Article

A Lexical Index of Electoral Democracy

Comparative Political Studies
2015, Vol. 48(12) 1491–1525

DOI: 10.1177/0010414015581050
cps.sagepub.com
SAGE

Svend-Erik Skaaning[1], John Gerring[2], and Henrikas Bartusevičius[1]

Abstract

Recent years have seen an efflorescence of work focused on the definition and operationalization of democracy. One debate concerns whether democracy is best measured by binary or graded scales. Critics of binary indices point out at that they are overly reductionist, while defenders counter that the different levels of graded measures are not associated with a specific set of conditions. Against this backdrop, we propose to operationalize electoral democracy as a series of necessary and sufficient conditions arrayed in an ordinal scale. The resulting "lexical" index of electoral democracy, based partly on new data, covers all independent countries of the world from 1800 to 2013. It incorporates binary coding of its subcomponents, which are aggregated into an ordinal scale using a cumulative logic. In this fashion, we arrive at an index that performs a classificatory function—each level identifies a unique and theoretically meaningful regime-type—as well as a discriminating function.

Keywords

political regimes, democratization, regime change, democracy index, measurement

Recent years have seen an efflorescence of work focused on the definition and operationalization of democracy. The concept serves as an ongoing touchstone in methodological discussions of concept formation (e.g., Goertz,

[1]Aarhus University, Denmark
[2]Boston University, MA, USA

Corresponding Author:
Svend-Erik Skaaning, Department of Political Science, Aarhus University, Bartholins Allé 7, 8000 Aarhus C, Denmark.
Email: skaaning@ps.au.dk

2006; Schedler, 2012; Seawright & Collier, 2014) and new democracy indices continually appear, which are periodically reviewed and critiqued (e.g., Armstrong, 2011; Coppedge et al., 2011; Gleditsch & Ward, 1997; Knutsen, 2010; Munck, 2009; Vermillion, 2006).

One way to categorize this growing corpus of indicators is by the type of scale used to measure the key concept (democracy)—binary, ordinal, or interval. Binary indices include the democracy–dictatorship ("DD") index produced by Przeworski and collaborators (Alvarez, Cheibub, Limongi, & Przeworski, 1996; Cheibub, Gandhi, & Vreeland, 2010) and an index produced by Boix, Miller, and Rosato (2013, hereafter "BMR"). Ordinal measures include the Political Rights ("PR") index and the Civil Liberty ("CL") index, both produced by Freedom House (2013), along with the Polity2 index drawn from the Polity IV database (Marshall, Gurr, & Jaggers, 2013). Interval measures include the Index of Democracy produced by Vanhanen (2000), the Contestation and Inclusiveness indices produced by Coppedge, Alvarez, and Maldonado (2008), and the Unified Democracy Scores ("UDS") produced by Pemstein, Meserve, and Melton (2010).

There is much more to a democracy index than its choice of scale. Even so, scaling is a critical issue in measurement and one that has garnered considerable controversy, especially as concerns the virtues and vices of binary measures (cf. Cheibub et al., 2010; Elkins, 2000). Critics of binary indices point out their reductionist elements: All features of a regime must be reduced to a single coding decision, producing binary sets that are highly heterogeneous and borderline cases that may not fit neatly into either category. Binary indices, by construction, lack discriminating power. Defenders counter that if the definition of these binary sets is properly grounded in theory, the two-part typology may succeed in identifying—from the multifarious elements of democracy—that condition, or set of conditions, that serves a crucial role in political life (see Collier & Adcock, 1999). However, this is not an easy claim to sustain—witness the proliferation of binary indices that identify different defining conditions of democracy.[1]

We take for granted that different sorts of scales are useful for different purposes. Our aim is thus not to subsume or replace extant measures of democracy. The discipline is well served by a variety of measures for this central concept. Instead, we propose a new method of scale construction that combines the differentiation of an ordinal scale with the distinct categories of a typology.

Specifically, we propose to operationalize electoral democracy as a series of necessary-and-sufficient conditions arrayed in an ordinal scale. We refer to this scaling procedure as "lexical." The resulting lexical index of electoral democracy, partly based on novel data construction, covers all independent

countries of the world from 1800 to 2013 and is thus the most comprehensive measure of democracy currently available.[2] It incorporates binary coding based on factual characteristics of regimes and in this way avoids the problem of subjective judgments by coders and the "mashup" quality of nonbinary indices (Ravallion, 2011). However, each binary coding is aggregated together using the cumulative logic of a lexical scale with seven levels. In this fashion, we arrive at an index that performs a classificatory function—each level identifies a unique regime-type—as well as a discriminating function. This approach to measurement offers theoretical and empirical advantages over other methods of representing the complex concept of electoral democracy that may be useful in certain settings.

The first section of the article "Discrimination Versus Meaningful Categories" shows that extant data sets of democracy fall short in simultaneously providing fine-grained discriminatory power and meaningful categories. The second section "Developing a Lexical Index" develops conditions that define our lexical index of electoral democracy. The third section "Coding" discusses how this index is coded through history and across the universe of independent states. The fourth section "Validity" deals with the anticipated validity of the coding. The fifth section "Distribution of Regime-Types Over Time" explores features of the lexical index, which is compared with extant indices in the sixth section "Contrasts With Extant Indices." The seventh section "The Lexical Index at Work: Democracy and State Repression" applies the new measure to the question of state repression, showing how its fixed meanings to the different levels inform the interpretation of statistical relationships in a way that is not accessible through conventional democracy indices. The eight section "Extensions" offers additional thoughts on the application of the lexical index to causal questions pertaining to democracy. The "Conclusions" section contains a summary of the value that a lexical approach to measurement may add to our understanding of electoral democracy.

Discrimination Versus Meaningful Categories

The Freedom House indices recognize seven categories each, and the Polity2 index 21 categories.[3] In contrast to binary indices, the levels in these ordinal indices are not qualitatively different from each other. A "3" on the PR, CL, or Polity2 scale signifies that a polity is more democratic than a country coded as "2" but it does not identify specific traits that distinguish polities falling into each of these categories. Extant ordinal indices, thus, identify countries with more or less democracy but not different kinds of democracy. In this respect, they resemble interval scales.

Interval indices of democracy are generally second-order indices. That is, they are constructed by aggregating together information provided by other democratic indices—through factor analysis (Contestation and Inclusiveness) or Bayesian latent variable models (UDS). The exception is Vanhanen's Index of Democracy. However, the distribution of data on this index is so highly skewed and so evidently censored—nearly 50% of the observations are at the zero point of a 100-point scale—that it loses discriminatory power.[4] Thus, our discussion of interval measures focuses on the Contestation, Inclusiveness, and UDS indices.

The purpose of a well-constructed interval index is to identify fine distinctions among entities. The Contestation, Inclusiveness, and UDS indices achieve this goal as well as can be expected. However, the goal of reducing the plenitude of characteristics associated with "democracy" to a single unidimensional index is elusive. It is elusive because the concept itself is multidimensional and because extant indicators are limited in their purview (Coppedge et al., 2011). An appropriate response is to define the resulting index in a carefully delineated way—as representing only one dimension of a multifaceted concept. Thus, Coppedge et al. (2008) describe one component from their principal components analysis as Contestation and the other as Inclusiveness. The UDS is simply described as a measure of democracy. However, these are ex post descriptions resulting from a rather ad hoc process, putting together myriad indices—whose definition and construction is often ambiguous and lacks justification—with a statistical model and labeling the central tendencies resulting from that model as "X." It is unclear, for example, whether an index labeled "Inclusiveness" includes all measures relevant to that concept and no measures irrelevant to that concept and whether the included elements are aggregated in an appropriate way. Aggregation techniques are virtually limitless, given that researchers must make many choices in the construction of a factor analysis or Bayesian latent variable model.[5]

If all indices are in some sense arbitrary, they are arbitrary in strikingly different ways. The arbitrariness of a binary scale lies in the choice of necessary condition(s) that define the two categories. The arbitrariness of an ordinal or interval scale lies in the choice of indicators to include as elements of the index and the choice of aggregation method to combine those indicators into a single index.

If all indices are informative, they are informative in strikingly different ways. The information contained in a binary index is classificatory, that is, it groups polities in a fashion that is (arguably) theoretically and empirically fecund. The information contained in an interval index is discriminating, that is, it identifies small differences between polities that allow us to distinguish

the degree to which they possess the core attribute of interest. Ordinal scales occupy a middle position in this respect. However, extant ordinal indices of democracy perform neither task very well, for reasons explained above.

Developing a Lexical Index

The core meaning of democracy is rule by the people; on this there is little dispute. One theory of democracy, which can be traced back to E. E. Schattschneider (1942) and Joseph Schumpeter (1950), among others, proposes that the mechanism by which people exert control over political decisions is electoral. Citizens are empowered to rule through competitive elections, which allow them to select leaders and discipline those leaders, establishing relationships of responsiveness and accountability. By electoral democracy, therefore, we mean a regime where leaders are selected through contested elections held periodically before a broad electorate.

Our proposed index of democracy focuses explicitly on this electoral model of democracy, sometimes referred to as a competitive, elite, minimalist, procedural, realist, "thin," or Schumpeterian conception of democracy (Møller & Skaaning, 2013; Przeworski, Alvarez, Cheibub, & Limongi, 2000; Schumpeter, 1950). We are not concerned with other aspects of democracy such as civil liberties, rule of law, constraints on executive power, deliberation, or nonelectoral mechanisms of participation. Electoral refers to elections, tout court.

