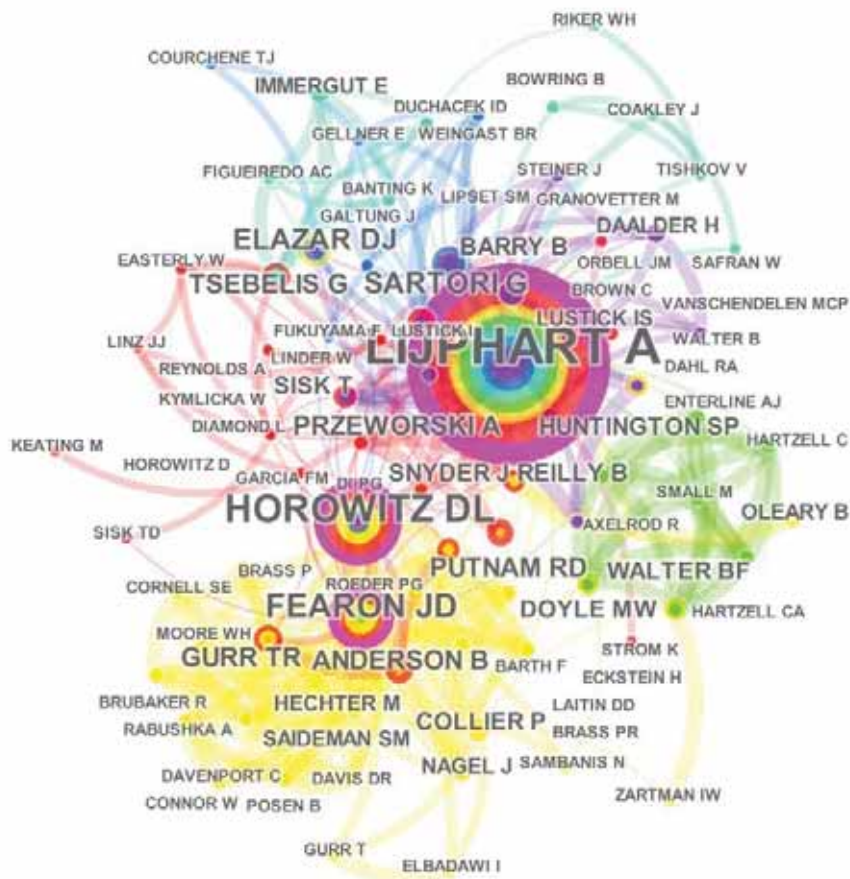


Figure 3: Author co-citation network

6 It must be noted, however, that the ACA does not take into account whether a source is cited agreeingly or critically or whether it is only referred to for strategic reasons. CiteSpace in Version 5.5 (2018) is used for analysis (Chen 2018). Thresholds are set $x > 2$ for cited references, $x > 3$ for co-citations, and $x > 0.2$ for the co-citation coefficient. For illustrative reasons, only a part of the co-citation network is displayed, which, however, shows the most important authors.

Lijphart is the most prominent academic in the field, who has an outstanding citation record, and whose work connects many different scholars in the field through co-citations. Lijphart's ideas on consociational democracy provide the central theoretical framework to which almost all authors in the research field refer to a greater or lesser extent. His "Patterns of Democracy – Government forms and performance in thirty-six countries" from 2012 alone reached more than 10,000 Google Scholar citations in August 2019 (Google 2019). Close to the Lijphartian center, we find several smaller thematic clusters driven by other academic stars, most notably Donald Horowitz and James Fearon, who both have produced pioneering studies on ethnic conflict and its regulation. They are the connecting points to another important co-citation network consisting of Ted Gurr, Benedict Anderson and Michael Hechter, among others. Another highly interconnected co-citation network consists of Barbara Walter, Caroline Hartzell or Brendan O'Leary, which also deals thematically with majority-minority conflicts and post-conflict institutions. Other networks, such as those around Daniel Elazar or Giovanni Sartori, are far less substantial. On the basis of these groundbreaking works, many others have decisively contributed to the research on autonomous regions. They constitute the many, more or less connected nodes of the intellectual network.

The importance of individual scholars can be measured with the h-index, which takes into account their citation impact as well as their overall productivity.⁷ Figure 3 gives an illustrative overview of the 2019 h-index of some of the most important scientists in autonomy studies.⁸

7 The index can be calculated as follows: $h\text{-index}(f) = \max_i \min(f(i), i)$ with f being the number of citations for each publication (Hirsch 2005).

8 Since the data was automatically extracted from Google Scholar, the h-index can only be used for authors with a Google Scholar profile.

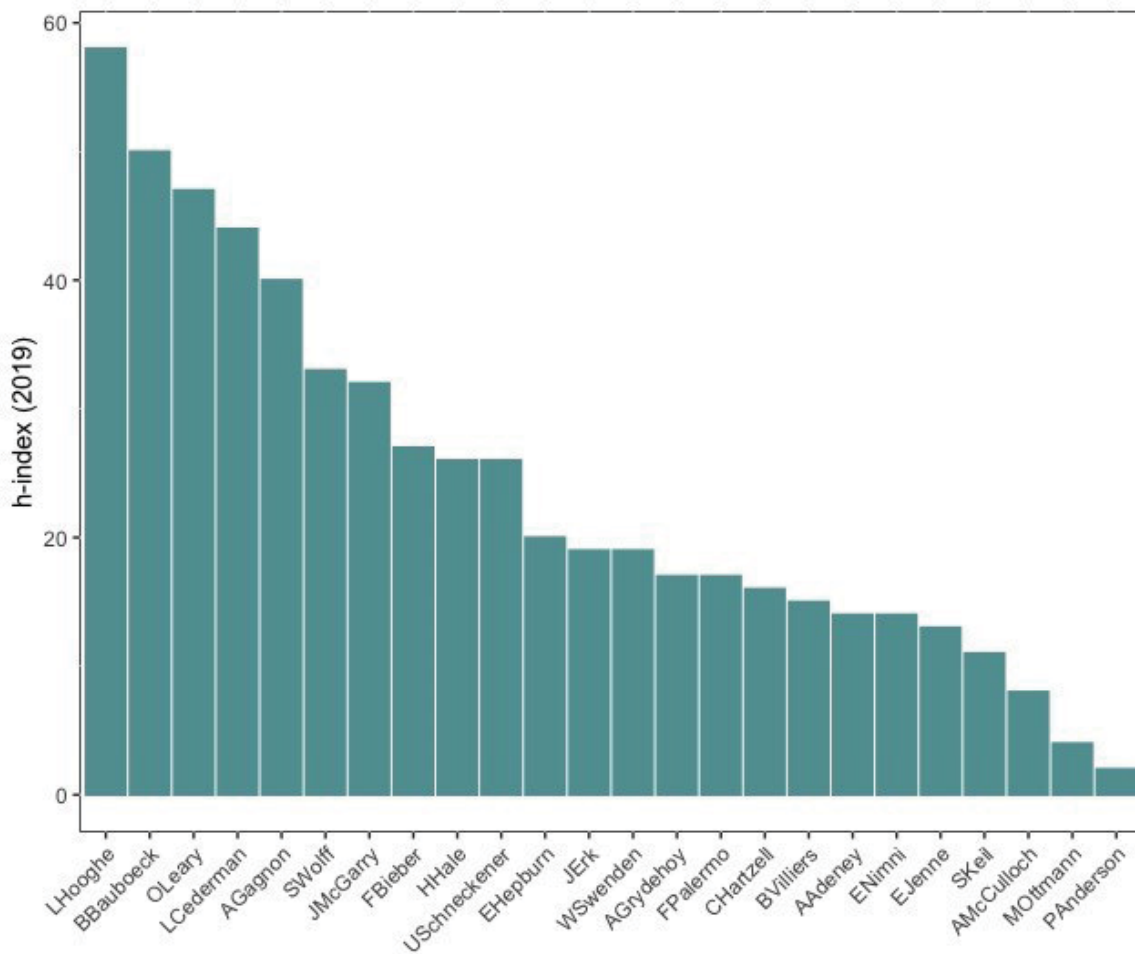


Figure 4: H-index scores (2019)

