



Measuring Electoral Democracy with  
V-Dem Data:  
Introducing a New Polyarchy Index

Jan Teorell, Michael Coppedge,  
Svend-Erik Skaaning, and Staffan I. Lindberg

March 2016

Working Paper

SERIES 2016:25

THE VARIETIES OF DEMOCRACY INSTITUTE



UNIVERSITY OF GOTHENBURG  
DEPT OF POLITICAL SCIENCE

**Varieties of Democracy (V-Dem)** is a new approach to the conceptualization and measurement of democracy. It is co-hosted by the University of Gothenburg and University of Notre Dame. With a V-Dem Institute at University of Gothenburg that comprises almost ten staff members, and a project team across the world with four Principal Investigators, fifteen Project Managers, 30+ Regional Managers, 170 Country Coordinators, Research Assistants, and 2,500 Country Experts, the V-Dem project is one of the largest-ever social science research-oriented data collection programs.

Please address comments and/or queries for information to:

V-Dem Institute  
Department of Political Science  
University of Gothenburg  
Sprängkullsgatan 19, PO Box 711  
SE 40530 Gothenburg  
Sweden  
E-mail: [contact@v-dem.net](mailto:contact@v-dem.net)

V-Dem Working Papers are available in electronic format at [www.v-dem.net](http://www.v-dem.net).

Copyright © 2016 by authors. All rights reserved.

# Measuring Electoral Democracy with V-Dem Data: Introducing a New Polyarchy Index\*

Jan Teorell  
Department of Political Science  
Lund University

Michael Coppedge  
Professor of Political Science  
University of Notre Dame

Svend-Erik Skaaning  
Professor of Political Science  
Aarhus University

Staffan Lindberg  
Professor of Political Science  
Director, V-Dem Institute  
University of Gothenburg

---

\* Previous versions of this paper were presented at the APSA Annual Meeting in Washington, DC, August 28-31, 2014, at the Carlos III-Juan March Institute of Social Sciences, Madrid, November 28, 2014, and at the European University Institute, Fiesole, January 20, 2016. Any remaining omissions are the sole responsibility of the authors. This research project was supported by Riksbankens Jubileumsfond, Grant M13- 0559:1, PI: Staffan I. Lindberg, V-Dem Institute, University of Gothenburg, Sweden; by Swedish Research Council, Grant C0556201, PI: Staffan I. Lindberg, V-Dem Institute, University of Gothenburg, Sweden and Jan Teorell, Department of Political Science, Lund University, Sweden; and by Knut and Alice Wallenberg Foundation to Wallenberg Academy Fellow Staffan I. Lindberg, Grant 2013.0166, V-Dem Institute, University of Gothenburg, Sweden; as well as by internal grants from the Vice-Chancellor's office, the Dean of the College of social Sciences, and the Department of Political Science at University of Gothenburg. Jan Teorell also wishes to acknowledge support from the Wenner-Gren Foundation and the Fernand Braudel Senior Fellowship at the European University Institute, Florence, which made it possible for him to work on this paper. We performed simulations and other computational tasks using resources provided by the Notre Dame Center for Research Computing (CRC) through the High Performance Computing section and the Swedish National Infrastructure for Computing (SNIC) at the National Supercomputer Centre in Sweden. We specifically acknowledge the assistance of In-Saeng Suh at CRC and Johan Raber at SNIC in facilitating our use of their respective systems.

## Abstract

This paper presents a new measure of electoral democracy, or "polyarchy", for a global sample of 173 countries from 1900 to the present based on the Varieties of Democracy (V-Dem) data, enabling us to address several deficiencies in extant measures of electoral democracy, such as Freedom House and Polity. The V-Dem data derive from expert polls of more than 2,600 country experts from around the world, with on average 5 experts rating each indicator. By measuring the five components of "Elected officials", "Free and fair elections", "Freedom of expression", "Associational autonomy" and "Inclusive citizenship" separately, we anchor this new index directly in Dahl's (1971) extremely influential theoretical framework, and can both show how well indicators match components as well as how components map the overall index. We also find that characteristics of the V-Dem country experts do not systematically predict their ratings on our indicators, nor differences between these ratings and existing measures such as FH and Polity, with which they are strongly correlated. Finally, we provide systematic measures of uncertainty (or measurement error) at every level. We showcase the usefulness of the new measure for understanding developments of electoral democracy over time, for comparing countries at a particular time point, and for understanding its relationship to economic modernization through disaggregation.

# 1. Introduction

Democracy may mean many things, but in this paper we shall only concern ourselves with the concept used as an empirical yardstick by which countries, or national political systems, may be compared. More specifically, we shall concentrate on a particular way of understanding democracy as a yardstick, one that has dominated the empirical social sciences since at least the second world war. Through the seminal work of Schumpeter (1942), Downs (1957) and Dahl (1956, 1971), a relative consensus has evolved that views elections – together with the institutions that uphold the democratic qualities of elections – as the core of the concept. We will call this the *electoral* conception of democracy, to distinguish it from other, more idealized, conceptions – such as liberal, participatory, deliberative, or egalitarian democracy (see Coppedge et al. 2015a, 2016, Lindberg et al. 2014) – that do not stress elections as democracy’s core institutional pillar.

Given this conceptual point of departure, we propose a new way of measuring electoral democracy. This measure is based on the methodology of the *Varieties of Democracy* project (Coppedge et al. 2015d, see also [www.v-dem.net](http://www.v-dem.net)), which draws on multiple expert perceptions of a large number of disaggregated indicators from a global sample of countries since 1900. More precisely, we use ratings provided by over 2,600 scholars and other experts on 36 specific indicators at country-year level to measure the core “institutional guarantees” in Dahl’s (1971, 1989, 1998) concept of “polyarchy.” We also launch an aggregate index of these components to measure electoral democracy. At all levels, we provide both point estimates and a measure of uncertainty.

We argue that this new index has several advantages over the widely used existing measures with similar coverage, such as Freedom House, Polity, Vanhanen (2000), Przeworski *et al.* (2000) and Boix *et al.* (2011). First, we anchor the new index directly in Dahl’s (1971) extremely influential theoretical framework. By measuring the five Dahlian (1998, 85) components of “Elected officials”, “Free and fair elections”, “Freedom of expression”, “Associational autonomy” and “Inclusive citizenship” separately, we can for the first time since Coppedge & Reinicke (1990) allow for an analysis of how these components hang together empirically. The data is public and our measure also allows for disaggregation, both to the level of the five components and down to the indicators constituting these. Second, we also allow for both minimalist and maximalist versions of electoral democracy (Munch and Verkuilen 2002), and can thus systematically explore for the first time how these different conceptions affect our estimates of democracy in the world. Third, our data generating process and aggregation scheme is fully transparent, and can be cross-examined. More specifically, we show below that characteristics of the V-Dem country experts do not systematically predict their ratings on our

indicators, nor differences between these ratings and existing measures such as FH and Polity, with which they are strongly correlated. Finally, we provide systematic measures of uncertainty (or measurement error) at every level.

In the following, the first section details the conceptualization of polyarchy followed by our approach to measurement. The third part outlines the V-Dem methodology followed by a section on how we measure the components of polyarchy. The fifth section deliberates on how we aggregate these components into the polyarchy index. After a section probing the validity of the underlying V-Dem data and indices, we end by providing some empirical illustrations results to highlight the added value of our new measure.

## **2. Conceptual Points of Departure: Polyarchy vs. Minimalist Conceptions of Democracy**

Dahl (1971, 2) famously defined democracy as “a political system one of the characteristics of which is the quality of being completely or almost completely responsive to all its citizens.” Yet he reserved this term, “democracy,” for an ideal system without empirical referents, instead preferring to label countries living up to the empirical requirements of democracy “polyarchies.” Originally eight, these requirements – or “institutional guarantees” – were in later work narrowed down to seven (Dahl 1989) and eventually six (Dahl 1998, 85). We have in turn collapsed two of these on theoretical grounds into the five guarantees presented in Table 1<sup>1</sup>.

---

<sup>1</sup> The two requirements from Dahl (1971, 3) that are missing in Dahl (1998, 85) are: “Eligibility for public office” and “Institutions for making government policies depend on votes and other expressions of preference” We agree with Dahl that these could be omitted, the first on the one hand because eligibility and suffrage tend to go hand in hand (as argued by Coppedge & Reinicke 1990, 53) , and on the other because several of the aspects of “Freedom of organization” such as freedom to organize political parties and for them to run in elections, captures much of the eligibility criteria. The second can be dropped because it is more of a summary proxy for all the other institutional requirements taken together, and not the least by the mechanism of free, fair, and regular multiparty elections. The remaining sixth guarantee that we include by collapsing it with freedom of expression is “alternative sources of information”. The latter is theoretically very closely connected to, and as we show empirically indistinguishable from, the former.

**Table 1. The Five Institutional Guarantees of Polyarchy**

	<i>Minimalist conception:</i>	<i>Maximalist conception:</i>
A. Elected officials	X	X
B. Free, Fair and Frequent Elections	X	X
C. Associational Autonomy	X	X
D. Inclusive Citizenship		X
E. Freedom of Expression		X

*Source:* Dahl 1998, 85 (“Alternative sources of information” collapsed with “Freedom of Expression”). The “minimalist” conception corresponds most closely to Schumpeter (1942).