As such, our definition of the topic is somewhat narrower than definitions of democracy adopted by most extant indices. This is especially true for indices that assume an ordinal and interval scale (e.g., PR, CL, Polity2, Contestation, Inclusiveness, UDS), which tend to range widely, including a broad range of features associated with the concept of democracy. This important definitional contrast should be highlighted from the outset, as it affects everything that follows.

A lexical approach to measurement is concept-driven (Gerring, Skaaning, & Pemstein, 2014). Thus, we begin with a survey of attributes associated with the key concept, electoral democracy, as defined. In identifying attributes for possible inclusion in our index, we are mindful of the vast literature on this topic, with special attention to linguistic studies of the concept (e.g., Held, 2006; Lively, 1975; Naess, Christophersen, & Kvalo, 1956) and foundational works in the electoral tradition (listed above).

To form a lexical scale one must arrange attributes so that each serves as a necessary-and-sufficient condition within an ordered scale. That is, each successive level is comprised of an additional condition, which defines the scale in a cumulative fashion. Condition *A* is necessary and sufficient for *L1*,

conditions *A&B* are necessary and sufficient for *L2*, and so forth. In achieving these desiderata, four criteria must be satisfied: (a) binary values for each condition, (b) unidimensionality, (c) qualitative differences, and (d) centrality or dependence (see Gerring et al., 2014).

First, each level in the scale must be measurable in a binary fashion without recourse to arbitrary distinctions. It is either satisfied or it is not. To be sure, the construction of a binary condition may be the product of a set of necessary and/or sufficient conditions. Collectively, however, these conditions must be regarded as necessary and sufficient.

Second, levels in a lexical scale must be understood as elements of a single latent (unobserved) concept. Conceptual multidimensionality must be eliminated, either by dropping the offending attribute and/or by refining the concept in a clearer and perhaps more restrictive fashion, as we have in moving from "democracy" to "electoral democracy."

Third, each level must demarcate a distinct step or threshold in a concept, not simply a matter of degrees. Levels in a lexical scale are intended to identify qualitative differences. A "3" on a lexical scale is not simply a midway station between "2" and "4." Indeed, each level may be viewed as a subtype of the larger concept. In this respect, the lexical index is reminiscent of "diminished subtypes" of democracy (Collier & Levitsky, 1997; Merkel, 2004). However, while subtypes revolve in a radial fashion around a central concept—possessing all the attributes of the ideal-type except one—the lexical index is more akin to a classical concept, where new concepts are created by cumulative combinations of attributes—*A, A&B, A&B&C*, and so forth.

The most challenging aspect of lexical scale construction is the ordering of attributes, which follows a conceptual, rather than empirical, logic. One attribute may be considered prior to another if it is more *central* to the concept of theoretical interest (from some theoretical vantage point). This follows a constitutive approach to measurement, where attributes are the defining elements of a concept (Goertz, 2006). Alternatively, one attribute may be considered prior if it is a logical, functional, or causal prerequisite of another. The *dependence* of *B* on *A* is what mandates that *A* assume a lower level on a scale. Whether responding to considerations of centrality or dependence, the levels of a lexical scale bear an asymmetric relationship to each other; some are more fundamental than others. This is the most distinctive feature of lexical scaling.[6]

Based on these considerations, we arrive at a lexical index of electoral democracy with six conditions and seven levels, as follows:

L0: No elections.
L1: No-party or one-party elections.

L2: Multiparty elections for legislature.
L3: Multiparty elections for legislature and executive.
L4: Minimally competitive, multiparty elections for legislature and executive.
L5: Minimally competitive, multiparty elections with full male or female suffrage for legislature and executive.
L6: Minimally competitive, multiparty elections with universal suffrage for legislature and executive.

Further elaboration of this minimalist approach to electoral democracy can easily be envisioned. For example, one might try to measure aspects of electoral integrity such as high respect for political liberties (see Howard & Roessler, 2006; Møller & Skaaning, 2013). For present purposes, we restrict ourselves to what might be considered the most basic properties of electoral democracy. Happily, these properties are also the most easily measured, allowing for an index that stretches back in time and across all independent states. The point is that the index proposed here is not the only index of electoral democracy that might be constructed. We trust that other approaches—either more detailed or more concise—would nonetheless be consistent with the judgments incorporated into this index, as discussed below.

Importantly, to qualify as an election (Condition 1), the electorate may be quite small but must be separable from, and much larger than, the group of officials it is charged with selecting. Examples include South Africa under Apartheid and virtually all national elections in Europe and the Americas during the 19th century. However, the selection of a king by a legislature or estates general, typical of the *Ständestaat* (Poggi, 1978), would not qualify, as the electorate is infinitesimal as a share of the citizenry (whom for present purposes we shall understand as permanent residents in whatever territory is claimed as sovereign), and difficult to distinguish from the chosen monarch because they both share royal blood and may all claim the title. "Indirect" elections count as elections unless there are multiple steps in between the electorate and the chosen representative(s), as in China today and Uganda in the 1970s. It follows that leadership positions filled through a one-stage electoral college (e.g., U.S. presidents, chosen by an electoral college, or prime ministers chosen by an elected parliament, who serve as an electoral college) are considered elective offices.

Having laid out the index, we now explore its rationale in relation to the four criteria presented above. The first criterion is that each condition be coded in a binary fashion (0/1). This criterion does little violence to reality as most of the conditions are naturally dichotomous. The exception is suffrage, a continuous variable. Note, however, that our understanding of an election

presumes an electorate that is considerably larger than the body it selects and separable from that body. An "election" where 0.0001% of citizens qualify for the vote would not qualify as an election under our definition. In the event, one does not find any modern examples of multiparty elections for national offices in independent countries where less than 5% of the electorate can vote. After the Reform Act of 1832, demarcating the introduction of significant contestation in England, more than 650,000 males—approximately 9% of the adult population—had the right to vote (Phillips & Wetherell, 1995). In the United States, around 60% to 70% of adult White men could vote by 1790 (Keyssar, 2009). Arguably, this feature of the historical record reflects a functional relationship. If the electorate is miniscule, there is less need for an electoral process by which to choose leaders and establish a relationship of accountability, and even if there is a perceived need it will be difficult to establish and maintain multiparty elections with a miniscule electorate (Gerring, Palmer, Teorell, & Zarecki, 2015). At the other end of the spectrum, nearly universal suffrage elections (where just a few, small categories of voters are excluded) are understood as universal, and are not, in any case, a stable category. Once suffrage has been granted to nearly all men or nearly all women, it becomes very difficult—and also rather pointless—to maintain the barrier. Again, there seems to be a functional logic at work. Thus, we find that in polities with competitive elections but without universal male or female suffrage, female suffrage is usually 0 and male suffrage is generally between 20% and 60% of the adult male population. By setting the bar for *L5* at 100%, we are thus comparing full (or nearly full) male suffrage with partial male suffrage.

The second criterion concerns unidimensionality, a feature that informs any index. The main challenge to this objective lies in the twin principles of inclusion and contestation, often regarded as constituting separate dimensions of electoral democracy (Dahl, 1971). Empirically, there is no question that these elements are distinct (Coppedge et al., 2008). Countries with high inclusion (as measured, for example, by suffrage rights) may have very low contestation, or none at all. However, the lexical index is theoretically driven rather than empirically driven. Our claim is that once a minimal level of inclusion has been attained—sufficient to constitute an electorate and hence the precondition for an election, as discussed—further increases in suffrage are irrelevant unless and until elections are competitive. This argument will be taken up below. For the moment, we note that the claim to unidimensionality is deductive rather than inductive.

The third criterion concerns qualitative differences across the identified levels. We want to claim that there is a degree of coherence to each category such that they can be considered as meaningful regime-types. That is to say,

members of each category constitute a set that shares additional (unmeasured) characteristics. This claim is addressed in §4, where we connect lexical types with research drawn from the literature on democratization. Relatedly, we suppose that each step in the index is consequential, at least for some outcomes. This claim is taken up (for one particular outcome) in §7.

The final criterion concerns the ordering of attributes into a lexical scale according to centrality or dependence. Recall that this is the most important and controversial aspect of lexical scaling, and its application to the concept of electoral democracy is by no means self-evident. We need to carefully explain and justify our choices.

The existence of elections is judged fundamental (*conditio sine qua non*), as other attributes associated with electoral democracy make no sense outside of an electoral context (Collier & Adcock, 1999; Merkel, 2004). Country *A* is not more of an electoral democracy than Country *B* if neither polity holds elections, regardless of what other characteristics[7] those polities might possess. Likewise, some attributes depend on other attributes in a logical manner. Specifically, an electoral regime is a necessary condition of multiparty elections, and multiparty elections are necessary conditions of competitive elections. Moreover, a regime in which both legislature and executive are elective is arguably more democratic than a regime in which only one of these offices is elective. These features of the lexical index may be regarded as self-evident.