4. Empirical cases and methods

The research interests in autonomy studies are relatively balanced. Primarily empirical, hypothesis-testing studies form a small majority compared to primarily theoretical and inductive studies. However, the proportion of primarily theoretical works is relatively high at about 30 percent, which has not significantly changed in recent years. Obviously, there is still relatively much conceptual work going on in the research field. What are the necessary and sufficient definition criteria for territorial autonomy? How can different sub-types be distinguished? Which cases fall under a specific definition? At the same time, de jure analyses and questions of international law also play a central role in autonomy studies when it comes to legal guarantees for autonomous regions or the transferred level of powers. About a third of the works contain both inductive and deductive elements and thus genuinely extend the borders of the field. About a quarter of the studies define dependent and independent variables and investigate causal relationships in an x- or y-centered

manner. In contrast, studies that are primarily methodologically oriented are practically irrelevant. The implementation and application of new empirical methods obviously takes place in more established fields. Figure 5 gives an overview of the development of research interests in recent years.

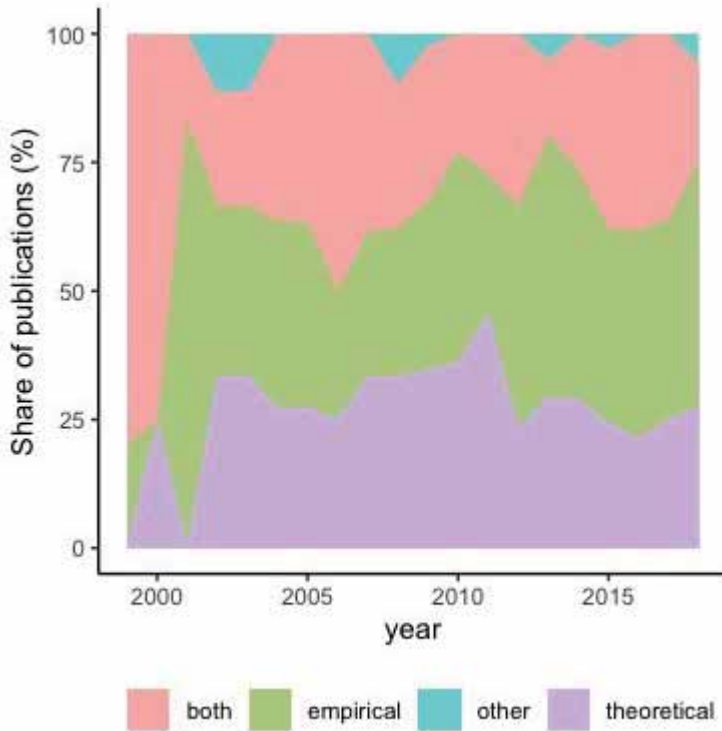


Figure 5: Research interests over time

A closer look at the research designs reveals that authors with a primarily empirical research interest mostly use qualitative methods on the basis of a small number of cases. Autonomy research is highly case-oriented and is more focused on the lower steps of the “ladder of abstraction” (Sartori 1970). Nearly 60 percent primarily study a single case. About 11 percent of articles are medium-n studies with a case number between 2 and 6 cases. The rest of the publications analyze 6 or more cases, with 18 percent reaching the number of 30 or more cases necessary for statistical analysis. While the methodologies of some rely on within-case process tracing, variable-based cross-case comparison or congruence analysis, a broad majority answers their research question in a descriptive and more illustrative manner. While these studies do often not develop new empirical hypotheses, these publications, such as John McGarry and Brendan O’Leary’s work on Northern Ireland or Rajat Ganguly’s on South Asian autonomies constitute the central repository of knowledge of autonomy research and reach relatively high citation scores (McGarry and O’Leary 2004; Ganguly 2013). 16.3 percent of autonomy studies use

descriptive statistics or regression-based analyses. Interestingly, almost all large-n studies belong to broader peace and conflict research that tends to look at autonomy arrangements as an independent variable for the (re-) occurrence of violent intra-state conflicts.

Statistical	Set-theoretical	Comparative	Single case	Mixed	Others	No Method
16.39 %	0.81 %	13.16 %	37.85 %	0.61 %	2.63 %	28.74 %

Table 1: Empirical methods in autonomy studies

Studies that use cross-case research designs to compare a medium number of cases on the basis of theoretically selected variables are relatively rare, with a total share of just over 13 percent. Such designs are able to provide for deep insights and to take the usually high diversity of autonomous sub-state regions into account. At the same time, since they reach a certain degree of generalizability of empirical findings, there seems to be some untapped methodological potential here. The same also applies to set-theoretical methods, which are hardly applied so far. Very few researchers use mixed-method designs and combine qualitative and quantitative methods to answer their research question. This may be a direct consequence of the low level of author networks and the resulting smaller repertoire of methods mastered by individual scientists. Here, too, an uptrend would be desirable in order to compensate for the disadvantages of methodological approaches and to increase the robustness of empirical findings. The proportion of publications that could not be assigned to a specific method type in coding process is comparatively high (28.74 percent). This again reflects the high proportion of purely theoretical studies that represent conceptual or legal papers and do not test empirical hypotheses in a strict sense.

The main focus of autonomy research is on sub-national regions in Western, Northern and Southern Europe, post-Soviet Eurasia, and the Balkan states, where a total of 57 percent of the cases are located (Figure 6). Sub-state regions in other parts of the world are studied far less frequently. While just under 10 percent of selected cases are to be found in Latin America/the Caribbean, the share is 8 percent in Southeast Asia and only 7 percent in sub-Saharan Africa. Cases of regional autonomy in East Asia, South Asia, the Middle East and North Africa, as well as North America and Oceania, each reach shares of only between 2 and 4 percent.