Dahl’s notion of polyarchy is sometimes juxtaposed to a more “minimalist” conception of democracy dating back to Schumpeter’s (1942, 269) famous assertion that “the democratic method is that institutional arrangement for arriving at political decisions in which individuals acquire the power to decide by means of a competitive struggle for the people’s vote.” This conception is in at least two senses more “minimal” than Dahl’s. First in that it excludes from its definition any reference to political liberties.<sup>2</sup> Second, in arguing that “disqualifications on grounds of economic status, religion and sex [are] compatible with democracy”, Schumpeter (1942, 244-5) clearly excluded the right to vote (or the extension of the suffrage) from his conception of democracy (Møller & Skaaning 2010, 268-9). In terms of Dahl’s (1971, 4) two famous “dimensions of democratization,” Schumpeter thus stresses competition or contestation at the expense of inclusiveness or participation. This implies that Schumpeter primarily defines democracy by reference to the first, second and third Dahlian institutional guarantees only (see Table 1).

Typically, the argument in favor of minimalist conceptions is to allow the relationship between democracy and other components to be empirically determined rather than settled by definitional fiat. Boix *et al.* (2013, 1527), for example, argue that “by not bundling in additional elements of democratic practice, such as civil liberties, it allows researchers to empirically relate these elements to regime type.” Similarly, Przeworski *et al.* (2000, 34) argue in favor of excluding suffrage from their conception of democracy in order to “be able to test theories about the effects of participation on the performance and the durability of democracy.” On the one hand, such empirical considerations cannot solely

---

<sup>2</sup> As pointed out by O’Donnell (2001, 9), Schumpeter (1942, 272) allows that freedom to compete for political leadership “will in most cases though not in all mean a freedom of discussion *for all*. In particular it will normally mean a considerable amount of freedom of the press.” The reservations “though not in all” and “normally” to us however signal that freedom of discussion and the press are for Schumpeter not definitional criteria, but likely outcomes or empirical implications of having a democratic system.

override theoretical issues of conceptualization. It remains a fact that the bulk of work attempting to measure democracy in the world has relied on Dahl's more maximalist conceptualization (e.g., Bollen 1980, 1990; Coppedge & Reinicke 1990; Hadenius 1992; Gasiorowski 1996; Mainwaring *et al.* 2001; Bowman *et al.* 2005). On the other hand, this is a false dichotomy given adequate data and disaggregation. An attractive feature of the approach we are proposing is that it allows room for both researchers wanting to stick with the maximalist conception and those wanting to test empirical relationships using one of the more minimalist ones or the individual components and indicators.

The argument against Dahlian "maximalist" conception is that too many features of the political system are built into the concept. Most importantly, why should a non-electoral dimension such as freedom of expression be made part of a measure of *electoral* democracy? This is a legitimate question, and some authors have accordingly suggested to reserve the term "electoral democracy" for the more minimalist Schumpeterian conception (e.g., Möller & Skaaning 2010, 268-171; Munck 2009, 55-56). Without denying the need for such diminished subtypes for certain analytical purposes, we maintain that polyarchy provides to most fully fleshed theory of what constitutes an electoral democracy. "By this conception," in Diamond's (2002, 21) words, "democracy requires not only free, fair, and competitive elections, *but also the freedoms that make them truly meaningful* (such as freedom of organization and freedom of expression, alternative sources of information, and institutions to ensure that government policies depend on the votes and preferences of citizens)" (italics added). To avoid the "fallacy of electoralism" (Karl 1986), even election-centered notions of democracy need to take into account some non-electoral aspects, most importantly freedom of organization and expression, in order to ascertain that elections work as intended.<sup>3</sup>

---

<sup>3</sup> O'Donnell (2001) also interprets Dahl's concept of polyarchy as listing some attributes of elections and then adding certain freedoms "deemed necessary for elections to be democratic" (12), although he then surmises that since what exact freedoms are required to maintain the democratic qualities of elections is an inductive question, this practice introduces a conceptual "can of worms" that already Schumpeter tried but failed to avoid (15). We agree with him on the first but not the second point.

### 3. Measurement Approach: Reflective vs. Constitutive Indicators of Democracy

Whether one adopts a reflective or constitutive model of how indicators and components are related to higher-level concepts (Goertz 2006) is critical for how to measure the electoral concept of democracy. In classical measurement theory, concepts are viewed as latent constructs that give rise to, or “cause,” their respective observable indicators. These are thus termed “effect” or *reflective* indicators. A person’s score on an IQ test, for example, is viewed as an “effect” caused by the underlying, unobserved attribute “intelligence.” As Goertz (2006) points out, however, there is another “ontological” approach to measurement where indicators are seen as *constituting* the concept. By defining “human development” as the composite of life expectancy, level of education and economic development, for example, UNDP is arguing that these three components constitute the overall concept – not that they are “effect” indicators “caused” by it.

Both the reflective and the constitutive view of democracy measures have been advocated.<sup>4</sup> Bollen (1990, 12) provides an eloquent presentation of the former view: “Conceptually, I view political rights and civil liberties as aspects of political democracy. I could imagine these as moving in response to changes in democracy. For instance, if democracy declines in a nation, I would expect this to lead to declines in political liberties and rights.” Democracy is thus viewed as a latent trait (Treier & Jackman 2008), living a life of its own as it were, causing observables such as the degree of political liberties to move as an effect of the underlying level of democracy changes. Among others (also see Munck 2009), Goertz (2006, 15) advocates the opposite view in writing that “to have competitive elections is not a symptom of democracy, it is not caused by democracy, but rather it constitutes what democracy *is*.”

The choice between viewing components and indicators as reflective or as constituent elements of higher-order concepts has important repercussions for aggregation (Munck 2009, 30-31). With the classical “effect indicators” approach that considers each indicator as at least partially substitutable and

---

<sup>4</sup> The distinction is similar to but not equivalent to the one between cause and effect indicators, or reflective vs. formative measurement models (Bollen 1989, 64-5; cf. Munck 2009, 148 n. 17). We prefer the term “constituent” to “cause” or “formative” indicator, however, since the language of causes and effects implies that the relationship is still somehow an empirical one. The virtue of Goertz (2009) approach is to highlight the fact that the relationship is theoretical or conceptual, not empirical.

“caused” by the latent trait, the conventional aggregation rule is to take the average.<sup>5</sup> With the constituent approach, by contrast, a theory of the relation between the indicators, and between them and the main concept must inform the choice of aggregation rule(s). If each indicator is viewed as perfectly substitutable for the others, and hence each on its own is a sufficient condition for the higher-level construct, taking the maximum is the warranted aggregation rule. But if theory specifies that each indicator is a non-substitutable and hence each is a necessary condition, the aggregation rule should be either taking the minimum or multiplying the indicators (Goertz 2009, 39-44, 11; cf. Munck & Verkuilen 2002, 24; Munck 2009, 32).

For our new measure of polyarchy, we argue in favor of a joint approach in which some parts of the higher-level construct are best measured with reflective indicators “caused” by a higher-level component, whereas other components should be treated as “constituted” by its indicators, on theoretical grounds. We thus argue in favor of the approach that makes most theoretical sense in each case, instead of a one-size-fits-all general solution.

## 4. The V-Dem Methodology

The large literature on existing attempts to measure polyarchy, or electoral democracy, has been reviewed extensively elsewhere (Munck & Verkuilen 2002; Hadenius & Teorell 2005; Munch 2009; Coppedge et al. 2011). Suffice to say here that in our view no extant measure fulfills all of the following five essential criteria: (1) covering all components of Dahl’s concept of polyarchy (as listed in Table 1) while; (2) providing disaggregated data to check dimensionality and allow users to “drill down” in order to evaluate what lower-level changes account for the shifts in higher-level indices, or what lower-level components that drive aggregate correlations; (3) for a global sample of countries *and* across long swaths of time; (4) using transparent data generating processes and aggregation rules; and (5) providing estimates of measurement uncertainty.

To be sure, selected extant measures live up to *some* of these criteria. Coppedge & Reinicke (1990), for example, provide measures of almost all of Dahl’s components, on a global sample using transparent aggregation rules – but only for the year of 1985, and with no measures of uncertainty. Przeworski *et al.* (2000), as well as Vanhanen (2000), are renowned for their incisive coding, clear

---

<sup>5</sup> Or a weighted average, with the weights made up of the strength of the “causal” relationship between each indicator and the higher-level construct (such as factor loadings).

aggregation rules, and reliance on strictly observable data. They also provide great geographic and temporal coverage – but again with no measures of uncertainty, and only measure a minimalist subset of the concept of polyarchy. Much the same goes for Boix *et al.* (2011), except that their exact coding rules for determining whether elections are “free and fair” are not fully transparent. Treier & Jackman (2008), finally, in one of the few efforts to furnish the point estimates of democracy with confidence intervals reflecting measurement uncertainty (cf. Pemstein *et al.* 2010), suffer from the Polity data’s limitations in terms of its minimalist conception of democracy, unjustified aggregation rules, and non-transparent coding scheme.

There is thus a need for a new measure of polyarchy that can live up to all theoretical and methodological criteria simultaneously. We argue that the *Varieties of Democracy* (V-Dem for short) data makes this possible. Three features in particular underpins this claim (for a fuller description of the V-Dem methodology, see Coppedge *et al.* 2015d, Pemstein *et al.* 2015). The first feature is radical disaggregation: we translate the highest-level principles of democracy into over 350 detailed questions with well-defined response categories or measurement scales. Although these have been designed to tap into a variety of conceptions of democracy, this still means we have constructed the polyarchy scale presented below on the basis of 36 individual variables. This not only implies that we measure all components of Dahl’s polyarchy concept, but also that each component save one is measured with multiple indicators, enhancing reliability and providing a basis for tests of dimensionality.