Some of the attributes of democracy depend for their meaning on other attributes in a functional manner. The most important of these involve the relationship of inclusion and contestation, referenced above. So long as the size of an electorate is nontrivial, we regard the extent of suffrage as irrelevant to electoral democracy unless and until elections count for something. The reasoning behind this assessment returns us to the electoral theory of democracy, according to which citizens are empowered through an electoral connection. To establish relationships of responsiveness and accountability between officials and the citizenry, the electoral theory suggests that it is essential for political offices to be elective, for citizens do the selecting, for there to be more than one choice, and for choices occur at regular intervals (introducing the threat of electoral punishment). If these elements are not present, the right of suffrage is meaningless and apt to serve as a tool of elite control rather than one of democratic accountability. This logic is apparent in classic theoretical work in the electoral tradition (e.g., Dahl, 1971; Przeworski et al., 2000; Schattschneider, 1942; Schumpeter, 1950) and is ratified by recent empirical work (reviewed in Gandhi & Lust-Okar, 2009).

To gain an intuitive sense for our prioritization of competitiveness over inclusion, let us consider several examples. We begin with electoral authoritarian regimes, where universal suffrage exists but elections lack multiparty

competition or the most important policymaking offices are nonelective (*L1-L3* in the lexical index). In our view, nothing of consequence distinguishes electoral authoritarian regimes that impose limits on suffrage from those that allow universal suffrage. Soviet-era Russia is not more democratic than prerevolutionary Russia, despite the inauguration of universal male suffrage in 1918. Likewise, if an electoral authoritarian regime like North Korea decided to restrict access to the ballot to certain classes of citizens, it would hardly be any less democratic. Similarly, we regard regimes with minimal competition but restricted suffrage such as Britain during the 19th century as more democratic than, say, present-day Rwanda, which is characterized by universal suffrage but not electoral competition. All of these examples seem to reinforce the notion that competitiveness stands prior to inclusion in the attainment of electoral democracy; the latter is functionally dependent upon the former.

Coding

To code the lexical index, we make use of five variables developed initially in the Political Institutions and Events (PIPE) data set (Przeworski, 2013): LEGSELEC, EXSELEC, OPPOSITION, MALE SUFFRAGE, and FEMALE SUFFRAGE. Because PIPE does not attempt to measure the quality of elections, we generate a sixth variable: COMPETITION. All variables are binary, coded 1 if the following circumstances obtain and 0 otherwise.

LEGSELEC: A legislative body issues at least some laws and does not perform executive functions. The lower house (or unicameral chamber) of the legislature is at least partly elected. The legislature has not been closed.
EXSELEC: The chief executive is either directly or indirectly elected (i.e., chosen by people who have been elected).
OPPOSITION: The lower house (or unicameral chamber) of the legislature is (at least in part) elected by voters facing more than one choice. Specifically, parties are not banned and (a) more than one party is allowed to compete or (b) elections are nonpartisan (i.e., all candidates run without party labels).
MALE SUFFRAGE: Virtually all male citizens are allowed to vote in national elections. Legal restrictions pertaining to age, criminal conviction, incompetence, and local residency are not considered. Informal restrictions such as those obtaining in the American South prior to 1965 are also not considered.[8]
FEMALE SUFFRAGE: Virtually all female citizens are allowed to vote in national elections. Similar coding rules apply.

COMPETITION: The chief executive offices and seats in the effective legislative body are filled by elections characterized by uncertainty (see Przeworski et al., 2000), meaning that the elections are, in principle, sufficiently free to enable the opposition to gain power if they were to attract sufficient support from the electorate. This presumes that control over key executive and legislative offices is determined by elections, the executive and members of the legislature have not been unconstitutionally removed, and the legislature has not been dissolved. With respect to the electoral process, this presumes that the constitutional timing of elections has not been violated (in a more than marginal fashion), nonextremist parties are not banned, opposition candidates are generally free to participate, voters experience little systematic coercion in exercising their electoral choice, and electoral fraud does not determine who wins. With respect to the outcome, this presumes that the declared winner of executive and legislative elections reflects the votes cast by the electorate, as near as can be determined from extant sources. Incumbent turnover (as a result of multiparty elections) is regarded as a strong indicator of competition but is neither necessary nor sufficient.[9] In addition, we rely on reports from outside observers (as reported in books, articles, and country reports) about whether the foregoing conditions have been met in a given election (see Svolik, 2012). Coding for this variable does *not* take into account whether there is a level playing field, whether all contestants gain access to funding and media, whether media coverage is unbiased, whether civil liberties are respected, or other features associated with fully free and fair elections. COMPETITION thus sets a modest threshold.

Although we use PIPE as an initial source for coding LEGSELEC, EXSELEC, OPPOSITION, MALE SUFFRAGE, and FEMALE SUFFRAGE, we deviate from PIPE—based on our reading of country-specific sources—in several ways. First, with respect to executive elections, in the PIPE data set, "Prime ministers are always coded as elected if the legislature is open." However, for our purposes, we need an indicator that also takes into account whether the government is responsible to an elected parliament if the executive is not directly elected—a situation generated by a number of European monarchies prior to World War I, by episodes of international supervision such as Bosnia–Herzegovina in the first years following the civil war, and by some monarchies in the Middle East and elsewhere (e.g., Liechtenstein, Monaco, and Tonga) in the contemporary era. To illustrate, PIPE codes Denmark as having executive elections from 1849 to 1900 although the parliamentary principle was not established until 1901. Before then, the government was accountable to the king. Among the current cases with elected

multiparty legislatures not fulfilling this condition are Jordan and Morocco. To achieve a higher level of concept-measure consistency, we have thus recoded all country-years (based on country-specific accounts) for this variable where our sources suggested doing so.

We also conduct original coding for countries whose coding is incomplete in PIPE and for additional countries such as the German principalities that are not covered in PIPE. In this fashion, we generate a complete data set for all six variables covering all independent countries of the world in the period under study (1800-2013). Whereas the numbers of observations for the PIPE variables range between 14,465 and 15,302, our data set provides 18,142 observations for all variables. Except for minor adjustments regarding executive elections (mentioned above), this additional coding follows the rules laid out in the PIPE codebook. Coding decisions are based on country-specific sources that are too numerous to specify. In rare instances, we stumbled upon information that required a recoding of PIPE variables, so the two data sets do not correspond exactly.

To generate the lexical index from these six binary variables, a country-year is assigned the highest score (*L0*-*L6*) for which it fulfills all requisite criteria, as follows:

L0: LEGSELEC = 0 & EXSELEC = 0.
L1: LEGSELEC = 1 or EXSELEC = 1.
L2: LEGSELEC = 1 & OPPOSITION = 1.
L3: LEGSELEC = 1 & OPPOSITION = 1 & EXSELEC = 1.
L4: LEGSELEC = 1 & OPPOSITION = 1 & EXSELEC = 1 & COMPETITION = 1.
L5: LEGSELEC = 1 & OPPOSITION = 1 & EXSELEC = 1 & COMPETITION = 1 & (MALE SUFFRAGE = 1 or FEMALE SUFFRAGE = 1).[10]
L6: LEGSELEC = 1 & OPPOSITION = 1 & EXSELEC = 1 & COMPETITION = 1 & MALE SUFFRAGE = 1 & FEMALE SUFFRAGE = 1.

Countries are coded across these conditions for the length of their sovereign existence within the 1800-2013 timespan, generating a data set with 221 countries. To identify independent countries, we rely on Gleditsch (2013) and Correlates of War (2011), supplemented from 1800 to 1815 by various country-specific sources. Importantly, electoral democracy does not presume complete sovereignty. A polity may be constrained in its actions by other states, by imperial control (as over a colony), by international treaties, or by world markets. Thus, to say that a polity is an electoral democracy is to say

Table 1. Frequency Distribution, 1800-2013.

Conditions	Satisfied	Lexical index						
		L0	L1	L2	L3	L4	L5	L6
A. Elections	13,584 (75%)	N	Y	Y	Y	Y	Y	Y
B. Multiparty elections for legislature	10,583 (58%)	I	N	Y	Y	Y	Y	Y
C. Multiparty elections for executive	9,080 (50%)	I	I	N	Y	Y	Y	Y
D. Minimally competitive elections	6,287 (35%)	I	I	I	N	Y	Y	Y
E. Full male or female suffrage	11,335 (62%)	I	I	I	I	N	Y	Y
F. Universal suffrage	9,553 (53%)	I	I	I	I	I	N	Y
		4,569 (25%)	2,964 (16%)	1,522 (8%)	2,808 (15%)	929 (5%)	480 (3%)	4,870 (27%)

Numbers represent observations (*N*) and share (%) of total sample (18,142 country-years) satisfying the specified condition(s), rounded to the nearest integer. *N* = No, *Y* = Yes, *I* = Irrelevant.

that it functions as such for policies over which it enjoys decision-making power. Scores for each indicator reflect the status of a country on the last day of the calendar year (31 December) and are not intended to reflect the mean value of an indicator across the previous 364 days.

Evidently, a lexical index reduces the potential property space of the component conditions. Exactly how this works can be seen in Table 1. The first column lists all six conditions, whereas the second column shows the number (*N*) and share (%) of total observations in our data set that meet that criterion. Thus, the first (positive) condition—the existence of elections for either the legislature or executive—is satisfied in 13,584 election-years, constituting 75% of the observations in our data set ($N = 18,142$). The second condition—multiparty elections for the legislature—is satisfied in 10,583 country-years, constituting 58% of our total observations, and so forth.