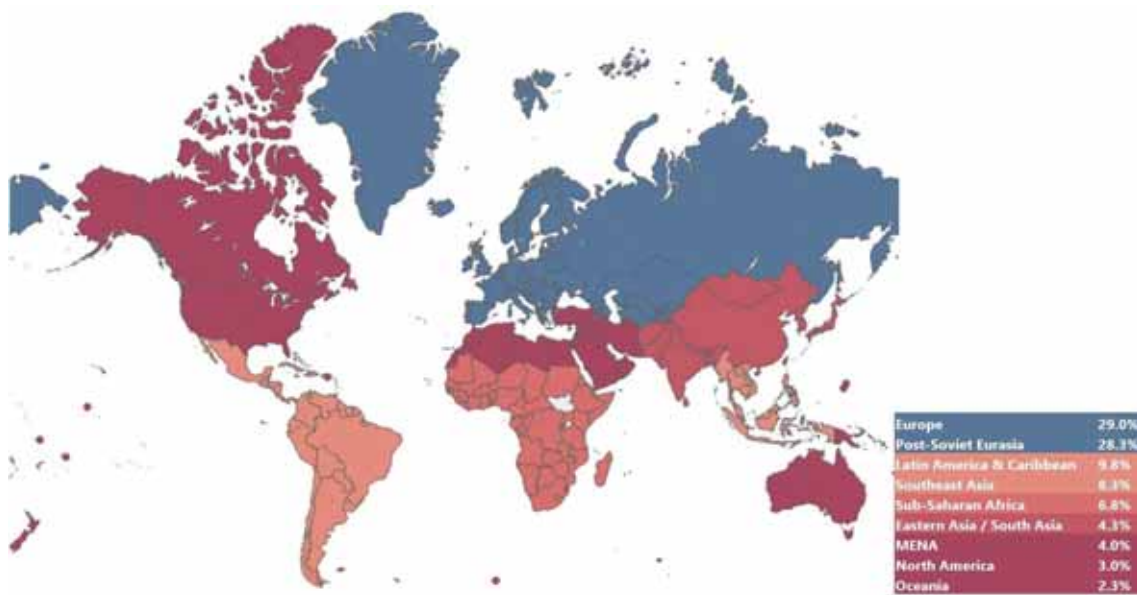


Figure 6: World map on regions of cases

Empirical research has shown that territorial self-governance is unequally distributed, with some world regions being significantly more decentralized than others (Trinn and Schulte 2019). While we find a disproportionately high number of autonomies in post-communist states, territorial self-government is indeed an empirically rather rare phenomenon in the Middle East, North America, or Oceania. A comparison with the Ethnic Regional Autonomy data set (ERAD, Panov and Semenov 2018) reveals that the current empirical focus is only partially related to the actual occurrence of forms of territorial self-government (Table 2). There is a clear misfit between empirically existing and scientifically studied cases when looking at Europe, sub-Saharan Africa or South Asia. While European autonomies are disproportionately often investigated compared to their actual number, South Asian and African cases are clearly underrepresented in the literature.

	Europe	Eurasia	L.-America	S-E-Asia	Sub-S.Africa	E-Asia	S-Asia	MENA	N.America	Oceania
Investigated (%)	29.0	28.3	9.8	8.3	6.8	4.3	4.3	4.0	3.0	2.3
ERAD (%)	13.7	28.1	5.0	8.6	19.4	3.6	16.5	0.7	1.4	2.9

Table 2: Investigated and existing cases in autonomy studies

It seems that autonomy scholars tend to study autonomous territories in the world region from which they themselves originate and of which they have a good case knowledge. Fisher's exact test, which assesses the significance of the association of two nominal variables, shows that the variables *origin* of first authors and the geographical location of the investigated cases (*cases_region*) are not independent of each other ($p=0.0004998$). Since most authors come from Western countries, sub-national entities in this part of the world are more frequently investigated. Selecting cases based on personal preferences and regional focus is a pragmatic and conventional practice and, at the same time, in-depth case-knowledge is indispensable. However, a too strong (eurocentric) bias on case selection can become problematic in the long term, especially if an external perspective is completely absent and even large-n analyses exclude non-European cases. Hence, a more systematic and theoretically meaningful case-selection is desirable.

77 percent of publications in TALD2 stem from authors who currently are employed at an institution in a country which is home to three or more politically relevant ethnic groups.⁹ A third of the home countries were affected by one or more ethnic self-determination movements since the Second World War. While this even constitutes a (weak) correlation between the proportion of authors from a country and the number of self-determination movements in that country, there is no evidence of a "personally-affected"-bias in autonomy research. Deviant cases such as China, India, or Russia, each with a high number of relevant ethnic groups and a correspondingly high number of self-determination movements, make up only a small proportion of the authors in TALD2. As already mentioned, high proportions are held by states such as the USA, Great Britain, and Canada, which are home to ethnic minorities and have experienced self-determination movements, but which are above all rich democracies with high educational expenditures and corresponding research opportunities and freedoms.

5. Topics

As argued above, research on territorial autonomies is a cross-cutting issue that is studied by many scientific disciplines. A Journal Co-Citation Analysis (JCA) demonstrates that autonomous regions as research subjects have made it into all major flagship journals of political science.¹⁰ Articles with the most citations have been published in the American

9 According to the Ethnic Power Relations Data Set (EPR) an ethnic group is politically relevant "if at least one significant political actor claims to represent the interests of that group in the national political arena, or if members of an ethnic category are systematically and intentionally discriminated against in the domain of public politics" (Lars-Erik Cederman, Brian Min, and Andreas Wimmer 2009)

10 The unit of analysis for co-citation analysis can also be scientific journals. JCA detects clusters of journals in which co-cited papers are published.

Political Science Review (48), the American Journal of Political Science (32), World Politics (38), the British Journal of Political Science (23), or the Journal of Democracy (22). A graphic illustration reveals several sub-clusters. The Journals African Affairs, Governance, as well as Publius - The Journal of Federalism, among others, form citation networks which have only limited connections to the dominant general-interest journals of the discipline.