Second, covering 173 country units across the entire globe (the list of countries covered is displayed in Appendix A), the bulk of the indicators stems from data collected from country experts, mostly academics from each country in question.<sup>6</sup> These experts have been recruited based on their academic or other credentials as field experts in the area<sup>7</sup> they code, on their seriousness of purpose and impartiality. The target is that at least 5 experts rate each indicator for each country and year going back to 1900. This means that more than 2,600 experts in all have helped us gather the data.

While we select experts carefully, they clearly exhibit varying levels of reliability and bias. Therefore – and this is the third unique feature of V-Dem – we use Bayesian item response theory (IRT) modeling techniques to estimate latent country coding unit characteristics from the collection of

---

<sup>6</sup> A notable feature of the V-Dem data is that we code a “country” throughout its history also as a semi-sovereign unit (until 1900). This implies that most colonies, and also some current non-independent territories such as Kosovo and the Palestines, are included in the sample.

<sup>7</sup> The questionnaire is subdivided into 11 different areas of expertise, and most experts code a cluster of three such areas.

expert ratings (see Pemstein *et al.* 2015). The underpinnings of these measurement models are straightforward: they use patterns of cross-rater (dis)agreement to estimate variations in reliability and systematic differences in threshold between ordinal response categories to adjust estimates of the latent concept in question. Since most coders rate one country only in areas based on their expertise, we also utilize “lateral” coders who rate multiple countries for a limited time period, as well as “bridge” coders who code the full time series for more than one country. As of December 2015, we have over 360 bridge coders – about 14 percent of all country experts – who have coded 6.1 surveys for 2.1 countries on average, as well as more than 350 country experts (about 15%) who have performed lateral coding, covering on average of 5.5 countries and 6.3 surveys. By this, every country is directly or indirectly connected across all surveys. Essentially, this allows us to mitigate the potential incomparability of coders’ thresholds and enhance cross-national comparability.

## 5. Measuring the Parts: Five Components of Polyarchy

Our description of the V-Dem measure of polyarchy proceeds as follows. In this section, we discuss the five component indices, and the indicators used to measure them, corresponding to Dahl’s “institutional guarantees” presented in Table 1. In the next section, we present the aggregation rules for combining them and the resulting index of electoral democracy (polyarchy).

### *A. Elected executive*

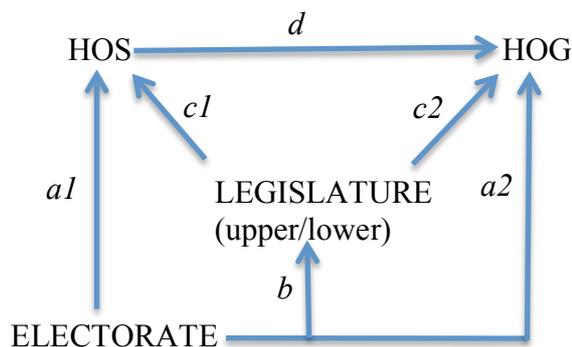
This index measures whether the chief executive is elected, either directly through popular elections or indirectly through a popularly elected legislature that then appoints the chief executive. In the literature on measuring democracy this concept is also referred to as the “effectiveness” (Hadenius 1992, 49) or “decisiveness” (O’Donnell 2001, 13) of elections. A “popular election” is thus minimally defined and also includes sham elections with limited suffrage and no competition. Similarly, “appointment” by legislature only implies selection and/or approval, not the power to dismiss.<sup>8</sup>

---

<sup>8</sup> The rationale for this is that counting dismissal powers would introduce a systematic “bias” in favor of parliamentary systems, which typically are the only systems relying on institutionalized powers of dismissal over and above the power to appoint or elect the chief executive. In presidential systems, for example, there typically is no recall vote at the national level.

The logic behind the construct is schematically portrayed in Figure 1. Unlike the following component indices, the index is based on a set of constitutive indicators that logically define the overall construct, regardless of whether the individual indicators hang together empirically or not. This is motivated out of purely theoretical concerns.

**Figure 1. The Logical Structure of the Elected Officials Index**



There are six different links of appointment/selection to take into account, all of which we scale to vary from 0 to 1. First and second, whether the head of state ( $a1$ ) and/or head of government ( $a2$ ) is directly elected (1) or not (0). Third, the extent to which the legislature is popularly elected ( $b$ ), measured as the proportion of legislators elected (if legislature is unicameral), or the weighted average of the proportion elected for each house, with the weight defined by which house is dominant (if legislature is bicameral).<sup>9</sup> Fourth and fifth, whether the head of state ( $c1$ ) and/or head of government ( $c2$ ) is appointed by the legislature (1), or the approval of the legislature is necessary for the appointment of the head of state (1), or not (0). Sixth, whether the head of government is appointed by the head of state ( $d=1$ ) or not ( $d=0$ ).

---

Also, in the assembly independent (albeit very unusual) Swiss system type, the legislature lacks the power to dismiss the cabinet.

<sup>9</sup> In terms of the V-Dem variables (see Appendix B):  $b=v2lgello$  (converted to fractions) in unicameral legislatures ( $v2lgbicam==1$ ), and a weighted average of  $v2lgello$  and  $v2lgelecup$  (converted to fractions) in bicameral legislatures ( $v2lgbicam==2$ ), where  $v2lgdomchm\_ord/4$  defines the weight. In the current version of V-Dem, no distinction is made between indirectly elected and appointed legislatures, which are thus treated equally. We hope to implement such a distinction in future versions.

In politics with unified executives, that is, where the head of state is also the head of government (Elgie 1998; Siaroff 2003), the complexity of this conceptual scheme reduces to the links  $a1$  and  $b*c$ . Since these are considered perfect substitutes (*either* a directly elected president *or* a president elected by a directly elected parliament suffices), the index value is thus achieved at by taking the maximum value of the two. In dual systems, where there is both a head of state and a head of government, the chief executive is determined by comparing the two executives' power over the appointment and dismissal of cabinet ministers. If the head of state and head of government share equal powers over the appointment and dismissal of cabinet ministers, the index averages across the extent to which both are directly or indirectly elected.<sup>10</sup>

As Figure C1 of Appendix C reveals, the resulting index (called  $v2x\_accex$ ) has a bimodal distribution. The rare values falling between 1 and 0 are mostly bicameral systems where the upper house is not directly elected, but also include cases such as Burma/Myanmar that (currently) has a unified executive with a president elected by parliament, in which only 75 % of the seats are directly elected.

Three potential limitations of the index should be noted. First, since this is an index of whether the executive is *appointed* through elections, the extent to which non-elected “accountability groups” (such as the military) may affect dismissal of the executive, or can veto important domestic policy proposals, has not been taken into account. Secondly, indirectly elected legislators, common in the upper house of federal systems, are treated the same as appointed legislators (that is, as unelected). Finally, since this index concentrates on the election/appointment of the chief *executive*, presidential systems with reserved, unelected seats in the legislature are not being penalized. For example, 8 out of 150 seats in the unicameral national assembly of Zambia, one third of the deputies of the “House of Elders” in Afghanistan, as well as 7 out of 47 seats in the “Senate” in Kazakhstan are presidential appointees, but since the chief executive in these systems (the president) is still directly elected, the index assumes the value 1.

---

<sup>10</sup> In technical terms, define  $bosw$  as the weight for the head of state. If the head of state is also head of government ( $v2exhoshog==1$ ),  $bosw=1$ . In dual systems ( $v2exhoshog==0$ ), if the head of state has more power than the head of government over the appointment ( $v2exdfcbhs\_rec>v2exdjcbhg$ ) and dismissal ( $v2exdfdmhs>v2exdfdshg$ ) of cabinet ministers, then  $bosw=1$ ; if the reverse is true,  $bosw=0$ . If the head of state has more appointment ( $v2exdfcbhs\_rec>v2exdjcbhg$ ) but the head of government more dismissal power ( $v2exdfdmhs<v2exdfdshg$ ), or vice versa,  $bosw=.5$ . Define the weight for the head of government as  $bogn=1-bosw$ . The index is then defined as  $bosw*[\max(a1, b*c1)]+bogn*[\max(a1*d, b*c1*d, a2, b*c2)]$ .

## B. Clean elections

The second component, which Dahl (1998, 85) calls “free, fair and frequent” elections, is designed to capture the level of integrity of elections measured as the absence of manipulation and bias in the administration of elections (e.g. Calingaert 2006, Lehoucq 2003, Bermeo 2010, Birch 2011, Donno 2013, Simpser 2013, Kelley 2013, Scheduler 2002, 2013, Norris 2015, van Ham and Lindberg 2015). Disregarding the “frequency” aspect, since that would build an unnecessary temporal aspect into the concept of polyarchy, this component is thus in essence tapping into whether an election could be considered free from manipulation. Or in other words, whether the election is “clean”.

The V-Dem dataset includes six indicators pertaining to this latent trait, rated by the country experts for each election (see Appendix B on exact question wording and response categories).<sup>11</sup> These measure the extent of (a) registration irregularities, (b) vote buying, (c) ballot fraud and intentional irregularities, (d) government-induced intimidation of opposition candidates, (e) other types of election violence (not instigated by the government or ruling party), as well as (f) an overall assessment of whether the election, all things considered, could be considered “free and fair.” Finally, two indicators are measured on annual basis: The (g) autonomy and (h) capacity of the election administration body (EMB) to conduct well-run elections.