Coding for the lexical index derives from these six conditions, as indicated in the second section of Table 1. A polity receives a score of 0 if the first condition is not met, that is, there are no elections for either the legislature or the executive. All other conditions are irrelevant. This situation obtains in 4,569 country-years, constituting 25% of our data set, as shown in the bottom row of Table 1. A polity receives a score of 1 if the first condition is met, that

is, there are national elections, but the second condition (multiparty elections for the legislature) is not satisfied. This situation obtains in 2,964 country-years, 16% of the country-years recorded in our data set. The highest (most demanding) score of 6 is accorded to a polity that satisfies all conditions, as shown in the final column of Table 1. This situation obtains in 4,870 country-years, 27% of the total observations in our data set.

In this fashion, any circumstance can be coded unambiguously into the typology. Of course, many attributes are irrelevant for this coding, as noted in Table 1. Specifically, as soon as a condition is not satisfied all higher conditions become irrelevant. If a polity does not allow for multiparty elections, the extent of suffrage is irrelevant, for example. This "deductive" quality is what distinguishes a lexical scale from a Guttman or Mokken scale.

Validity

When contrasted with most continuous measures of democracy, the lexical index is relatively simple, enhancing its transparency and reproducibility. Coding decisions are generally factual in nature, resting on institutional features that require historical knowledge but not subjective judgments on the part of the coder. To be sure, uncertainties are introduced when source material for a country is weak. But we assume that this sort of bias is random rather than systematic (as it might be if coder judgments involved questions of meaning and interpretation). In this respect, the lexical index echoes a feature of most binary indices (e.g., DD and BMR). Indeed, it is quite similar to these indices insofar as it relies on binary codings, which are combined to form a cumulative index.

Another important feature of the coding procedure is its separability from other factors that sometimes confound our ability to measure political institutions. When coding democracy and governance indices—particularly those that assume a continuous distribution—there is a strong possibility that coders may view the state of democracy or governance in Country *X* as inseparable from the general state of affairs in that country, including its economic performance. When things are going well, *X* may receive a higher score. When things are going poorly, it may receive a lower score, even if its political institutions are substantially unchanged (Kurtz & Schrank, 2007). The coding of the lexical index offers little opportunity for this species of measurement error because coding decisions rest on clear-cut thresholds and because the features that are being coded are not amenable to "state of affairs" confounders.

To provide an empirical check on reproducibility we conducted an inter-coder reliability test. By design, one of the authors (H.B.) was not involved in the construction of the index or the original coding of the data set and was

Table 2. Intercoder Reliability.

	Agreement (%)	Cohen's κ	Krippendorff's α
LEGSELEC	93.72	.831	.831
EXSELEC	95.63	.903	.903
OPPOSITION	94.54	.888	.888
MALE SUFFRAGE	95.49	.898	.898
FEMALE SUFFRAGE	96.38	.926	.926
COMPETITION	96.66	.920	.920
LEXICAL INDEX	87.37	.840	.943

Countries (randomly) selected into the intercoder reliability test: Equatorial Guinea, Cameroon, Zambia, Iraq, South Korea, Lebanon, Korea (pre-1910), Hesse Darmstadt (Grand Duchy of Hesse), Solomon Islands, St. Lucia, Malta, Kyrgyzstan, Peru, Trinidad and Tobago, Romania, Israel, Parma, West Germany, Burundi, Marshall Islands, Argentina, and Syria.

not informed of codings arrived at by the other authors or by the PIPE data set. He was then assigned the task of recoding 21 countries (10% of the sample), chosen at random, based on the coding rules presented above and using only country-specific sources (which he chose based upon his review of the extant literature).

Three standard statistical measures of intercoder reliability are presented in Table 2: percent agreement, Cohen's kappa, and Krippendorff's alpha. These are calculated at the variable level (for LEGSELEC, EXSELEC, OPPOSITION, MALE SUFFRAGE, FEMALE SUFFRAGE, and COMPETITION) and at the composite level (for the lexical index). All measures report high levels of intercoder reliability, suggesting that the index is readily reproducible. It is worth noting that this conclusion applies no less to the new competition indicator, although some might consider it to be less reliable because it is less directly observable.[11]

Distribution of Regime-Types Over Time

A frequency distribution of scores across the entire 1800-2013 period is provided in the bottom row of Table 1. It will be seen that the most populated categories are *L0, L1, L3*, and *L6*, whereas others (notably *L5*) have fewer occupants. A fairly high proportion of cases stack up at the two ends of the index, in common with many ordinal and interval indices (Cheibub et al., 2010; Treier & Jackman, 2008).

The distribution of cases changes over time, as one might expect. To get a feel for the application of the lexical index, we provide country scores for the median year in our sample, 1904, as shown in Table 3. At that time, there

Table 3. Frequency Distribution, 1904 ($N = 53$).

0	1	2	3	4	5	6
− National elections	+ National elections	+ Multiparty elections for legislature	+ Multiparty elections for executive	+ Competitive elections	+ Male or female suffrage	+ Universal suffrage
Afghanistan	Colombia	Austria	Argentina	Chile	Belgium	Australia
China	Cuba	Hungary	Bolivia	Costa Rica	Canada	
Ethiopia	Haiti	Bulgaria	Brazil	Denmark	France	
Iran	Honduras	Germany	Dominican Republic	Luxembourg	Greece	
Korea	Liberia	Japan	Ecuador	Netherlands	Panama	
Montenegro	Mexico	Portugal	El Salvador	United	Switzerland	
Nepal	Nicaragua	Romania	Guatemala	Kingdom		
Oman	Paraguay	Spain	Italy	Uruguay		
Ottoman Empire	Venezuela	Sweden	Peru	United States		
Russia			Serbia			
Thailand						

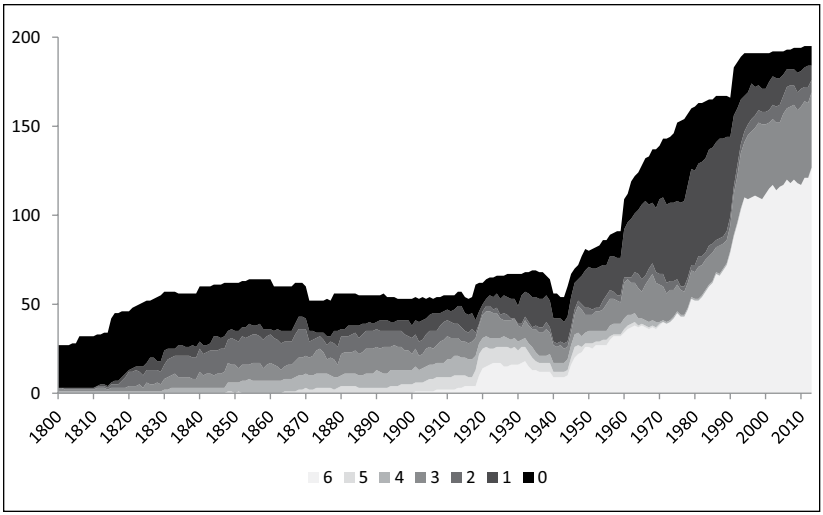

Figure 1. Absolute distribution of political regimes, 1800-2013.
L0 = nonelectoral regimes; *L1* = one- and no-party regimes; *L2* = nonparliamentary constitutional monarchies; *L3* = limited multiparty authoritarian regimes; *L4* = exclusive democracies; *L5* = male democracies; *L6* = electoral democracies.

were 53 independent countries in the world. These were distributed fairly evenly across the seven categories of the lexical index, with the exception of the most democratic category (*L6*), which has only one occupant. Only Australia granted universal suffrage to both men and women, while satisfying the other criteria stipulated in the index. (New Zealand—often considered as the first country to introduce universal suffrage—did not become independent before 1907 according to our criteria.)

A comprehensive picture of change over time is portrayed in a stacked graph of the regime-types across each year, shown in Figure 1. Note that our sample grows over time—from 27 in 1800 to 195 in 2013—due to the appearance of newly sovereign states (e.g., in Africa) and the break-up of sovereign states (e.g., the Soviet Union).

At an aggregate level, Figure 1 highlights those periods in which electoral democracy advanced throughout the world—notably, at the end of World War I, World War II, and the Cold War—as well as those periods in which it declined—notably, the 1930s. The more important feature of this diagram, however, is the disaggregated picture of regime evolution it presents. By decomposing the concept of electoral democracy into constituent parts, we can view changes in membership across regime-types over time.

In 1800, polities were predominantly of Type 0 (no elections), which we call *nonelectoral regimes*. Later in the 19th century, we see the rise of Types 1 to 5 and the concomitant decline of nonelectoral regimes. This is the most diverse period, when no single type is dominant, as illustrated by our snapshot of the world in 1904 (see Table 3).

Over the course of the 20th century, we can see the extraordinary rise of Type 1 (elections without multiparty competition), often referred to as *one- and no-party regimes* (Hadenius & Teorell, 2007). A steep decline for these regime-types begins in the 1980s, coincident with the Third Wave of democratization (Huntington, 1991).

Apart from some transitional regimes, Type 2 regimes (multiparty elections for legislature but not the executive) correspond for the most part to what Therborn (1977) has called *nonparliamentary constitutional monarchies*. This regime-type, widespread in 19th-century Europe, falls into desuetude in the contemporary era, describing just a few polities at the present time.