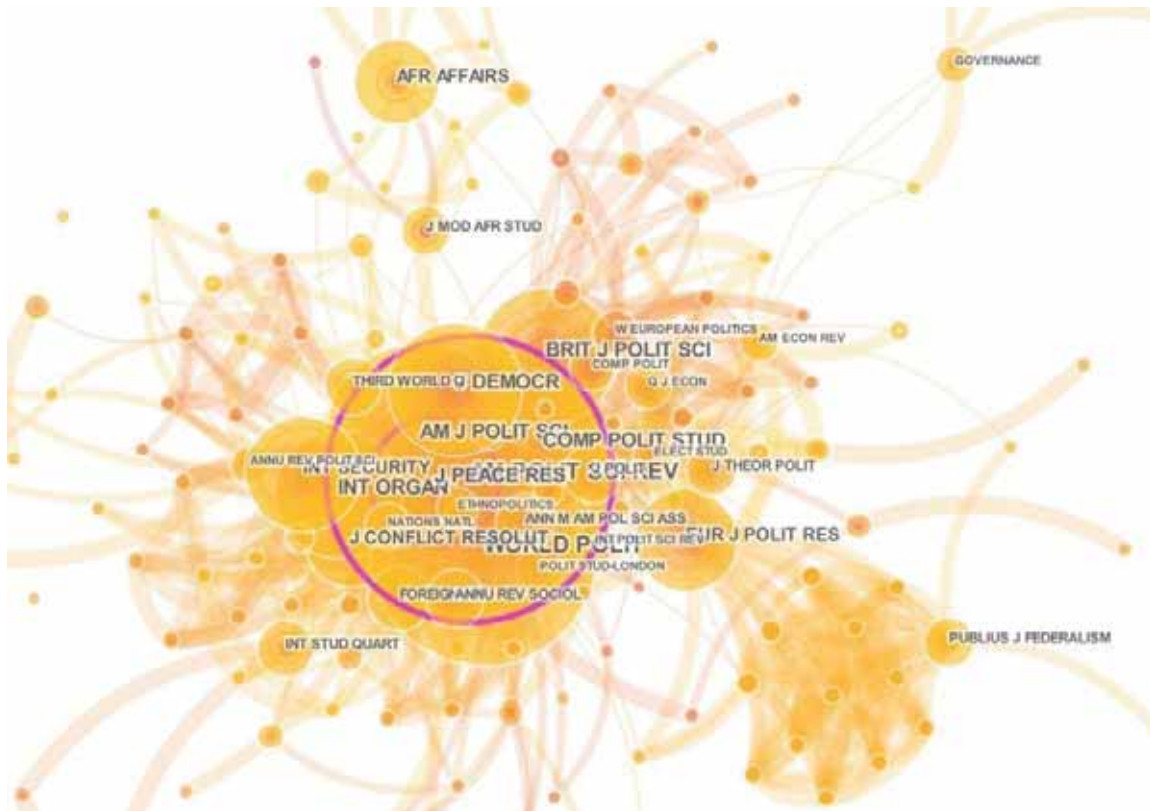


Figure 7: Journal co-citation network

This indicates a diverse and sub-field spanning usage of autonomy as a concept in political science research. Autonomy is included as a dependent or independent variable, as a broad or narrow concept, subsumed under different definition criteria, and studied in various contexts. To further test this assumption, all keywords assigned by the authors are collected and divided into subgroups based on their similarity using a clustering algorithm.¹¹ The results are displayed in a dendrogram (Figure 7).

The dendrogram seems to illustrate a high thematic diversity of the articles included in TALD2. Thematic agglomerations include clusters of *Europe/European Union/*

¹¹ The R package cluster in Version 2.1.0 is used for analysis (Rousseeuw et al. 2019)

Europeanization (blue), of Region/Regionalism/Regional identity (green) and Language/ Language Rights/ Language Policy (brown). A closer look, however, reveals that most clusters are thematically relatively close and revolve around a specific higher-order topic, namely multiculturalism and politics in deeply divided societies. Despite the rather broad application of territorial autonomy as an empirical concept, the research field follows a very clear thematic direction, which is addressed by almost all publications covered by TALD and, as outlined above, largely influenced by Lijphart's work as a key driving force. A simple frequency analysis of the article abstracts supports this assumption. Among the 10 most frequent nouns in the article abstracts, the terms *state*, *minority*, *government*, and *conflict* are found (Figure 9).

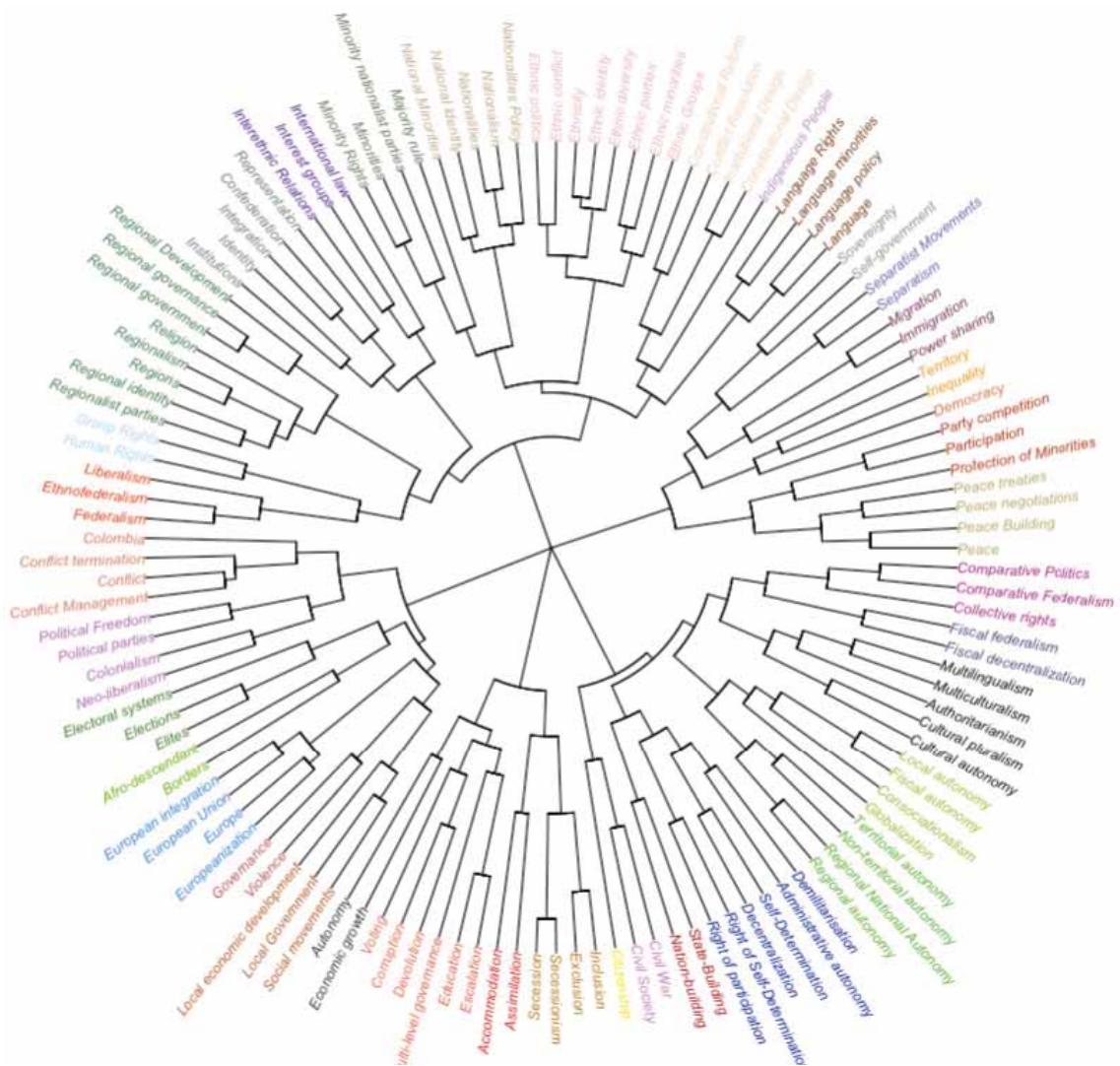


Figure 8: Dendrogram of keyword clustering

Ethnic, regional, national or *cultural* are among the most often used adjectives. The most often occurring noun phrases, text sequences comprising an adjective and a noun, are *territorial autonomy* and *regional autonomy*, followed by *ethnic group*, *civil war* and *international law*.¹²

Frequency analyses are a suitable, yet imperfect way to assess the main topic in a research field, as it may be the case that authors assign different keywords to thematically similar articles or select keywords on a strategic basis. An alternative way is to determine the content of research texts by automatically generating keywords from them. Therefore, the RAKE algorithm (Rapid Automatic Keyword Extraction) is applied to all article abstracts included in TALD2. RAKE produces keywords by analyzing contiguous word sequences and excluding irrelevant terms. Each candidate keyword or parts of such gets a score based on its frequency and co-occurrence frequency in the texts.