We view the aggregated country-year data on these eight indicators as reflective indicators caused by the unobserved latent trait “clean elections”, and therefore partially substitutable. To test this measurement proposition, we ran a Bayesian factor analysis model and the results are reported in Table 2.<sup>12</sup> As can be seen, the vote buying and election violence (not conducted by the government) indicators have somewhat weaker loadings and a larger share of their variance unaccounted for (uniqueness). By and large, however, the fit to a unidimensional model seems adequate.

---

<sup>11</sup> If legislative (lower house or constituent assembly) and presidential elections were held concurrently, the questions pertain to both of them. If multiple elections (or rounds of elections) were held in the same year, however, we asked about each election separately. For present purposes, the estimates have been averaged across multiple elections within a year to arrive at country election-year estimates. Election types and dates were compiled from multiple sources (see codebook at [www.v-dem.net](http://www.v-dem.net)).

<sup>12</sup> To run the Bayesian factor analysis models, we have used the *factanal()* command in the MCMC package for R (Martin *et al.* 2014). We owe thanks to Yi-ting Wang and Eitan Tzelgov at University of Gothenburg for helping us out in setting up the R code for these analyses.

**Table 2. Measuring Clean Elections (BFA estimates)**

<i>Indicator</i>	<i>Loadings (<math>\Lambda</math>)</i>	<i>Uniqueness (<math>\Psi</math>)</i>
Proper voter registry (v2elrgstry)	.811	.344
Vote buying (v2elvotbuy)	.680	.538
Ballot fraud/vote irregularity (v2elirreg)	.873	.240
Government intimidation (v2elintim)	.866	.251
Other electoral violence (v2elpeace)	.656	.570
Overall judgment “free & fair” (v2elfrfair)	.892	.207
EMB autonomy (v2elembaut)	.871	.243
EMB capacity (v2elembcap)	.809	.347

*Note:* Entries are factor loadings and uniqueness from a normal theory Bayesian factor analysis model;  $n=3,271$  election years.

The index of clean elections is based on the point estimates from this Bayesian factor analysis model (distributed as in Figure C2), accompanied by measures of uncertainty.<sup>13</sup> Since this index is only observed for election years, however, we face the issue of how to extrapolate its values over time. Although one could contemplate extrapolation rules that take election frequency into account (for example by adding a yearly depreciation rate, such as .95), the approach taken here is to simply repeat the index values from the last election until either another election occurs or there is an “electoral interruption,” defined as either (i) the dissolution, shut-down, replacement or in any sense termination of the elected body (such as after coups or violent takeovers of the government, etc), or (ii) that the elected body in question, while still intact or in place, is no longer to be appointed through (direct) elections (as after an autogolpe).<sup>14</sup>

A thornier issue concerns what value to assign to the index during periods of such electoral interruptions. Recall that, being based on the scores from the item response theory measurement model, the index has no natural zero point and a theoretically defined minimum score cannot be identified. Our solution is first (this step pertains to all V-Dem indices composed using BFAs) to

<sup>13</sup> More specifically, we randomly select 100 draws from each variable’s posterior distribution (see Pemstein et al 2015 for on the measurement model), and use a unidimensional Bayesian factor analysis (BFA) to measure this latent concept sequentially for each randomly-selected draw in each grouping of variables. We then combine the posterior distributions of the latent factor scores in each variable group to yield the latent factor scores. In all analyses the variables generally load highly on the underlying factor, and by this procedure we take full advantage of the estimates of uncertainty at the indicator level, to provide realistic (not overly narrow) estimates of uncertainty at the component-level.

<sup>14</sup> The V-Dem dataset has a specific indicator for this, capturing electoral interruptions (v2x\_elecrg).

convert the index score to a probability (0-1) score by using the normal cumulative distribution function (cdf). Second, we substitute all values during periods of electoral interruption with zero. This transformed index score could thus be interpreted as the probability of observing the corresponding BFA point estimate score or lower (a BFA score of  $-1.65$  thus corresponds to a probability of .05, a score of 0 to a probability of .5, and so on). Figure C3 (of Appendix C) shows how this transformed 0-1 index (called *v2xel\_frefair*) map onto the original BFA estimates. As can be seen, what the normal cdf accomplishes is a non-linear transformation that slightly compresses the original BFA estimates at the higher and lower ends. We believe this distortion of the original BFA estimate is tolerable.

### *C. Freedom of Organization*

The first two component indices of elected officials and clean elections does not take into account the degree of pluralism. They could at least in theory reach high values even in a totalitarian, single-party state. This is why we need a separate index capturing the degree of what Dahl (1998, 85) calls “associational autonomy,” or freedom of organization. As a component of polyarchy, the core of this construct is party-centered: are political parties free to form; operate autonomously from the ruling regime; and field candidates in national elections? In addition, associational autonomy in the political sphere also requires that there be no barriers to the entry and exit, or state repression of, a wider set of civil society organizations providing alternative means for voice and political activity making policy depend on votes and preferences.<sup>15</sup> We have thus opted for the set of reflective indicators of freedom of association indicated in Table 3. They clearly load on a single underlying dimension with the somewhat weaker loading for civil society repression as the only partial exception.<sup>16</sup> The resulting index (called *v2x\_frasoc*), again rescaled to 0-1 through the normal cdf, is bimodally distributed (see Figure C5).

---

<sup>15</sup> In this particular respect, however, the dividing line between a more minimalist “Schumpeterian”, which would most likely argue in favor of ignoring these additional civil society features, and a more maximalist “Dahlian” concept of freedom of organization cuts within one of Dahl’s five components. For empirical purposes, this distinction makes very little difference however: a thinner freedom of organization index only based on the first four electoral indicators correlates at .99 with the thicker version we opted for.

<sup>16</sup> Since the multiparty elections indicator is only observed in election years, we face the same extrapolation problem as for the clean elections index above. The solution we have adopted (again) is to repeat the values observed in the last election and then replace all observations with 0 during electoral interruptions.

**Table 3. Measuring Freedom of Organization (BFA estimates)**

<i>Indicator</i>	<i>Loadings (<math>\Lambda</math>)</i>	<i>Uniqueness (<math>\Psi</math>)</i>
Party ban (v2psparban)	.898	.194
Barriers to parties (v2psbars)	.911	.170
Opposition parties autonomy (v2psoppaut_fill)	.889	.210
Elections multiparty (v2elmulpar_fill)	.850	.278
CSO entry and exit (v2cseeorgs)	.887	.213
CSO repression (v2csreprss)	.883	.305

*Note:* Entries are factor loadings and uniqueness from a normal theory Bayesian factor analysis model;  $n=16,196$  country-years.

#### *D. Inclusive Citizenship*

For the purpose of measuring electoral democracy, Dahl’s (1971, 1989, 1998) component called “inclusive citizenship” in effect amounts to the extension of the suffrage. The V-Dem indicator is based on an estimate of the proportion of adult citizens eligible to vote, roughly based on the Paxton *et al.* (2003) methodology (see Appendix B, section D, for further details), measured for all country units since 1900. The resulting 0-1 index (called *v2x\_suffr*) is again bimodally distributed (see Figure C6).

#### *E. Freedom of Expression*

Finally, Dahl’s concept of polyarchy includes some non-electoral aspects, most conspicuously having to do with the freedom of expression. Staying true to his widely accepted concept necessitates borrowing some indicators from the “liberal” sphere for this component. A core set of indicators captures media freedom (Behmer 2009) such as active state censorship of print/broadcast media, media self-censorship, and harassment of journalists. In addition, there should be freedom of discussion in society at large, for both men and women (Skaaning 2009).<sup>17</sup> Also included in this index are four indicators more geared towards measuring content rather than government restrictions. Following Dahl (1971) it could be called “alternative sources of information”: whether the media is biased against opposition

---

<sup>17</sup> One could imagine an even more “maximalist” version of the freedom of expression component also including two indicators of physical integrity rights (as is done by, for example, Hadenius 1992). On the one hand, it makes sense to say that a government that practices torture or murder opposition supporters are clearly violating the freedom of expression *de facto*. On the other, one might think of these as two restrictions on what governments can do, and instead be treated as part of the liberal-democratic conception. This is the approach followed here.

parties and candidates, whether major print and broadcast outlets routinely criticize the government, and whether they represent a wide range of political perspectives, as well as general repression of cultural and academic expressions of political dissent.

Table 4 demonstrates that these nine (reduced to eight) indicators are well conceived as partially substitutable, reflective indicators caused by the unobserved latent trait “freedom of expression”. Hence, our index of freedom of expression (called *v2x\_freexp\_thick*) is based on the point estimates from this Bayesian factor analysis model. To allow for simple averaging or multiplication as aggregation rules (see below), we have again converted this index score to a probability (0-1) score by using the normal cdf, resulting in a 0-1 scaled component index distributed as in Figure C7.

**Table 4. Measuring Freedom of Expression (BFA estimates)**

<i>Indicator</i>	<i>Loadings (<math>\Lambda</math>)</i>	<i>Uniqueness (<math>\Psi</math>)</i>
Print/broadcast censorship (v2mecenefm)	.900	.190
Harassment of journalists (v2meharjrn)	.879	.228
Media self-censorship (v2meslfcen)	.887	.214
Freedom of academic/cultural expr. (v2clacfree)	.844	.288
Freedom of discussion (v2cldiscm+v2cldiscw)	.893	.202
Media bias (v2mebias)	.928	.140
Print/broadcast media critical (v2mecrit)	.924	.146
Print/broadcast media perspectives (v2merange)	.920	.153

*Note:* Entries are factor loadings and uniqueness from a normal theory Bayesian factor analysis model; all loadings have been inverted from negative to positive values;  $n=15,594$  country-years.