Type 3 regimes (multiparty executive and legislative elections without real competition), a modestly sized category a century ago, began to grow in the late 20th century to the point where it constitutes today the second-most dominant regime-type. This regime-type is similar to polities described as *electoral, competitive*, or *limited multiparty* authoritarian regimes (Hadenius & Teorell, 2007; Levitsky & Way, 2002; Schedler, 2002). We prefer the latter label because it captures the intension of the concept.

Exclusive democracies and *male democracies*, respectively, have been suggested as plausible labels for Types 4 and 5 (see Collier & Levitsky, 1997; Merkel, 2004). With the growing illegitimacy of suffrage restrictions, these regime-types have become virtually extinct in the 21st century, though they constituted a significant share of all polities prior to World War II.

A final and equally striking pattern in evidence over the past century is the rise of Type 6 regimes, the highest level of our lexical index, corresponding to polities that satisfy all assessed criteria for electoral democracy. This category, largely capturing what democratization scholars have referred to as *electoral democracies* (Diamond, 2002; Møller & Skaaning, 2013), now comprises over half of all polities in the world.

Contrasts With Extant Indices

Table 4 summarizes salient features of the lexical index alongside the nine extant democracy indices introduced at the outset. It will be seen that the lexical index has much broader historical coverage than DD, PR, CL, Contestation, Inclusiveness, and UDS—all of which are focused on the

Table 4. Indices Compared.

	Scale		Coverage			Correlation (with Lexical)	
Index	Type	Range	Countries	Years	Observations	All obs.	Lexical < 6
Lexical (authors)	Lexical	0-6	221	1800-2013	18,142		
DD (Cheibub et al.)	Binary	0-1	197	1946-2008	9,115	.84 (S)	.37 (S)
BMR (Boix et al.)	Binary	0-1	210	1800-2007	16,308	.78 (S)	.45 (S)
Polity2 (Marshall et al.)	Ordinal	−10 to 10	185	1800-2012	16,327	.80 (S)	.59 (S)
PR (Freedom House)	Ordinal	1 to 7	200	1972-2012	7,040	.85 (S)	.42 (S)
CL (Freedom House)	Ordinal	1 to 7	200	1972-2012	7,040	.79 (S)	.31 (S)
Democracy index (Vanhanen)	Interval	0 to 100[a]	192	1810-2010	15,149	.78 (P)	.48 (P)
Contestation (Coppedge et al.)	Interval	−1.84 to 1.96	197	1950-2000	7,534	.91 (P)	.65 (P)
Inclusiveness (Coppedge et al.)	Interval	−3.04 to 1.91	197	1950-2000	7,534	.59 (P)	.54 (P)
UDS (Pemstein et al.)	Interval	−2.11 to 2.26	201	1946-2012	9,850	.87 (P)	.62 (P)

Country counts are based on COW country codes (extended with additional, unique country codes for Orange Free State, Transvaal, Tibet, and United Provinces of Central America, as suggested by Gleditsch), whereas years and observations are taken from the original data sets. DD = Democracy–Dictatorship; S = Spearman's rho; BMR = Boix, Miller, and Rosato; PR = Political Rights; CL = Civil Liberty; P = Pearson's correlation coefficient; UDS = Unified Democracy Scores.

a. Theoretical range.

contemporary era—and slightly better coverage than Polity2, BMR, and Vanhanen's Democracy index.

As is to be expected, the lexical index generally covaries with other indices. For example, it correlates with Polity2 at .80 and with the Political Rights index at .85 (Spearman's rho). However, when the highest scoring cases (Lexical = 6) are dropped from the sample, these correlations drop to .59 and .42, respectively. If we split the sample, distinguishing between years before and after 1900, intercorrelations between the full lexical index and the BMR are .46 for the 19th century and .83 for the 20th century, whereas intercorrelations with Polity2 are .65 and .82. Thus, while the lexical index overlaps with other indices of democracy, it is by no means redundant.

A more detailed look at the relationships between extant binary (DD, BMR) and ordinal (PR, Polity2) indices and the lexical index is portrayed in cross-tabulations in Table 5. This confirms that while various measures of electoral democracy are related, they are not very highly correlated.

One might infer that the lexical index is an outlier among democracy indices. However, a principal components analysis, shown in Table 6, reveals that this is not the case. Again, we find a striking contrast between full sample and partial sample results. In the full sample, 83% of the variance across these 10 indices is explained by the first component. In the partial sample (Lexical < 6), only 52% of the variance can be explained by the first component. However, in neither analysis is the lexical index an outlier, as shown in the eigenvalues.

In elucidating the distinctive features of our lexical index, a useful point of comparison is provided by binary indices. The latter generally combine several of the features identified in our ordinal scale. For example, DD may be said to combine *L1-L4* while BMR combines *L1-L5*, with suffrage understood as a majority of men rather than all men. In doing so, the authors suggest that a polity cannot be called an electoral democracy until it has satisfied a number of conditions—though these conditions do not exactly map onto the condition utilized to score the lexical index, as shown in Table 5.

Our index does not take issue with this determination. However, a lexical approach to scaling suggests that polities that fail to pass all four or five of these conditions may nonetheless be regarded as partial members of the class "electoral democracy." For example, a polity with elections is closer to the electoral ideal than a polity without elections. And it suggests that this distinction—along with others identified along the seven-level index—has consequences, consequences that can be understood as greater/lesser possession of various traits associated with electoral democracy. Thus, rather than insisting that a number of necessary conditions be met, we regard each condition as providing a threshold on a single ordinal scale.

Table 5. Cross-Tabulations.

	Lexical							
	L0	L1	L2	L3	L4	L5	L6	Sum
DD								
0	1,378	1,949	290	1,181	50	2	259	5,109
1	0	10	26	170	101	49	3,587	3,943
Sum	1,378	1,959	316	1,351	151	51	3,846	9,052
BMR								
0	3,796	2,878	1,242	2,410	403	90	183	11,002
1	8	10	43	92	480	390	3,906	4,929
Sum	3,804	2,888	1,285	2,502	883	480	4,089	15,931
PR								
0	421	502	18	113	0	0	0	1,054
1	373	483	49	350	1	0	2	1,258
2	150	202	78	275	15	0	30	750
3	22	54	51	257	4	0	206	594
4	3	16	12	67	0	0	499	597
5	2	0	21	26	0	0	959	1,008
6	0	0	21	1	0	0	1,734	1,756
Sum	971	1,257	250	1,089	20	0	3,430	7,017
Polity2								
0	1,126	123	40	0	0	0	0	1,289
1	256	624	114	126	0	0	4	1,124
2	131	216	77	75	0	1	0	500
3	658	779	212	109	0	0	0	1,758
4	677	222	160	202	0	0	1	1,262
5	120	186	41	205	20	0	9	581
6	58	78	258	197	24	3	1	619
7	245	263	88	426	45	52	15	1,131
8	21	24	34	166	39	1	3	288
9	64	125	51	253	6	4	25	528
10	119	44	11	143	8	9	22	356
11	146	34	76	100	8	7	3	374
12	15	41	44	186	66	22	39	413
13	9	24	8	81	76	14	45	257
14	87	10	28	131	161	8	81	506
15	6	23	11	103	50	19	179	391
16	7	3	26	31	53	22	315	457
17	8	13	1	53	52	37	340	504
18	5	12	8	33	36	46	571	711
19	0	1	1	6	50	56	437	551
20	3	2	0	10	187	176	1,896	2,274
Sum	3,761	2,847	1,289	2,633	881	477	3,986	15,874

PR and Polity2 rescaled so that 0 = lowest value. DD = Democracy–Dictatorship; BMR = Boix, Miller, and Rosato; PR = Political Rights.

Table 6. Principal Components Analysis.

	Full sample				Partial sample			
Component	Eigenvalue	Difference	Proportion	Cumulative	Eigenvalue	Difference	Proportion	Cumulative
1	8.32	7.59	0.83	0.83	5.18	3.85	0.52	0.52
2	0.73	0.40	0.07	0.90	1.33	0.11	0.13	0.65
3	0.33	0.11	0.03	0.94	1.22	0.55	0.12	0.77
4	0.22	0.09	0.02	0.96	0.67	0.24	0.07	0.84
5	0.13	0.04	0.01	0.97	0.43	0.03	0.04	0.88
6	0.10	0.03	0.01	0.98	0.40	0.11	0.04	0.92
7	0.07	0.01	0.01	0.99	0.29	0.04	0.03	0.95
8	0.06	0.03	0.01	1.00	0.25	0.11	0.03	0.98
9	0.03	0.01	0.00	1.00	0.14	0.07	0.01	0.99
10	0.02		0.00	1.0000	0.08		0.01	1.00

EIGENVECTORS

	Full sample		Partial sample	
Variable	Comp1	Unexplained	Comp1	Unexplained
Lexical	0.33	0.10	0.32	0.46
DD	0.31	0.18	0.21	0.78
BMR	0.32	0.14	0.16	0.87
Polity2	0.33	0.09	0.35	0.35
PR	0.33	0.08	0.36	0.34
CL	0.32	0.17	0.32	0.49
Democracy index	0.32	0.17	0.33	0.43
Contestation	0.34	0.03	0.41	0.14
Inclusiveness	0.20	0.66	0.18	0.83
UDS	0.34	0.06	0.41	0.13

Principal factor analysis of democracy indices (unrotated). Number of observations = 4,028 (full sample) and 2,431 (partial sample, where Lexical < *L6*). Factors retained: 1. DD = Democracy–Dictatorship; BMR = Boix, Miller, and Rosato; PR = Political Rights; CL = Civil Liberty; UDS = Unified Democracy Scores.