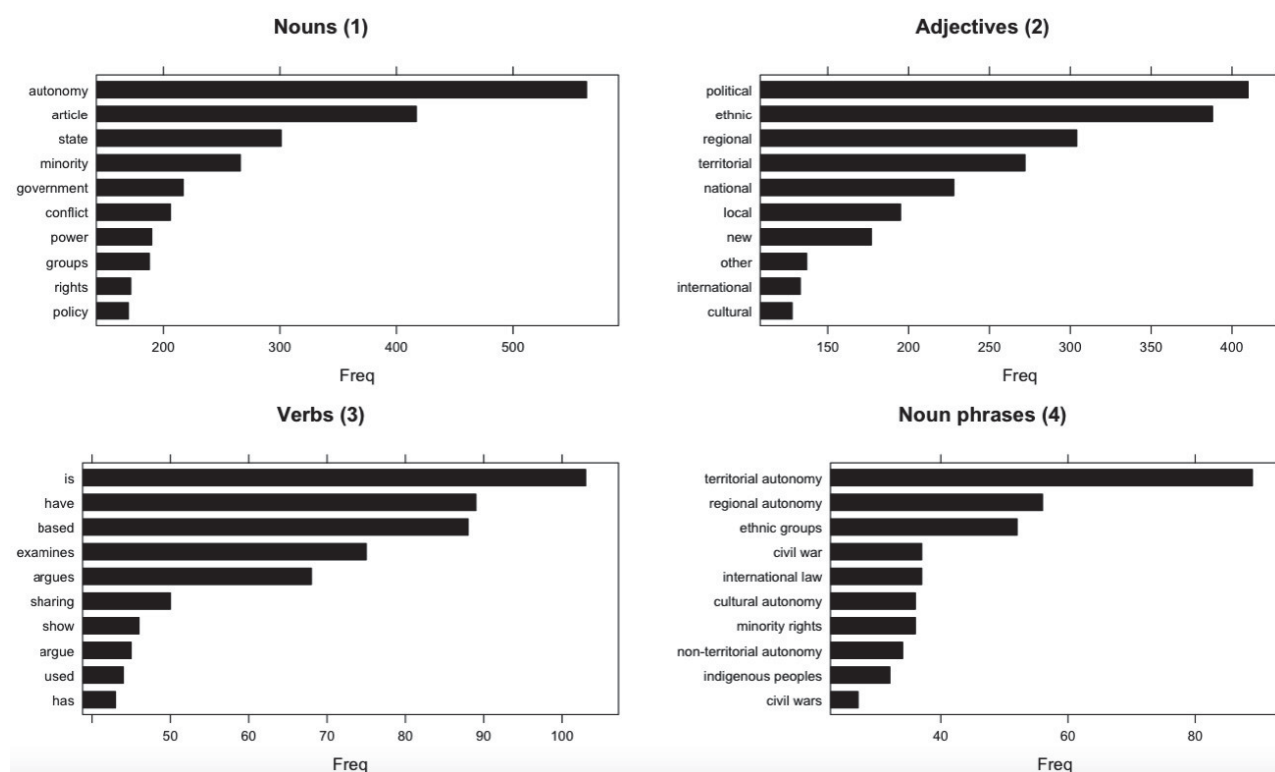


Figure 9: Frequency analysis

12 The natural language processing R toolkit UDPipe in Version 0.83 is used for analysis (Wijffels, Straka, and Straková 2019).

The RAKE score is then the score of each of the words which define the keyword. (Smith, Phil, and Jones 2017). The list of the 15 keywords with the highest RAKE score contains more or less exclusively terms that relate very closely to the aforementioned topic of minority protection and constitutional or institutional engineering in multi-ethnic contexts. In addition to the automatically generated keywords *civil war*, *armed conflict* or *ethnic conflict*, the terms *international law* or *institutional design* also receive high RAKE scores. The equally frequent occurrence of definitory terms (e.g. *cultural autonomy*, *local autonomy* or *local government*) once again makes it clear that the conceptual debates in the field have not yet ended and scholars tend to use different terminologies when studying autonomous regions.

By extracting the vocabulary of the abstracts, we can analyze the correlations between word pairs to see how strongly two terms are linked within a sentence in an abstract.¹³ The word pairs “self-determination” and “civil war”, but also “power-sharing”, reach high co-occurrence scores (cooc). The latter, as argued above, is a derivative from consociational theory, and is frequently used to describe territorial autonomy arrangements (Wolff 2013; Graham, Miller, and Strøm 2017; Cederman, et al. 2015). Highly correlated word pairs illustrate the basic cornerstones of the thematic map of autonomy studies: their primary focus is on autonomy solutions in the context of intrastate conflicts (“*civil war(s)*”) in deeply divided societies between the majority population and regional minorities (“*center – periphery*”) over self-determination (“*self – determination*”) and minority rights, which are often caused by marginalization (“*deprivation – rebellion*”; “*income – inequality*”) assimilation or majority dominance.

Keyword	Frequency	RAKE
civil war	31	2.589
ethnic minority	26	2.233
ethnic group	42	2.170
international law	33	2.134
national identity	17	2.124
institutional design	16	2.080
armed conflict	17	2.068
regional autonomy	41	2.050
local autonomy	21	2.044
ethnic conflict	25	2.026
non-territorial autonomy	27	2.002
territorial autonomy	72	1.986
minority rights	23	1.981
local government	24	1.971
cultural autonomy	27	1.899

Table 3: Automatically generated keywords

¹³ Excluded are trivial word pairs such as “GDP per capita” or “20th century.”

Finally, this raises the question of whether the thematic focus of publications has changed over time, whether new topics have been addressed, and whether the field of research, like many others in political science, has become more diverse and fragmented. For this purpose, the sample is split in two: a sample of texts published between 2000 and 2010, and a second sample of articles published between the beginning of 2010 and the end of 2018.¹⁴ In order to compare the contents of the publications in both investigation periods I again count word co-occurrences and explore the frequency of nouns and adjectives in the paper abstracts. In order to test the robustness of the analysis, it is also examined for both periods which words follow each other how often. Both analyses can be graphically illustrated using a word network, whereby the thickness of the lines indicates the relative frequency of a word pair (Figure 10 and 11).

Both word networks show a very similar picture. It provides no indication that the thematic orientation of the publications has changed substantially in the past decades, nor is there any evidence of any greater thematic fragmentation over time. While the top six co-occurrences are identical in both time periods, this also applies to almost 70 percent of word sequences. This is quite remarkable in spite of the short period of investigation, especially since “unequal” sequences are quite similar in terms of content, as e.g. *ethnic group* instead of *minority group*, etc. Despite the rather sparse cooperation networks, the degree of thematic fragmentation is surprisingly low. In this respect, autonomy research appears as a rather coherent and small intellectual world.