## 6. Measuring the Whole: Aggregating the Components

Before delving into the question of how these five separately measured components of polyarchy should best be aggregated, it might prove useful to know how they co-vary. The lower diagonal in Table 5 displays the correlation coefficients using the entire set of country-year observations. The covariation is moderate to strong in most instances, but one particular cluster of stronger associations stands out: clean elections, freedom of organization, and freedom of expression. Both the elected executive index and suffrage indicator show systematically lower correlation to these three components. One could suspect this pattern to be the result of the zeros we impose on the clean elections index and on the multiparty elections indicator of the freedom of organization index for periods of electoral interruptions. The upper diagonal in Table 5 therefore presents the correlations for election-years only. As expected, the correlation with the indices of clean elections and freedom of organization becomes

significantly weaker. Yet, the cluster with stronger bivariate correlations remain, separating the elected executive and suffrage components on the one hand from the clean elections, freedom of organization and freedom of expression components on the other.

**Table 5. Correlations among Polyarchy Components**

	Elected Officials	Clean Elections	Freedom of Organization	Suffrage	Freedom of Expression
Elected Officials		.163	.280	.300	.306
Clean Elections	.604		.756	.228	.769
Freedom of Organization	.577	.801		.145	.923
Suffrage	.589	.497	.412		.178
Freedom of Expression	.540	.774	.924	.386	

*Note:* Entries are correlation coefficients, in the lower diagonal for  $n=15,432$  country-years, in the upper diagonal for  $n=3,231$  election years.

The results indicate that polyarchy is multidimensional,<sup>18</sup> as argued by Coppedge & Reinicke (1990) and Coppedge, Alvarez & Maldonado (2008). This in turn complicates the choice of which aggregation rule to apply. Had the correlations between all components been strong and consistent, the exact choice of aggregation rule would matter less. Since the different components demonstrably point in slightly different directions for different countries and time periods, however, the way we combine information from all of them will be of larger consequence for the resulting index value.

As argued above, theoretical considerations should take precedence over empirical in the choice of aggregation rules. The first and most important question is thus whether the five components are reflective indicators being “caused” by an underlying polyarchy trait, or if they are constitutive of that construct. As it turns out, the literature on electoral democracy and how it should be measured is divided on this issue. On one hand, there is a strong rationale in the literature for treating the components as constitutive elements. More precisely, as argued by Przeworski *et al.* (2000) and even

<sup>18</sup> In a Bayesian factor analysis model on these six measures, the fit to a one-dimensional model is decent (but with elected officials and suffrage having weaker loadings). If the model is re-run on election years only, however, the fit is utterly poor, clearly suggesting a two-dimensional solution.

more clearly so by Munck (2009) and Boix *et al.* (2013), these components should be considered as non-compensatory, necessary conditions for polyarchy. In other words, the degree of suffrage is not relevant if there is no freedom of organization, if the election results are completely fabricated, *or* if the executive is not elected. Similarly, the freedom and fairness of elections should not count if only a tiny fraction of the population is enfranchised, and so on. Theoretical guidance thus rule out taking the maximum as the aggregation rule, since that is tantamount to treating them as substitutable and each component as a sufficient condition for polyarchy. Theory also speaks against the use of simple (or weighted) averaging for aggregation, since it implies that high scores on one component at least partially compensates for low scores on another.

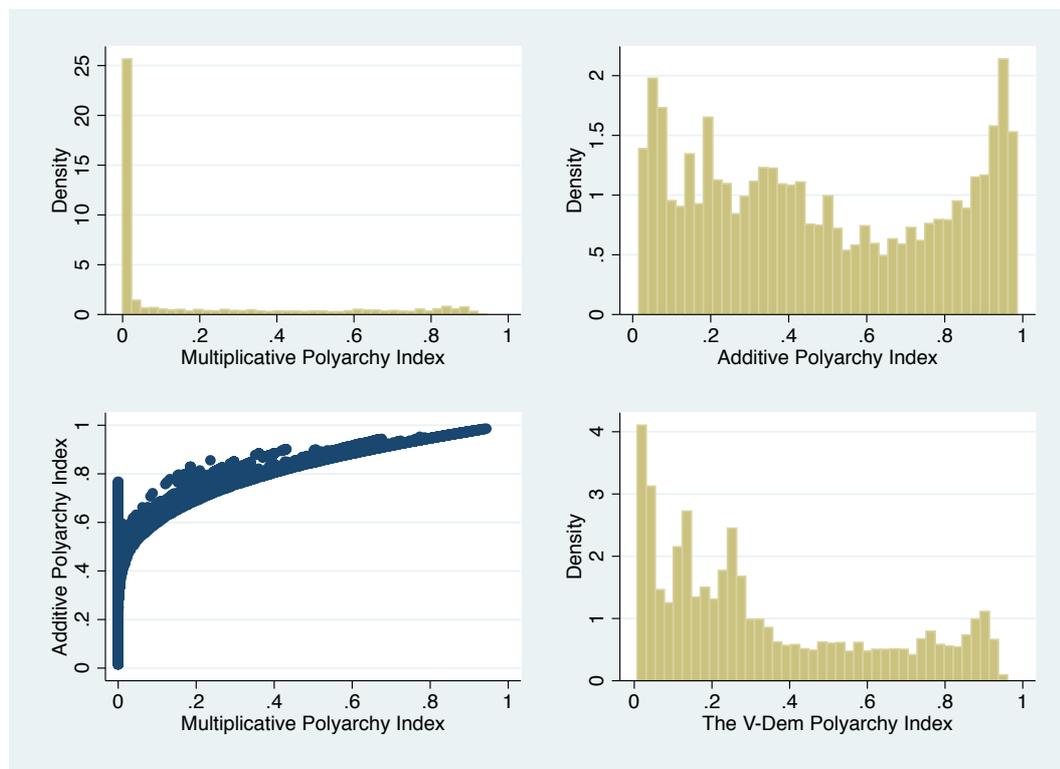
Remains then two viable aggregation rules, which in the literature are typically considered the most appropriate for capturing a set of necessary conditions: either multiplication, as favored by Munck & Verkuilen (2002, 24) and Munck (2009, 32, 40-51), or taking the minimum, as favored by Bowman *et al.* (2005, 956) and Goertz (2006, 111-115). With dichotomous measures these two aggregation rules reduce to the exact same thing, but with graded information this is generally not the case. We agree that from the perspective of fuzzy set/Boolean algebra, the minimum is the most strict interpretation of what a necessary condition implies. But an important drawback is that the value of a single indicator completely determines the aggregated index value. The relative virtue of multiplication is that it combines information from all constitutive elements and hence relies on and retains more information. When all index components are scaled from 0-1, as in our case, it is also fairly easily interpretable, with each component indicator working as a weight variable, as it were, for the others.

According to the “constitutive” logic, then, what might be called the Multiplicative Polyarchy Index (MPI) is constructed as follows:

$$(1) \text{ MPI} = \text{Elected Officials} * \text{Clean Elections} * \text{Freedom of Organization} * \text{Suffrage} * \text{Freedom of Expression}$$

By the logic of multiplication, a low score on any of the component indices suppresses the value of the overall index. As a result, the distribution of this index is heavily skewed towards zero (see the upper left quadrant of Figure 2). The measurement uncertainty of the lower-level component indices are propagated into this component index by the same aggregation rule. Since both Elected Officials and Suffrage are currently assumed to be measured without error, this means that the standard errors from the BFA posteriors of the Clean Elections, the Freedom of Organization and the Freedom of Expression Indices are multiplied.

**Figure 2. Aggregating to Polyarchy**



However, there is also a second well established strand in the literature on how to measure electoral democracy going back to Bollen (1980), Coppedge & Reinicke (1990) and Hadenius (1992). It argues that the polyarchy components as independent reflections (“effects”) of the same underlying construct. From this perspective, the aggregation rule should thus be additive rather than multiplicative. This logic seems to have its strongest theoretical rationale when it comes to the components based on “freedoms” of different sorts, particularly with respect to O’Donnell & Schmitter’s (1986) concept of “liberalization” – the phase in a transition to democracy when the first opening of the authoritarian regime occurs (such as a lifting of media censure, of bans on political parties and wider acceptance for expressions of popular discontent), before the first “founding election” has been held. If the extent to which such “liberalizations” should count is made conditional on the electoral side of the equation – as implied by the multiplicative logic – then we will not be able to observe them in the data. The additive or averaging logic, however, allows such openings to be counted in and of themselves.