Clearly, the lexical index allows one to represent more information than is possible in a binary scale. At the same time, the sensitivity of a seven-level ordinal scale is lower than that provided by a longer ordinal scale (e.g., Polity) and much lower than an interval scale (e.g., UDS). In terms of discriminatory ability, the lexical index occupies a midway point.

The advantage of lexical scaling relative to more differentiated ordinal scales or interval scales is in clarity. While the latter are derived from complex models (e.g., UDS) or less formulaic but often opaque weightings across dimensions (e.g., Freedom House and Polity), the lexical value affixed to a country in a particular year is immediately interpretable. We know what a "5" means because there is only one combination of attributes that will yield a score of 5 on a lexical scale.

Likewise, we can understand the categories of the scale as indicating discrete regime-types, which can be tracked through time, as in Figure 1. By way of contrast, longer ordinal scales (e.g., Polity) and interval indices (e.g., UDS) allow one to track the overall trends—more or less democracy through time—by examining changes in the mean over time. But they cannot indicate anything about the specific content (quality) of regimes or about which regime-types expanded or contracted at different points in time. The latter information is both substantively important as well as useful for tracing causal mechanisms, as discussed below.

The Lexical Index at Work: Democracy and State Repression

One purpose of the lexical index of electoral democracy is descriptive: to differentiate regime-types in the world (Table 2) and to portray changes over time (Figure 1). Another use is to probe causal relationships between regime-type and other factors. As an example of this sort of work, we shall explore the relationship between regime-type and state repression of personal integrity rights (Davenport & Armstrong, 2004).

Democracies are expected to be less repressive than autocracies for a variety of reasons. First, a democratic framework is thought to promote tolerance. Second, low respect for human rights may be punished by the electorate at the ballot box. Finally, political participation and contestation provide an outlet for protests and secure legitimacy in the broader population, alleviating the extra-constitutional challenges that often spur violent government repression. Extant theory thus presents a strong prima facie case for political regime-type as an influence on state repression.

However, it is not clear what the precise empirical relationship might be. Extant work on the subject suggests three possible patterns. As summarized

by Davenport and Armstrong (2004: 538–39): (1) "with every step toward democracy, the likelihood of state-related civil peace is enhanced"; (2) "human rights conditions are not only improved when full democracy exists but also when full autocracy is present"; or (3) "there may…be some threshold of domestic democratic peace, below which there is no effect of democracy on repression, but above which a negative influence can be found."

Our interest in this question is heuristic. We probe the empirical relationship between electoral democracy and state repression to demonstrate how the lexical index may be brought to bear on a causal hypothesis where countries are the relevant units of analysis. Specifically, we wish to utilize the special qualities of the lexical index to gain insight into the mechanisms at work in this (putatively) causal relationship.

To simplify things, we adopt the empirical format employed by Davenport and Armstrong (2004), with some minor modifications to update the analysis through 2004.[12] We readily grant that there are other approaches to causal modeling that might be adopted in this instance. However, as our purpose is to compare extant indices—rather than to make causal claims—difficult choices among estimators, specifications, and samples may be put aside.

Following Davenport and Armstrong, state repression is measured by the Political Terror Scale (PTS), in turn based on the State Department human rights country reports (Wood & Gibney, 2010). We enlist Ordinary Least Squared (OLS) regression to assess the model and use a battery of covariates including interstate armed conflict (Uppsala Conflict Data Program [UCDP]/Peace Research Institute Oslo [PRIO]), internal armed conflict (UCDP/PRIO), military dictatorship (Cheibub et al., 2010), population (ln) (Penn World Tables [PWT]), Gross Domestic Product (GDP)/cap. (ln) (PWT), and a one-period lag of the outcome. Democracy is measured in the first instance by the 10-point Polity Democracy index (scaled from 0 to 10) drawn from the Polity IV data set (Marshall et al., 2013).

Our second measure of democracy is the lexical index, with one notable coding change. Data on state repression are available only from 1976, meaning that there is little variation in suffrage laws during the observed period. Distinctions across *L4-L6* of the lexical index are therefore rendered moot, prompting us to collapse them into a single category (*L4*). The resulting index has five levels—*L0-L4*, with roughly equal membership—and is otherwise identical to the index described above.

To probe possible links between regime-type and state repression, we adopt a series of approaches, summarized in Table 7. First, we test the possibility of a linear relationship. Polity (Model 1) and Lexical (Model 2) both indicate a negative relationship: more democracy is correlated with less repression, corroborating the general theory but leaving the problem of causal explanation opaque. Next, we test the possibility of a curvilinear relationship

Table 7. Electoral Democracy as a Predictor of State Repression: Lexical and Polity Compared.

	Linear		Curvilinear		Disaggregated	
	1	2	3	4	5	6
Polity	−.025 (.003)***		.051 (.011)***			
Polity^2			−.009 (.001)***			
Polity, *L1*					.120 (.045)**	
Polity, *L2*					−.067 (.050)	
Polity, *L3*					−.010 (.067)	
Polity, *L4*					.073 (.067)	
Polity, *L5*					−.033 (.058)	
Polity, *L6*					−.011 (.040)	
Polity, *L7*					−.064 (.045)	
Polity, *L8*					−.119 (.037)***	
Polity, *L9*					−.170 (.041)***	
Polity, *L10*					−.400 (.037)***	
Lexical		−.037 (.007)***		−.002 (.029)		
Lexical^2				−.008 (.007)		
Lexical, *L1*						−.132 (.031)***
Lexical, *L2*						−.128 (.060)*
Lexical, *L3*						−.044 (.033)
Lexical, *L4*						−.223 (.031)***
$PTSsd_{t-1}$	.691 (.011)***	.715 (.010)***	.669 (.011)***	.714 (.010)***	.659 (.012)***	.702 (.010)***

(continued)

Table 7. (continued)

	Linear		Curvilinear		Disaggregated	
	1	2	3	4	5	6
Interstate conflict	.057 (.054)	.019 (.053)	.086 (.053)	.021 (.053)	.088 (.053)*	.031 (.053)
Internal conflict	.371 (.026)***	.375 (.025)***	.387 (.026)***	.377 (.025)***	.388 (.026)***	.381 (.025)***
Military dictator	.042 (.027)	.057 (.026)*	.096 (.028)***	.056 (.026)*	.086 (.028)**	.041 (.026)
Population (ln)	.051 (.007)***	.042 (.006)***	.055 (.007)***	.042 (.006)***	.056 (.007)***	.044 (.006)***
GDP/cap. (ln)	−.051 (.009)***	−.073 (.008)***	−.025 (.010)**	−.070 (.008)***	−.022 (.010)*	−.071 (.008)***
Constant	.727 (.095)***	.929 (.090)***	.497 (.100)***	.896 (.094)***	.499 (.101)***	.975 (.095)***
R^2	.748	.770	.752	.770	.754	.772
n	3,553	3,924	3,553	3,924	3,553	3,924

Sample period: 1976-2004. *Countries*: 155/165. *Estimator*: OLS. Standard errors in parentheses.
L = levels on an ordinal scale (not lags). PTS = Political Terror Scale.
*p < .1. **p < .01. ***p < .001 (two-tailed test).
Note: Superscript 2 refers to squared in Table 7.

by introducing a multiplicative term. The coefficients for Polity (Model 3) and Lexical (Model 4) are similar, though only Polity offers support for the notion that democracy's impact on repression is nonlinear.

Finally, we attempt to explore each category of these indices separately through the use of dummy variables representing each level (with the first level omitted as a reference category). Results, shown in Models 5 and 6, are again broadly similar across the two indices, though there are some important differences. The coefficient for *L1* in the polity index is significantly more repressive than the reference category, *L2-L6* do not show results are statistically distinguishable from the null, and *L7-L10* show negative, and statistically significant coefficients. By contrast, *L1, L2*, and *L4* in the lexical index are statistically significant from the reference category, but not *L3*. Additional tests (not reported) show that the differences between *L4*, on one hand, and *L1, L2*, and *L3*, on the other, are significant. The only additional significant difference is that between *L1* and *L3*. We have rerun all the analyses holding sample constant across the parallel models based on Polity and Lexical, respectively. The results (not reported) are virtually identical, meaning that varying coverages do not account for the differences.

Leaving aside for a moment the question of which index offers a truer representation of the relationship between democracy and repression, let us consider what might be learned from Models 5 and 6. Davenport and Armstrong (2004) conclude that "there are important differences between the political systems associated with the highest levels of the Polity measure . . ." (p. 548). This is a reasonable conjecture. But they cannot follow this statement up with any speculation about what is distinctive about the higher levels of the polity index or what might be driving the apparently curvilinear relationship between democracy and repression. This is because the levels of the polity index are not individually interpretable. In this respect, ordinal indices of democracy such as Polity, PR, and CL function very much like interval indices. They inform us about quantities (more or less of some latent trait) but not about qualities (categorical differences across levels).