¹⁴ Since before 2000 substantially fewer articles were published and therefore no comparable sub-samples can be compiled, the analysis is limited to a rather short investigation period of 18 years.

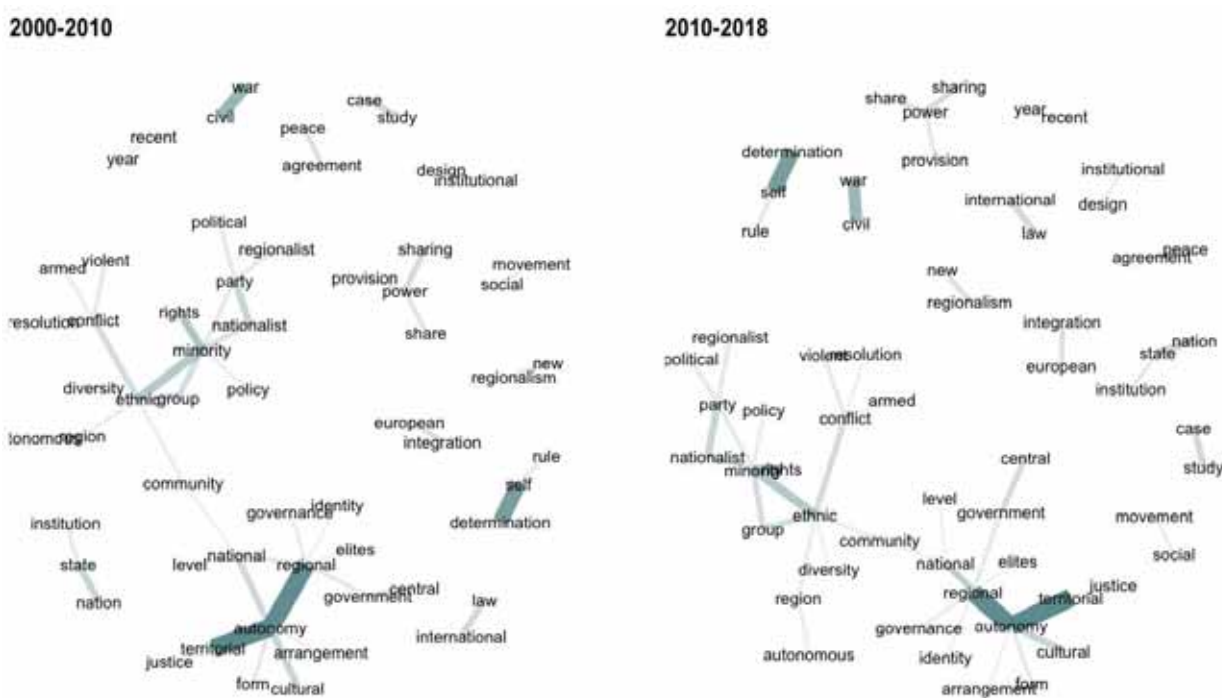


Figure 10: Word co-occurrences (2000–2010 and 2010–2018)

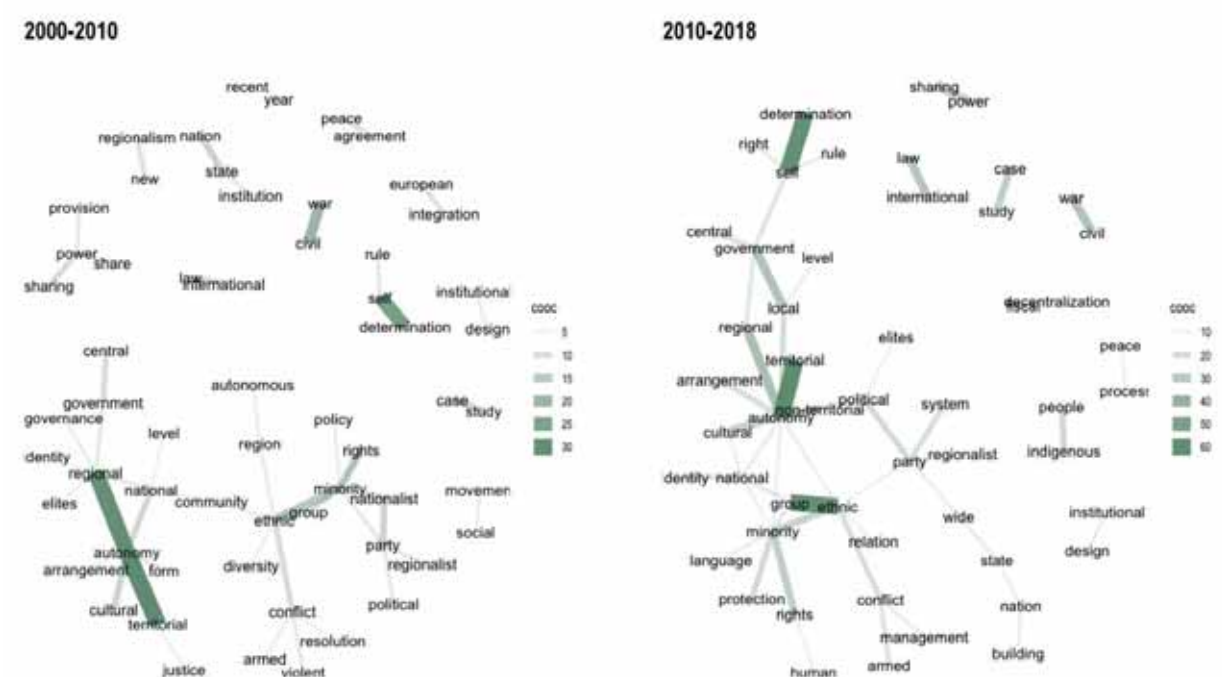


Figure 11: Words following one another (2000–2010 and 2010–2018)

Term A	Term B	Correlation
self-	determination	0.757
civil	war	0.697
turnout	variation	0.566
constituent	unit	0.529
power	sharing	0.469
center	periphery	0.449
geographical	distance	0.432
deprivation	rebellion	0.423
immigrant	citizenship	0.402
education	school	0.389
native	titular	0.386
civil	wars	0.353
income	inequality	0.348

Table 4: Word co-occurrence scores

6. Conclusion

Scientific progress presupposes a certain degree of self-reflection. This requires discussing not only cutting-edge findings and theoretical controversies, but also to critically debate methodological developments and trends in collaboration and publication. As far the author is aware, this research note provides the first comprehensive bibliometric analysis in the field of autonomy studies. It has introduced the Territorial Autonomy Literature Data Sets, TALD1 and TALD2, which include several hundred peer-reviewed and non-peer-reviewed articles from 1945 to 2018, and which provide us with the necessary data to discuss the thematic, as well as methodological directions and to detect fruitful avenues for further research.