Yet as already Dahl (1971) argued, and as has been recently demonstrated also by Coppedge et al. (2008) and Miller (2013), there have been other paths to the installation of partial polyarchy. Another prominent such path has been the introduction of executive elections with universal suffrage, but with severe electoral manipulation, little or no competition at the polls or in society at large, and with severe

repression of the freedom of expression. In effect, it is the existence of this other mostly Marxist-Leninist or post-Communist mode of institutionalizing the trappings of electoral democracy – without its safeguards, as it were – that is driving the low correlations between elected officials and suffrage, on the one hand, and the other three polyarchy components in Table 5. If we were to generalize the additive logic to the full set of components, these two different paths would weigh approximately as much in the resulting index. While recognizing the value of an aggregation rule that also lets the polyarchy components influence the overall score independent of each other, we therefore favor a slightly adjusted approach, by letting the two more “façade-like” components (elected officials and suffrage) together weigh half as much as their more “liberal” and hard-fought counterparts (clean elections, freedom of organization and expression).<sup>19</sup>

An Additive Polyarchy Index (API) is then constructed as follows:

$$(2) \text{ API} = [(Elected\ Officials + Suffrage) + 2 * (Clean\ Elections + Freedom\ of\ Organization + Freedom\ of\ Expression)] / 8 = .125 * Elected\ Officials + .125 * Suffrage + .25 * Freedom\ of\ Organization + .25 * Clean\ Elections + .25 * Freedom\ of\ Expression$$

The resulting index is mostly bimodally distributed, as shown in the upper right quadrant of Figure 2. The measurement uncertainty is again taken into consideration by the same rule, implying that the standard errors from the BFA estimates of Clean Elections are averaged with the standard errors of Freedom of Expression and Freedom of Organization.

The lower left quadrant of Figure 2 contrasts the two aggregation rules. In essence, they discriminate at two different ends of the underlying electoral democracy scale. Thus, what the additive index mostly does is to differentiate different degrees of democracy at the lower end of the multiplicative scale. Even when the latter is zero, the additive index can achieve as high a score as .77. Conversely, the multiplicative index mostly discriminates among countries already achieving high values on the additive scale. Thus, the variation on the multiplicative scale is from .18 to .94 when the additive scale is above .80.

---

<sup>19</sup> The authors wish to thank Carl-Henrik Knutsen, Oslo University, for intellectual assistance in devising this aggregation rule. The assignment of weights can also be supported empirically by fitting a single-factor BFA to the components (see former footnote). The loadings of elected officials and suffrage is then roughly half that of freedom of organization, freedom of expression and freedom of organization (results available upon request).

Since both the multiplicative and the additive logic has support in the literature, and since they evidently have the virtue of discriminating at different ends of the spectrum that one would want to measure in full, we argue that a compromise – more specifically, the average – between the two is the preferred solution.

The V-Dem polyarchy index is thus constructed by averaging (1) and (2), or more precisely:<sup>20</sup>

$$\begin{aligned}
 (3) \text{ Polyarchy} &= \\
 &.5 \text{ MPI} + 0.5 \text{ API} \\
 &= .5(\text{Elected Officials} * \text{Clean Elections} * \text{Freedom of Organization} * \text{Suffrage} * \text{Freedom of Expression}) \\
 &\quad + (1/16 \text{Elected Officials} + 1/8 \text{Clean Elections} + 1/8 \text{Freedom of Organization} + 1/16 \text{Suffrage} + 1/8 \\
 &\quad \text{Freedom of Expression}) \\
 &= .5(\text{Elected Officials} * \text{Clean Elections} * \text{Freedom of Organization} * \text{Suffrage} * \text{Freedom of Expression}) \\
 &\quad + .0625 * \text{Elected Officials} + .125 * \text{Clean Elections} + .125 * \text{Freedom of Organization} + .0625 * \text{Suffrage} \\
 &\quad + .125 * \text{Freedom of Expression}
 \end{aligned}$$

The measurement uncertainty is again propagated into this overall index by averaging the standard errors from the API and MPI. As shown in the lower right quadrant of Figure 2, this index is still positively skewed, but far less so than the multiplicative version.

---

<sup>20</sup> Due to missing data on the media indicators in parts of the world prior to 1946, we have in this index also imputed missing values from a version of the index not based on these indicators.

## 7. Validating the Polyarchy Index

Following McMann, Pemstein, Teorell, Zimmerman and Lindberg (2016), there are two fundamentally different ways in which we could provide positive evidence for the validity of the polyarchy index based on the output of the data generating process.<sup>21</sup> The first is to compare the ratings of different country experts for the same indicators, countries and years; the second to compare these coder-level ratings, as well as the aggregated index at the country-year level, to other similar measures of electoral democracy from other datasets. Under the general heading of “convergent validity” (Campbell and Fiske 1959), agreement among coders or among datasets could both be considered as evidence of measurement validity. Moreover, following Donchev and Ujhelyi (2014), understanding the sources of disagreement can provide an additional tool for assessing validity.

### *Comparing coders*

Assuming that our experts exercises independent judgment when translating their perceptions of the world into numerical ratings (Schedler 2012), coder agreement could be interpreted as a sign of validity.

As Steenbergen & Marks put it:

What would happen if our worst fears about expert judgments were to materialize? We would find different experts judging different objects, on different dimensions, at different points in time... Absent such conditions, high *variance* would be the necessary consequence of experts basing their judgments on different foundations. As a corollary, the correlation between expert judgments would be reduced. The key to assessing expert judgments, then, is to assess the variance in those judgments (2007, 351).

In this spirit, Table 6 decomposes the variance in the 23 expert-coded indicators (out of the 36 in total) underlying the polyarchy index: clean elections (8), freedom of organization (6), and freedom of expression (9). Results are pooled across indicators for each component, controlling for country- and year-fixed effects. Expressed as standard deviations, the random effects are measured on the same ordinal 0-4 scale as the original indicators to ease interpretation. Overall they do not suggest conspicuously high levels of disagreement among our country experts. One should also recall that a decent share of this “raw” coder disagreement is known to be a reflection of varying thresholds for how to map perceptions on to the ordinal coding categories in the survey, and that the measurement

---

<sup>21</sup> A third approach is “case face validation” (McMann et al. 2016), that is, to compare the ratings with the qualitative literature on democracy in a particular country or set of countries. This is an approach not pursued in this paper.

model corrects a substantial share of this variance when aggregating these raw scores to the country-year level (Pemstein et al. 2015).

**Table 6. Variance decomposition of coder disagreement**

	<i>Clean elections</i>	<i>Freedom of organization</i>	<i>Freedom of expression</i>
Grand mean	1.40*** (.038)	2.23*** (.060)	1.65*** (.059)
Coder-random effects	.511*** (.015)	.409*** (.019)	.544*** (.013)
Indicator-random effect	.802*** (.007)	.904*** (.009)	.591*** (.005)
No of indicators	8	6	9
No of experts	1292	1863	1687
No of observations	288,910	438,269	725,251

\*  $p < .10$ , \*\*  $p < .05$ , \*\*\*  $p < .01$

Note: Entries are variance components expressed as standard deviations, with standard errors in parentheses. Country- and year-fixed effects included but omitted from the table. Estimates obtained from the “mixed” command in Stata13.

Granted that country experts disagree at times, is there meaningful variation in that incongruity? Arguably, we should expect coders to disagree more, the more difficult the coding task (Steenbergen and Marks 2007; Coma and van Ham 2015). In the V-Dem data, we may distinguish at least three sources of difficulty (McMann et al. 2016). First, all else equal, judging indicators of electoral democracy should be more difficult further back in time, when experts cannot rely as much on their academic and lived experience. Second, and relatedly, difficulty should increase when there is little independent information from which to form ones judgment. Third, extreme conditions should be more easily identifiable than the “muddled middle”, hence coders should be likely to disagree more at intermediate levels of whatever aspect of electoral democracy they are assessing. This third expectation may work counter to the first to the extent that the situation in many countries was clearly “bad” in terms of democratic aspects in the early years of the 20<sup>th</sup> century.

In Table 7, we test these three propositions by regressing the country-year standard deviation in coder ratings, pooled across the same 23 expert-coded indicators as in Table 6. We present results by component controlling for indicator-fixed effects (not displayed in the table). We also control for the

number of coders that may in itself be a systematic driver of the amount of disagreement.<sup>22</sup> All three expectations are borne out by evidence. Coders disagree more when coding electoral democracy indicators further back in time, when there is less available media information, and at intermediate levels of the indicators they are assessing. The only exception is the insignificant estimate for year for the clean elections indicators. We believe this is because election quality is difficult to assess whatever the slice of history (at least when holding the amount of available information constant). We conclude that the level of coder disagreement varies in a meaningful and predictable way, something that lends further support for the validity of the V-Dem data.

**Table 7. Predicting coder disagreement**

	<i>Clean elections</i>	<i>Freedom of organization</i>	<i>Freedom of expression</i>
	(1)	(2)	(3)
Year	-.000 (.000)	-.001*** (.000)	-.001** (.000)
Media access	-.003*** (.001)	-.003*** (.001)	-.003*** (.000)
Level	.053*** (.011)	-.052*** (.010)	.030*** (.008)
Level <sup>2</sup>	-.112*** (.006)	-.097*** (.005)	-.083*** (.003)
No of coders	.006** (.003)	.010** (.005)	.019*** (.005)
Mean standard deviation	.860	.804	.794
No of indicators	8	6	9
Adjusted R-squared	.289	.331	.334
No of countries	171	173	173
No of observations	49,390	80,450	139,046

\*  $p < .10$ , \*\*  $p < .05$ , \*\*\*  $p < .01$

*Note:* Entries are regression coefficients, with Standard errors, clustered on countries, in parentheses. Indicator-fixed effects included but omitted from the table.