By contrast, the lexical index provides ample fodder for theorizing because each level defines a discrete category and each category is plausibly approached as a regime-type. Let us begin by reviewing the information contained in Model 6. No level in the lexical index reveals higher levels of state repression than level *L0* (no elections). While it is unsurprising to discover that a nonelectoral state has high levels of repression (for all the reasons set forth in our initial theory), it is somewhat surprising to find that there is no (statistically significant) difference in levels of repression across *L0* and *L3*. If the model is correct, repression decreases significantly when a polity moves (hypothetically) from no national (*L0*) elections to a situation of national elections (*L1*), national multiparty elections for the legislature (*L2*), or—most effectively—minimally competitive elections for the legislature and the executive, while the degree of state repression in a situation of multiparty elections for legislative and executive offices that are not minimally competitive (*L3*) is not significantly different from a situation without national elections.

An explanation may be found in the hybrid nature of the *L3* regimes, which are characterized by many of the constitutional features of democracy without the crucial missing step in which elections are allowed to become competitive. That is, *L3* polities look as if they are democratic—and undoubtedly are portrayed by their leaders as democratic. But even though opposition groups are free to organize and to participate in the political system, they are not allowed to win government power (Schedler, 2002). Some of the hybrid features of this setting are likely to engender more repression than in the other settings characterized by national elections. Because the opposition is free to organize, it is likely to pose a significant challenge to the government. And because the elections are not free, the opposition is likely to pursue extraconstitutional measures, which in turn are likely to provoke government

repression. In short, both government and opposition have means and motive to engage in a cycle of protest and retaliation, a setting that is likely to feed levels of state repression that are indistinguishable from settings without national elections.

Extensions

In this section, we discuss possible applications of the lexical index for understanding causal questions about democracy. To begin, let us reemphasize that the short explanatory sketch offered in the previous section is not intended to convince. To be fully convincing, a causal argument would need to be accompanied by a much longer theoretical discussion intended to make sense of case-based evidence and extant theorizing on this well-trodden subject—not to mention a battery of robustness tests. Our purpose is illustrative. We hope to have shown that a lexical approach to measurement provides a useful tool for gaining insight into causal relationships and specifically into the causal mechanisms that may be at work. This feature derives from the fact that the levels of a lexical scale are individually meaningful.

By contrast, binary scales are generally too diffuse to be useful in this context. Country-years scored as 0 are different from country-years scored as 1 in *many* ways. It is not clear which of these differentiating features might be responsible for a causal effect or whether their impact is combinatorial (a compound treatment). Extant ordinal scales can, in principle, be disaggregated into their component parts, as we have done with the Polity2 index. However, as these components are not uniquely defined, they are not very informative. We know that *L3* is higher than *L2*, but this is about all we know. Interval scales may also be disaggregated. However, establishing the break-points is a highly arbitrary affair, and the resulting categories contain no useful information.

We suspect that the same aspects that render the lexical index useful in the context of state repression might also be useful in the context of other research questions where regime-type lies on the right side of a causal model. Consider the vaunted democratic peace hypothesis (Brown, Lynn-Jones, & Miller, 1996). Although a new scale of democracy will assuredly not solve this obdurate research question by itself, it does allow a more nuanced test of the thesis (at least as pertains to the electoral components of democracy). Specifically, we can explore whether there is a specific level in the lexical index beyond which conflict between nations ceases to occur and whether one or both members of the dyad must surpass that threshold. This is arguably more informative than a binary or ordinal/interval analysis of the problem.

As a second example, one might consider the contested relationship between development and democracy (Przeworski et al., 2000). With democracy on the left side of the model, one may investigate whether the empirical relationship of socioeconomic development to electoral democracy is different at various points in the lexical index. Do increases in per capita GDP have a greater impact on electoral democracy at certain thresholds? With democracy on the right side of the model, one might investigate whether different thresholds of electoral democracy have varying relationships to economic growth. For example, does the initial transition to multiparty elections have a different impact on growth performance than the transition to competitive elections?[13]

Conclusion

The lexical approach to index construction is unique in that each level in the ordinal scale is defined by an additional attribute of the core concept (electoral democracy). These cumulative attributes are assigned according to theoretical expectations rather than empirical distributions, as would be the case for Guttman and Mokken scales. This produces levels that correspond to distinct types. These types are of great value in understanding the progress/regress of democracy around the world, grouping regimes into similar categories, and explaining various outcomes of interest. Note that while any ordinal scale can be disaggregated into categories corresponding to each level, this does not normally reveal groupings that have much in common with each other. (There are many ways to receive a "−6" on the Polity2 index, for example.) In addition, the lexical index has greater country and historical coverage than any extant index of democracy. It may also claim greater precision than most indices by virtue of the largely factual coding criteria (demonstrated by high intercoder reliability) and simple aggregation procedures.

It should be clear that in launching a new index of electoral democracy, we are not proposing that the lexical index has any claim to ontological priority over other sorts of indices, each of which represent certain aspects of reality (while occluding others) and each of which has its uses. Sometimes, relationships are continuous and hence are best measured with an interval scale. Sometimes, they have only one threshold and hence are best measured with a binary scale. Our claim is that, sometimes, descriptive and causal relationships are ordinal in character or require an ordinal scale to test various threshold possibilities. In these settings, which may apply to many theories about democratic development (as cause or effect), a lexical scale may be appropriate. Here, ordinal levels are constructed to represent qualitatively different

categories. These categories are informative insofar as they are fecund, attaining the desiderata of any classificatory scheme, that is, to group phenomena in categories that are mutually exclusive and exhaustive.

Note that the utility of a lexical definition of democracy (like that of all others) rests ultimately on how well it explains the world around us. The electoral interpretation of democracy presumes that one dimension of democracy—grounded in elections—has the greatest impact on governance, well-being, and perhaps also on other aspects of democracy (liberal, deliberative, etc.). It treats the electoral component as causally exogenous. Likewise, our lexical index is premised on a notion of which features of electoral democracy are likely to be most fundamental. On this basis, we included some attributes and excluded others and arrived at a lexical ordering of those that were included. Whether this construction of the world is fruitful rests on empirical investigations that unfold over time. Our attempt in this study is to demonstrate that this approach to conceptualization and measurement bears further exploration.

Acknowledgement

Previous versions of this article were presented at the annual meetings of the American Political Science Association, the University of Amsterdam, Aarhus University, and the University of Heidelberg. We thank participants at these meetings for their comments. We also acknowledge valuable feedback received from Michael Coppedge, Adam Glynn, Gary Goertz, Staffan Lindberg, James Mahoney, Jørgen Møller, Gerardo Munck, Alexander Schmotz, Andries van der Ark, and the anonymous reviewers for the journal.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors received financial support for the research from the Danish Council for Independent Research [11932].

Notes

1. A short list would include democracy–dictatorship (DD) and Boix, Miller, and Rosato (BMR)—already discussed—as well as Bernhard, Nordstrom, and Reenock (2001) and the overview of electoral democracies by Freedom House (2013).

2. The data set (and future editions) can be downloaded at www.ps.au.dk/dedere and http://thedata.harvard.edu/dvn/dv/skaaning
3. The Polity2 index is less discriminating than it appears; countries tend to bunch in two areas—toward the bottom of the index and at the very top—producing a strongly bimodal distribution.
4. In addition, Vanhanen's index has been criticized for low construct validity (Munck, 2009).
5. Here, the term *factor analysis* is used in a general fashion to refer to a large class of models including principal components analysis.
6. Where lexical ordering is unclear a priori (according to considerations of centrality and dependence), one is well advised to consider the shape of the empirical universe. Specifically, if *A* is always (or almost always) present where *B* is present, there may be grounds for considering *A* as more central or more fundamental than *B*. However, any conclusions reached on the basis of an exploration of empirical properties must be justified as a matter of centrality or dependence. Thus, we regard the relative prevalence of attributes as a clue to asymmetric relationships among the properties of a concept, not as a desideratum. In constructing a lexical scale, deductive considerations trump data distributions.
7. Such as the nonelectoral powerbase or the level of civil liberties, the rule of law, or socioeconomic equality.
8. This is consistent with usage of the suffrage concept by Schumpeter and Przeworski and also with many extant indices such as BMR.
9. It is not necessary as an incumbent party can be sufficiently popular to win a long sequence of genuinely contested elections, as happened for decades in, for example, Botswana, Japan, and Sweden. It is not sufficient because the opposition can gain power through a flawed election if the incumbents have only weak control on power or have stepped down. Moreover, the fact that the incumbents step down after a particular election does not necessarily mean that previous elections under their leadership were competitive—as it is assumed by the DD if the previous elections took place under the same electoral rules. That said, in all but a few cases executive turnover in conjunction with elections is associated with a coding of 1 for COMPETITION.
10. In no extant cases was universal female suffrage introduced before universal male suffrage, so in practice this level is reserved for countries with male (only) suffrage.
11. In case of disagreements, we searched for additional sources and revised the coding if additional information suggested doing so.
12. Apart from the lexical index, all data used are taken from the Quality of Government (QoG) standard data set (Teorell et al., 2013).
13. In placing the lexical index on the left side of a causal model, one would of course want to use an appropriate estimator. Traditionally, ordinal outcomes are tested with ordered logit or ordered probit models. One might also construct binary variables using different cut-points on the lexical index, which could then be analyzed with logit or survival models.

References

Alvarez, R. M., Cheibub, J. A., Limongi, F., & Przeworski, A. (1996). Classifying political regimes. *Studies in Comparative International Development*, *31*(2), 3-36.