The analysis has shown that since the 1980s, autonomy studies have developed into a highly productive and dynamic field of research. Territorial autonomy is a widely used concept in political science and has entered the flagship journals of the discipline. Based on the seminal works by Hurst Hannum, Ruth Lapidot, Stefan Wolff, John McGarry, and Brendan O’Leary, many have collected and analyzed examples of autonomy arrangements. The field is very much dominated by single-authored and small-n studies. Most autonomy scholars focus on the in-depth analysis of single cases or a very small number of sub-state regions in Europe, post-Soviet Eurasia, and the Balkans. There is evidence that the case selection is heavily influenced by the author’s country of origin. While some progress towards greater diversity can be observed in recent years, research on territorial autonomy is still dominated by male and Western-based political scientists. With very few exceptions, the research on regional autonomy concentrates on minority rights and diversity management and is heavily guided by Lijphart’s ideas on democracy in deeply

divided places. Consociationalism provides the fundamental theoretical basis for almost all scholars in the field, whether they are comparativists, peace researchers, or scholars who study autonomous regions from an international law or philosophical perspective. However, the rather vague conceptualization of the autonomy concept in the original consociational theory seems to have led to a certain conceptual ambiguity, demonstrated by prominent word creations such as 'complex power-sharing' (Wolff) or 'territorial power-sharing' (Cederman et al. or Hartzell and Hoddie) for sub-national entities with a special status. From an ontological point of view, further research should draw a clearer line between forms of integrative power-sharing and forms of power-dividing, as well as combinatory institutional arrangements.

While a frequently recurring methodological discussion revolves around the increased prevalence of mixed-method research designs in the social sciences, autonomy studies are still primarily oriented towards case-based qualitative methods and hardly rely on multi-method approaches. This may be due to the fact that it is a comparatively young field which deals with a truly cross-cutting issue, and which is also often examined from less empirically oriented legal or philosophical perspectives. As mentioned above, ongoing conceptual debates make up a significant part of the literature.

The research field attracts more and more authors who produce an increasing number of publications. Authors have become more diverse in recent years when it comes to the proportion of women in first authorship or the proportion of non-Western authors. Interestingly, there has been no significant diversification in methodologies, research interests, or topics covered. From a methodological point of view, a trend towards more cross-case comparisons or quasi-experimental research designs would be particularly beneficial to address the trade-off between generalizability and in-depth case analysis. The combination of qualitative case studies and regression-based analyses also offers untapped potential for further developing research on autonomous regions. More attention should be paid to detailing the case selection procedures in small-n research articles, which are not fully documented in many publications. A large number of articles study 'typical cases' when testing theoretical arguments, which are often the well-studied European cases such as Åland, South Tyrol, Northern Ireland, or Spain's autonomous regions. Including less considered cases of territorial autonomy in Asia or Sub-Saharan Africa would benefit the robustness and generalizability of empirical findings. I, therefore, encourage the development of stronger international and interdisciplinary collaboration networks in autonomy research.

Bibliography

- Achouri, Cyrus (2013). *Kultur und Autonomie. Ein Beitrag zur orientalisches-abendländischen Kulturverständigung*. Wiesbaden.
- Åkermark Sia S. and Sarah Stephan (2013). Rethinking Territorial Arrangements in Conflict Resolution – Introduction to the Special Issue of the International Journal on Minority and Group Rights. In: *International Journal on Minority and Group Rights* 20.1
- Arzheimer, Kai and Harald Schoen (2009). Isoliert oder gut vernetzt? Eine vergleichende Exploration der Publikationspraxis in der PVS”. In: *Politische Vierteljahresschrift* 50.3.
- Bednar, Jenna (2011). The Political Science of Federalism. In: *Annual Review of Law and Social Science* 7.1.
- Brancati, Dawn (2008). *Peace by Design. Managing Intrastate Conflict through Decentralization*. Oxford.
- Cancela, João, Thierry Dias Coelho, and João Pedro Ruivo (2014). Mapping Political Research in Portugal: Scientific Articles in National Academic Journals (2000–2012). In: *European Political Science* 13.4.
- Cederman, Lars-Erik, Kristian Skrede Gleditsch, and Julian Wucherpfennig (2017). “Predicting the decline of ethnic civil war: was Gurr right and for the right reasons?” In: *Journal of Peace Research* 54.2.
- Cederman, Lars-Erik, Simon Hug, Andreas Schädel, and Julian Wucherpfennig (2015). Territorial Autonomy in the Shadow of Conflict: Too Little, Too Late? In: *American Political Science Review* 109.2.
- Cederman, Lars-Erik, Brian Min, and Andreas Wimmer (2009). Ethnic Power Relations Dataset. In: *Harvard Dataverse Network*, url: <https://dataverse.harvard.edu>
- Cederman, Lars-Erik, Gleditsch Kristian Skrede, and Julian Wucherpfennig (2018). The Diffusion of Inclusion: An Open-Polity Model of Ethnic Power Sharing. In: *Comparative Political Studies* 51.10.
- Chen, Chaomei (2018). *CiteSpace - Visualizing Patterns and Trends in Scientific Literature*. url: <http://cluster.cis.drexel.edu/>.
- Dinstein, Yoram (1981). *Models of Autonomy*. Tel Aviv.
- Fisher, Bonnie S. et al. (1998). How many authors does it take to publish an article? Trends and patterns in political science. In: *PS - Political Science and Politics* 31.4.
- Gagnon, Alain-G.; Keating, Michael (2012). *Political Autonomy and Divided Societies: Imagining Democratic Alternatives*. New York.
- Ganguly, Rajat. (2013). *Autonomy and ethnic conflict in South and South-East Asia*. London.
- Gerolemou, M. (2017). *Thinking of Autonomy as Automatism: The Case of Autonomy in Thucydides’ History*. In: *Araucaria* 19: 199–211.
- Google (2019). *Google Scholar* url: <https://scholar.google.com>.
- Goyal, Nihit and Michael Howlett (2018). Lessons Learned and Not Learned: Bibliometric Analysis of Policy Learning. In: *Learning in Public Policy*. Cham.