As a final test comparing coder estimates, following Dahlström, Lapuente and Teorell (2012), we assess if coder characteristics systematically affect ratings. These characteristics, deriving from the V-Dem post-survey questionnaire that all country experts were asked to fill out, are both sociological

<sup>22</sup> This is particularly important when considering the lateral coding performed by a large number of coders for many countries, all in 2012. This means that the number of coders for most countries is higher, or much higher for that particular year.

and ideological in nature. We control for gender (female=1, male=0), age (in years, also including the squared polynomial), level of education (PhD=1, 0=less than PhD), employment (1=government, 0=other), whether the expert was born in or resides in the country he or she is coding (1=yes, 0=no), as well as whether the expert is of “Western” origin or not. Moreover, we control for three potential sources of ideological bias. The first is support for free market economy as a proxy for left-right leaning. Second, we look at two measures of support for different normative understandings of democracy. The first we call the “conventional understanding”, which entails support for electoral and liberal democracy. The second we call the “alternative understanding”, which entails support for participatory, liberal or egalitarian types of democracy.<sup>23</sup> In order to fix comparisons across coders to the same countries and years, we also control for country- and year-fixed effects (results omitted).

**Table 8. Predicting coder ratings**

	<i>Clean elections</i>		<i>Freedom of organization</i>		<i>Freedom of expression</i>	
	(1)	(2)	(3)	(4)	(5)	(6)
Gender	-.079 (.050)	-.018 (.046)	-.012 (.041)	-.029 (.041)	-.060 (.041)	-.090*** (.033)
Age	-.017 (.013)	-.029* (.015)	-.004 (.011)	.002 (.012)	.004 (.011)	.004 (.009)
Age <sup>2</sup>	.000 (.000)	.000** (.000)	.000 (.000)	-.000 (.000)	.000 (.000)	-.000 (.000)
PhD education	-.037 (.056)	-.028 (.056)	-.020 (.039)	-.041 (.047)	-.039 (.038)	-.028 (.036)
Government employee	-.105 (.088)	-.104 (.089)	.087 (.080)	.089 (.080)	-.085 (.077)	.082 (.059)
Born in country	.180*** (.065)	.145** (.070)	.085* (.044)	.082 (.053)	.060 (.051)	.068 (.050)
Resides in country	.020 (.062)	.005 (.063)	-.064* (.036)	-.058 (.045)	.027 (.048)	-.023 (.043)
Born in Western country	-.021 (.070)	.007 (.084)	-.083 (.051)	-.004 (.060)	-.041 (.063)	-.041 (.055)
Free market support	.019 (.021)	.030 (.023)	.022* (.013)	-.007 (.015)	.031* (.016)	.006 (.019)
Conventional understanding of democracy	-.006 (.023)	-.014 (.053)	-.001 (.015)	-.042 (.034)	-.002 (.016)	-.022 (.032)
Alternative understanding of democracy	-.008 (.022)	.013 (.050)	-.029* (.015)	.035 (.038)	-.033** (.016)	-.016 (.031)
Western country		.135* (.050)		-.319*** (.038)		-.174 (.031)

<sup>23</sup> This grouping of the five principles of democracy is supported by a principal component factors analysis run at the coder level (results available upon request).

		(.072)		(.085)		(.129)
Born in Western country		-.084		.115		.125
×Western country		(.121)		(.091)		(.114)
Openness to trade		.024*		-.020*		-.002
		(.014)		(.010)		(.010)
Free market support		-.006*		.002		-.002
×Openness to trade		(.003)		(.003)		(.003)
Conventional democracy		1.921***		2.890***		2.819***
score		(.322)		(.315)		(.261)
Conventional under-		.027		.056		.060
standing ×score		(.085)		(.060)		(.049)
Alternative democracy		1.835***		-.243		.903***
score		(.356)		(.349)		(.258)
Alternative under-		-.075		-.105*		-.025
standing ×score		(.075)		(.062)		(.045)
Country-fixed effects?	Y	N	Y	N	Y	N
R-squared	.211	.458	.104	.299	.161	.545
No of countries	173	147	174	149	174	149
No of observations	245,039	160,998	383,763	228,557	635,567	370,224

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

*Note:* Entries are regression coefficients, with standard errors, clustered on countries, in parentheses. Year and indicator-fixed effects included but omitted from the table. Models 1, 3 and 5 includes country-fixed effects.

The results in models 1, 3 and 5 are overall comforting. With one exception, coder characteristics are not related to how V-Dem experts rate electoral democracy indicators. The exception is a tendency among coders born in a country to provide a more positive assessment of election quality in that country than non-native experts (independent of country of residence). This is an interesting pattern worthy of further inquiry, but it is the exception to the rule: experts do not provide different assessments of election quality, freedom of association or of expression, depending on who they are.

Models 2, 4, and 6, tests for another type of coder bias that Bollen and Paxton (2000, 72) labels “situational closeness”. This is the idea that “judges will be influenced by how situationally and personally similar a country is to them.” More specifically, one could imagine that strong believers in the free market has no general bias in their rankings but a specific tendency to rate countries with free markets better. Similarly, situational closeness bias implies that strong supports of liberal-electoral democracy would give overly positive ratings of countries that are strong on liberal-electoral democracy; or that believers in other conceptions of democracy rank countries more proximate to their ideals more favorably also on the electoral indicators. Finally, this kind of bias could manifest such that even if Western coders display no systematic bias in and of themselves, bias might be directed toward

non-Western countries. This translates into controlling for interaction effects between particular coder and country characteristics.<sup>24</sup>

We find no evidence of situational bias in the V-Dem data. The two tendencies that come closest to statistical significance at conventional levels are that market believers tend to rate elections as of *lower* quality in countries with more openness to trade (the latter is our proxy for free market establishment in practice). This is in the opposite direction from “situational closeness.” The second is that supporters of alternative (participatory/deliberative/egalitarian) forms of democracy tend to rate countries that are strong on such dimensions lower in terms of freedom of organization. The dominant pattern is that situational closeness does not introduce bias in the V-Dem experts’ ratings of electoral democracy.

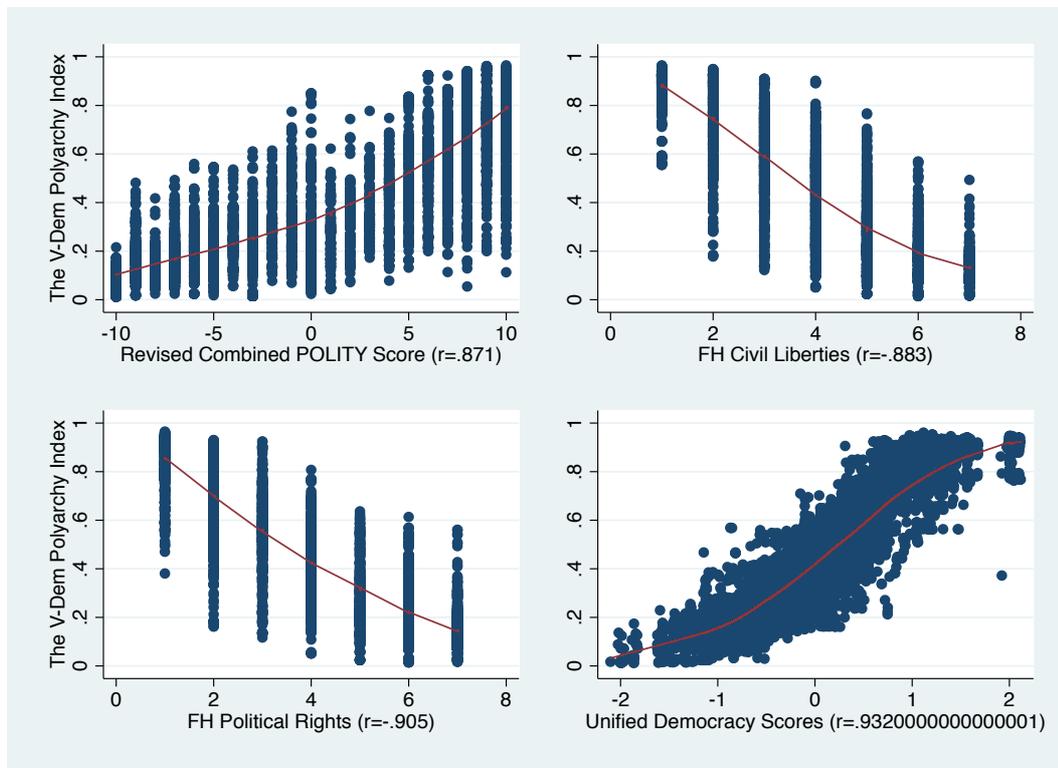
### *Comparing datasets*

Our second set of validation exercises focus on comparisons between our measure of electoral democracy and extant alternatives. Figure 4 displays the bivariate descriptive pattern at country-year level, comparing the V-Dem polyarchy index to the Polity and Freedom House ratings, as well as the “Unified Democracy scores” (UDS) respectively. The latter is based on just about every pertinent measure of electoral democracy covering multiple years and countries (Pemstein et al. 2010). Convergence could be said to be the overall pattern. The pairwise correlations range from .85 for Polity, .87 for Freedom House civil liberties, .90 for Freedom House political rights, to .93 for the UDS ratings. The red smoothed lowess-line of best fit also indicates a consistent pattern of monotonically increasing levels of polyarchy, the higher the extant measure of electoral democracy. In the convergent validity sense, these different measures thus validate each other.

---

<sup>24</sup> This means that we have to drop the country-fixed effects from these models

**Figure 3. Comparing Polyarchy with Alternative Electoral Democracy Indices**



*Note:* The red lines are smoothed lowess regression lines of best fit.