Armstrong, D. A. II. (2011). Stability and change in the Freedom House political rights and civil liberties measures. *Journal of Peace Research*, *48*, 653-662.

Bernhard, M., Nordstrom, T., & Reenock, C. (2001). Economic performance, institutional intermediation, and democratic survival. *Journal of Politics*, *63*, 775-803.

Boix, C., Miller, M., & Rosato, S. (2013). A complete dataset of political regimes, 1800–2007. *Comparative Political Studies*, *46*, 1523-1554.

Brown, M. E., Lynn-Jones, S. M., & Miller, S. E. (Eds.). (1996). *Debating the democratic peace*. Cambridge, MA: MIT Press.

Cheibub, J. A., Gandhi, J., & Vreeland, J. R. (2010). Democracy and dictatorship revisited. *Public Choice*, *143*, 67-101.

Collier, D., & Adcock, R. (1999). Democracy and dichotomies: A pragmatic approach to choices about measures. *Annual Review of Political Science*, *2*, 537-565.

Collier, D., & Levitsky, S. (1997). Democracy with adjectives: Conceptual innovation in comparative research. *World Politics*, *49*, 430-451.

Coppedge, M., Alvarez, A., & Maldonado, C. (2008). Two persistent dimensions of democracy: Contestation and inclusiveness. *Journal of Politics*, *70*(3), 335-350.

Coppedge, M., Gerring, J., Altman, D., Bernhard, M., Fish, S., Hicken, A., . . . Teorell, J. (2011). Conceptualizing and measuring democracy: A new approach. *Perspectives on Politics*, *9*, 247-267.

Correlates of War. (2011). *State system membership (v2011)*. Retrieved from http://correlatesofwar.org/data-sets/state-system-membership/state-system-membership

Dahl, R. A. (1971). *Polyarchy: Participation and opposition*. New Haven, CT: Yale University Press.

Davenport, C., & Armstrong, D. (2004). Democracy and the violation of human rights: A statistical analysis from 1976–1996. *American Journal of Political Science*, *48*, 538-554.

Diamond, L. (2002). Thinking about hybrid regimes. *Journal of Democracy*, *13*(2), 21-35.

Elkins, Z. (2000). Gradations of democracy? Empirical tests of alternative conceptualizations. *American Journal of Political Science*, *44*, 287-294.

Freedom House. (2013). *Freedom in the world 2013*. Retrieved from http://www.freedomhouse.org/report/freedom-world-2013/methodology

Gandhi, J., & Lust-Okar, E. (2009). Elections under authoritarianism. *Annual Review of Political Science*, *12*, 403-422.

Gerring, J., Palmer, M., Teorell, J., & Zarecki, D. (2015). Demography and democracy: A global, district-level analysis of electoral contestation. *American Political Science Review*, *109(3)*, 574-591.

Gerring, J., Skaaning, S.-E., & Pemstein, D. (2014). *A concept-driven approach to measurement: The Lexical Scale*. Unpublished manuscript, Boston University, MA.

Gleditsch, K. S. (2013). *List of independent states*. Retrieved from http://privatewww.essex.ac.uk/~ksg/statelist.html

Gleditsch, K. S., & Ward, M. D. (1997). Double take: A re-examination of democracy and autocracy in modern polities. *Journal of Conflict Resolution*, *41*, 361-383.

Goertz, G. (2006). *Social science concepts: A user's guide*. Princeton, NJ: Princeton University Press.

Hadenius, A., & Teorell, J. (2007). Pathways from authoritarianism. *Journal of Democracy*, *18*(1), 143-156.

Held, D. (2006). *Models of democracy*. Stanford, CA: Stanford University Press.

Howard, M., & Roessler, P. (2006). Liberalizing electoral outcomes in competitive authoritarian regimes. *American Journal of Political Science*, *50*, 365-381.

Huntington, S. P. (1991). *The third wave: Democratization in the late twentieth century*. Norman: University of Oklahoma Press.

Keyssar, A. (2009). The right to vote: The contested history of democracy in the United States. New York, NY: Basic Books.

Knutsen, C. H. (2010). Measuring effective democracy. *International Political Science Review*, *31*, 109-128.

Kurtz, M., & Schrank, A. (2007). Growth and governance: Models, measures, and mechanisms. *Journal of Politics*, *69*, 538-554.

Levitsky, S., & Way, L. (2002). The rise of competitive authoritarianism. *Journal of Democracy*, *13*(2), 51-55.

Lively, J. (1975). *Democracy*. New York, NY: St. Martin's.

Marshall, M. G., Gurr, T., & Jaggers, K. (2013). *Polity IV project: Dataset Users' Manual*. Retrieved from http://www.systemicpeace.org/inscr/p4manualv2012.pdf

Merkel, W. (2004). Embedded and defective democracies. *Democratization*, *11*(1), 33-58.

Møller, J., & Skaaning, S.-E. (2013). Regime types and democratic sequencing. *Journal of Democracy*, *24*(1), 142-156.

Munck, G. L. (2009). *Measuring democracy: A bridge between scholarship and politics*. Baltimore, MD: Johns Hopkins University Press.

Naess, A., Christophersen, J., & Kvalo, S. (1956). *Democracy, ideology, and objectivity*. Oslo, Norway: Universitetsforlaget.

Pemstein, D., Meserve, S., & Melton, J. (2010). Democratic compromise: A latent variable analysis of ten measures of regime type. *Political Analysis*, *18*, 426-449.

Phillips, J., & Wetherell, C. (1995). The Great Reform Act of 1832 and the political modernization of England. *American Historical Review*, *100*, 411-436.

Poggi, G. (1978). The development of the modern state: A sociological introduction. Stanford, CA: Stanford University Press.

Przeworski, A. (2013). *Political Institutions and Political Events (PIPE) data set*. Retrieved from https://sites.google.com/a/nyu.edu/adam-przeworski/home/data

Przeworski, A., Alvarez, M., Cheibub, J. A., & Limongi, F. (2000). *Democracy and development: Political institutions and material well-being in the world, 1950–1990*. New York, NY: Cambridge University Press.

Ravallion, M. (2011). Mashup indices of development. *World Bank Research Observer*, *27*, 1-32.
Schattschneider, E. E. (1942). *Party government*. New York, NY: Rinehart.
Schedler, A. (2002). The menu of manipulation. *Journal of Democracy*, *13*(2), 36-50.
Schedler, A. (2012). Judgment and measurement in political science. *Perspectives on Politics*, *10*, 21-36.
Schumpeter, J. A. (1950). *Capitalism, socialism and democracy*. New York, NY: Harper & Row.
Seawright, J., & Collier, D. (2014). Rival strategies of validation: Tools for evaluating measures of democracy. *Comparative Political Studies*, *47*, 111-138.
Svolik, M. (2012). *The politics of authoritarian rule*. New York, NY: Cambridge University Press.
Teorell, J., Charron, N., Dahlberg, S., Holmberg, S., Rothstein, B., Sundin, P., & Svensson, R. (2013). *The quality of government dataset, version 20Dec13*. Gothenburg, Sweden: University of Gothenburg, the Quality of Government Institute. Retrieved from http://www.qog.pol.gu.se/
Therborn, Göran (1977). The rule of capital and the rise of democracy. *New Left Review, 103*, 3-41.
Treier, S., & Jackman, S. (2008). Democracy as a latent variable. *American Journal of Political Science*, *52*, 201-217.
Vanhanen, T. (2000). A new dataset for measuring democracy, 1810–1998. *Journal of Peace Research*, *37*, 251-265.
Vermillion, J. (2006). Problems in the measurement of democracy. *Democracy at Large*, *3*(1), 26-30.
Wood, R., & Gibney, M. (2010). The Political Terror Scale (PTS): A re-introduction and a comparison to CIRI. *Human Rights Quarterly*, *32*, 367-400.

Author Biographies

Svend-Erik Skaaning, PhD, Aarhus University, 2007, is professor of Political Science at Aarhus University. His research interests include democratization, human rights, and comparative methods. He is the co-author of *Requisites of Democracy* (Routledge, 2011), *Democracy and Democratization* (Routledge, 2013), and *The Rule of Law* (Palgrave, 2014), along with numerous articles.

John Gerring, PhD, University of California at Berkeley, 1993) is professor of Political Science at Boston University, where he teaches courses on methodology and comparative politics. He is the author of *Party Ideologies in America*, 1828-1996 (Cambridge University Press, 1998), *Case Study Research: Principles and Practices* (Cambridge University Press, 2007), *A Centripetal Theory of Democratic Governance* (Cambridge University Press, 2008), *Concepts and Method: Giovanni Sartori and His Legacy* (Routledge, 2009), *Social Science Methodology: A Unified Framework*, 2nd ed (Cambridge University Press, 2012), *An Applied Guide to Social Science Methodology* (Cambridge University Press, 2016), along with numerous articles.

Henrikas Bartusevičius (PhD, Aarhus University, 2013) is an assistant professor at the Department of Political Science at Aarhus University. He was previously a post-doctoral research fellow at the Department of Government at University of Essex. His research interests include violent and non-violent interactions in civil conflict, democratization and civil conflict, and psychological biases and human aggression. His work has been published in the *Journal of Peace Research and Conflict Management and Peace Science*.