- Graham, Benjamin, Michael K Miller, and Kaare W Strøm (2017). Safeguarding Democracy: Power-sharing and Democratic Survival. In: *American Political Science Review* 111.4.
- Hannum, Hurst (1996). *Autonomy, Sovereignty, and Self-Determination: The Accommodation of Conflicting Rights*. Philadelphia.
- Hartzell, Caroline A and Matthew Hoddie (2015). The Art of the Possible: Power Sharing and Post-Civil War Democracy. In: *World Politics* 67.1.
- Hirsch, Jorge E. (2005). An index to quantify an individual's scientific research output. In: *Proceedings of the National Academy of Sciences of the United States of America* 102.46.
- Lapidoth, Ruth (1997). *Autonomy. Flexible Solutions to Ethnic Conflicts*. Washington, D.C.
- Leifeld, Philip and Karin Ingold (2016). Co-authorship Networks in Swiss Political Research. In: *Swiss Political Science Review* 22.2.
- Leifeld, Philip, Sandra Wankmüller, Valentin Berger, Karin Ingold, and Christiane Steiner (2017). Collaboration patterns in the German political science co-authorship network. In: *PLOS ONE* 12.4.
- Lijphart, Arend (1999). *Patterns of Democracy. Government Forms and Performance in Thirty-Six Countries*. New Haven-London.
- Malloy, Tove H, Alexander Osipov, and Vizi Balázs (2015). *Managing Diversity through Non-Territorial Autonomy. Assessing Advantages, Deficiencies, and Risks*. Oxford.
- McGarry, John and Brendan. O'Leary (2004). *The Northern Ireland Conflict: Consociational Engagements*. Oxford.
- Metz, Thomas and Sebastian Jäckle (2017). Patterns of Publishing in Political Science Journals: An Overview of Our Profession Using Bibliographic Data and a Co-Authorship Network. In: *PS: Political Science & Politics* 50.1.
- Osipov, Alexander (2018). Can 'Non-territorial Autonomy' Serve as an Analytical Term? Between 'Thick' and 'Thin' Approaches. In: *International Journal on Minority and Group Rights* 25.4.
- Panov, Petr and Andrei Semenov (2018). The World of Ethnic Regional Autonomies: Introducing the New Dataset. In: *Nationalism and Ethnic Politics* 24.4.
- Pehl, Malte (2012). The study of politics in Germany: A bibliometric analysis of subfields and methods. In: *European Political Science* 11.1.
- Pelke, Lars and Paul Friesen (2019). Democratization Articles Dataset: an introduction. In: *Democratization* 26.1.
- Rousseuw, Peter et al. (2019). *Package cluster*. url: <https://svn.r-project.org/R-packages/trunk/cluster/>.
- Sambanis, Nicholas et al. (2018). SDM: A New Data Set on Self-determination Movements with an Application to the Reputational Theory of Conflict. In: *Journal of Conflict Resolution* 62.3.
- Sartori, Giovanni (1970). Concept Misformation in Comparative Politics. In: *American Political Science Review* 64.4.
- Smith, Andrew, Maintainer Phil, and Mike Jones (2017). Package rakeR. url: <https://philmikejones.github.io/rakeR/>.

- Sophocles (1672). *The Three Theban Plays. Antigone/Oedipus the King/Oedipus at Colonus*. Trans. by Bernard MacGregor Walker Knox Robert Fagles. 1984
- Springer Rudolf (Renner), Karl) (1902). *Der Kampf der österreichischen Nationen um den Staat*. Wien/ Leipzig.
- Strøm, Kaare W., Scott Gates, Benjamin Graham, and Håvard Strand (2017). Inclusion, Dispersion, and Constraint: Power sharing in the World's States, 1975–2010. In: *British Journal of Political Science* 47.1.
- Szöcsik, Edina and Christina I. Zuber (2015). EPAC - A New Dataset on Ethnonationalism in Party Competition in 22 European Democracies, In: *Party Politics* 21.01
- Taylor, James S. (2019). *Autonomy*. url: <https://www.britannica.com/topic/autonomy>.
- Trinn, Christoph und Schulte, Felix (2019): Untangling Autonomy – Introducing the Territorial Self-Governance Dataset (TERRGO). Paper presented at IPSA June 2019, Sarajevo University.
- Wang, Binglu, Yi Bu, and Win-bin Huang (2018). Document- and Keyword-based Author Co-citation Analysis. In: *Data and Information Management* 2.2.
- Web of Science (2019). *Social Science Citation Index* url: <http://wcs.webofknowledge.com>.
- Wijffels, Jan, Milan Straka, and Jana Straková (2019). Package udpipe. url: <https://bnosac.github.io/udpipe/>
- Willett, Joanie and Arianna Giovannini (2014): The Uneven Path of UK Devolution. Top-Down vs. Bottom-Up Regionalism - Cornwall and the North-East Compared. In: *Political Studies* 62.2
- Wimmer, Andreas, Lars-Erik Cederman, and Brian Min (2009). Ethnic Politics and Armed Conflict: A Configurational Analysis of a New Global Data Set. In: *American Sociological Review* 74.2.
- Wolff, Stefan (2013). Conflict Management in Divided Societies: The Many Uses of Territorial Self-Governance. In: *International Journal on Minority and Group Rights* 20.1.
- Wolff, Stefan and Christalla Yakinthou, eds. (2012). *Conflict Management in Divided Societies: Theories and Practice*. Abingdon.

Appendix

Variable	Description
origin1	The origin of the first author coded by region. (1) Northern, Western and Southern Europe (2) Post-Soviet Eurasia (3) North America (United States of America, Canada) (4) Latin America and the Caribbean (5) East Asia (China, Taiwan, Korea, Japan) (6) Southeast Asia (ASEAN 10, Timor) (7) South Asia (incl. Afghanistan) (8) Middle East and North Africa (incl. Israel, Turkey) (9) Sub-Saharan Africa (incl. South Africa) (10) Oceania (AUS, NZ, PNG etc.) (99) missing
origin2-5	The origin of the second/third/fourth/fifth author coded by region
number_aut	number of individuals listed as article authors
gender1	The gender of the first author (1) male (2) female (99) missing
gender2-5	The gender of the second/third/fourth/fifth author
Home	Country of home institution (university, research institute)
Heterog1	Number of relevant ethnic groups in country of home institution (EPR)
Heterog2	Country of home institution of first author was affected by violent self-Determination movement since 1945 (Sambanis et al. 2017) (1) Yes (0) No
Heterog3	Number of ethnic autonomies in country of home institutions (Panov/Semenov 2017)
Interest	Primary research interest of article (1) theoretical or legal The research agenda primarily focuses on theoretical problems or legal issues without empirical tests (2) empirical The research agenda primarily focuses on one or more empirical phenomena. No original theory is developed. (3) theoretical & empirical The research agenda focuses on both theory or conceptual design as well as testing the theoretical arguments through empirical analysis. (4) methodological The article primarily focuses on methodological questions or the implementation and application of methods. (99) other primary research interest.

Causal claim	<p>Causal claim</p> <p>(1) Descriptive: No attempt is made to identify causal relationships between independent and dependent variable(s).</p> <p>(2) Causal: The article aims to identify causal relationships between independent and dependent variable(s). (99) missing</p>
Method	<p>(1) statistical analysis (descriptive or causal)</p> <p>(2) set theoretic methods (including QCA)</p> <p>(3) comparative case studies</p> <p>(4) single case study</p> <p>(5) mixed methods design</p> <p>(6) others</p> <p>(7) no method</p>
Cases_region	<p>The world region of focus included in the empirical study.</p> <p>(1) Northern, West and Southern Europe</p> <p>(2) Post-Soviet Eurasia</p> <p>(3) North America (United States of America, Canada)</p> <p>(4) Latin America and the Caribbean</p> <p>(5) Eastern Asia (China, Taiwan, Koreas, Japan)</p> <p>(6) Southeast Asia (ASEAN 10, Timor)</p> <p>(7) South Asia (incl. Afghanistan)</p> <p>(8) Middle East and North Africa (incl. Israel, Turkey)</p> <p>(9) Sub-Sahara Africa (incl. South Africa)</p> <p>(10) Oceania (AUS, NZ, PNG etc.)</p> <p>(11) Interregional (Study includes two or more world regions)</p> <p>(99) missing</p>
Number_cases	<p>Number of countries included.</p> <p>(1) number of cases 1</p> <p>(2) number of cases 2</p> <p>(3) number of cases 3-5</p> <p>(4) number of cases 6-30</p> <p>(5) number of cases 31+</p> <p>(99) missing</p>
Keyword1-5	Keywords assigned by author(s)
Abstract	Article abstract

Table 5: Codebook for TALD2