However, one could also argue that the many and sometimes fairly large deviations is more eye-opening. In particular Polity and Freedom House regularly produces scores that differs significantly from the V-Dem polyarchy index. Evidence presented in Table 9 refute two potential explanations for these deviations that, if they were true, would cast doubt over the V-Dem measure. The first is the potential effects of coder characteristics, but now in terms of the extent to which differences between the V-Dem polyarchy measure and Polity or Freedom House can be explained by the aggregate composition of coders for each country, estimated as the average coder characteristics at the country-year level. To avoid potential ecological fallacy bias (Huckfeldt and Sprague 1993), we also control for the individual-level tendencies unraveled in Table 8. Second, we assess whether differences can be explained by three other prime suspects of the V-Dem data that varies by country and year: the average level of disagreement among coders, the number of coders, and the number of bridge/lateral coders. To the extent these features of the data collection process systematically predicts deviations from Polity and Freedom House, we have a reason to worry about a biased data generating process.

**Table 9. Predicting deviations from Polity and Freedom House**

	<i>Polity</i>		<i>Freedom House</i>	
	<i>Raw scores</i> (1)	<i>Absolute</i> (2)	<i>Raw scores</i> (3)	<i>Absolute</i> (4)
Polity score	1.949*** (.076)	-.135*** (.028)		
Freedom House score			2.204*** (.064)	-.225*** (.028)
<i>Coder characteristics:</i>				
Gender	-.050** (.024)	.027** (.012)	-.058** (.023)	.015 (.011)
Age	-.000 (.007)	-.003 (.003)	-.002 (.007)	.000 (.003)
Age <sup>2</sup>	.000 (.000)	.000 (.000)	.000 (.000)	.000 (.000)
PhD education	-.041* (.025)	.007 (.013)	-.045* (.024)	-.005 (.013)
Government employee	.014 (.046)	.019 (.027)	-.001 (.044)	.016 (.021)
Born in country	.116*** (.036)	.002 (.017)	.073** (.035)	.017 (.016)
Resides in country	-.011 (.036)	.027 (.017)	.016 (.032)	.017 (.014)
Born in Western country	-.008 (.035)	-.070*** (.019)	-.062** (.027)	-.058*** (.017)
Free market support	.012 (.011)	-.001 (.005)	.002 (.011)	.001 (.005)
Conventional under- standing of democracy	.004 (.010)	.008 (.005)	.004 (.011)	.002 (.005)
Alternative under- standing of democracy	-.042*** (.010)	.008 (.005)	-.021** (.010)	.006 (.005)
<i>Coder composition:</i>				
Share female coders	-.024 (.125)	-.084** (.040)	-.107 (.100)	-.047 (.038)
Average age of coders	.000 (.030)	.009 (.015)	.034 (.031)	-.004 (.012)
Average age <sup>2</sup>	.000 (.000)	-.000 (.000)	-.000 (.000)	-.000 (.000)
Share of PhD coders	.133 (.086)	-.060 (.042)	-.033 (.082)	-.022 (.032)
Share of coders employed by government	.054 (.182)	-.060 (.076)	-.055 (.170)	-.066 (.071)
Share of coders born in country	.114 (.135)	.068 (.047)	-.010 (.095)	.054 (.039)
Share of coders residing in country	.253** (.102)	-.039 (.036)	.226** (.091)	-.110*** (.041)
Share of Western coders	.160* (.084)	.048 (.031)	.149*** (.057)	.009 (.026)

Average free market support	.007 (.043)	.004 (.013)	.041 (.032)	-.035** (.014)
Average conventional understanding	-.047 (.057)	-.019 (.019)	-.104** (.048)	-.007 (.018)
Average alternative understanding	-.038 (.045)	-.026 (.016)	-.057* (.035)	-.003 (.014)
<i>Country characteristics:</i>				
Coder disagreement	.044 (.050)	.223*** (.015)	.054 (.052)	.214*** (.015)
No of coders	.002 (.003)	-.002** (.001)	-.000 (.002)	-.003*** (.001)
No of lateral coders	.017** (.009)	-.003 (.005)	.007 (.007)	.003 (.004)
R-squared	.399	.159	.426	.190
No of countries	163	163	168	168
No of observations	930,161	930,161	529,367	529,367

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

*Note:* Entries are regression coefficients, with standard errors, clustered on countries, in parentheses. Year and indicator-fixed effects included but omitted from the table.

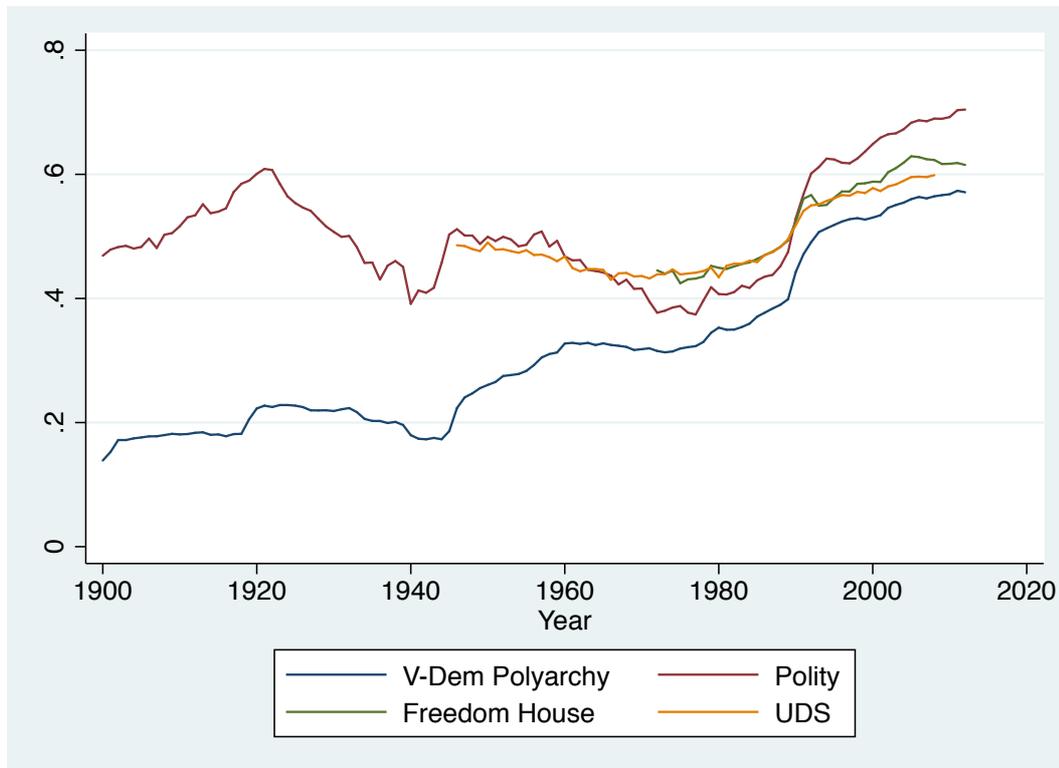
Overall the evidence is pointing clearly to negative findings. We do not focus on individual-level characteristics in this analysis, since they simply replicate the results from Table 8 but with increased precision (mostly due to the larger sample size). Admittedly, when controlling for these individual-level tendencies, there are a few significant aggregate coder composition effects. When a large share of coders are born in the countries they are coding, V-Dem rates countries as less democratic than both Polity and Freedom House, but with respects to Freedom House this is offset by a tendency to produce more similar scores in absolute terms. A large share of coders with alternative conceptions of democracy rank countries as less democratic than Polity, but again the substantive effect is miniscule (recall that the dependent variable is measured on a 0-4 scale). Coder disagreement at the country-year level is one of the few positive predictors of absolute deviations, both from Polity and Freedom House. This is noteworthy, but more V-Dem coders shrinks the distance. On the whole, evidence clearly suggests that divergences between the V-Dem polyarchy index scores and Polity and Freedom House are not due to the composition of our sample of coders, or how these coders assess electoral democracy.

## 8. The Added Value: Some Empirical Illustrations

We first showcase the V-Dem polyarchy index's development over time in comparison to some of the most popular extant indices (here rescaled to vary from 0 to 1). As shown in Figure 4, the new index track all of Huntington's (1991) three waves of democracy, although with differences not seen before. First, the first wave in the early 19<sup>th</sup> century is much less pronounced. This is mainly a result of the fact that we treat semi-sovereign territories such as colonies on par with independent states. This is an important corrective to our understanding of the level and development of democracy in the world. Polity (along with all other historical regime datasets we are aware of) portrays the state of democracy only in the part of the world that consisted of independent states. As we know, in the early 20<sup>th</sup> century a vast majority of the world's populations lived under colonial rule. V-Dem's measure of polyarchy reflects also how oppressive these empires were, not only the level of electoral democracy of the core, independent states. This also provides a unique opportunity that no other dataset offers. We can for the first time analyze moves toward independence as processes of democratization (which they often were).

Second, in contrast to other indices, the V-Dem polyarchy index show a much more gradual second wave stretching into the 1960s rather than ending abruptly after the Second World War, and display very few signs of a reversed second wave. Rather, the level of electoral democracy in the world seemed to have been leveled out prior to the start of the third wave in the 1980s. The reasons for this need yet to be fully explored but both the fact that V-Dem measures also semi-sovereign territories and that the polyarchy index captures aspects beyond the partial, electoralist components covered by Polity are prime suspects.

**Figure 4. Aggregate Trends in the Major Democracy Indices**



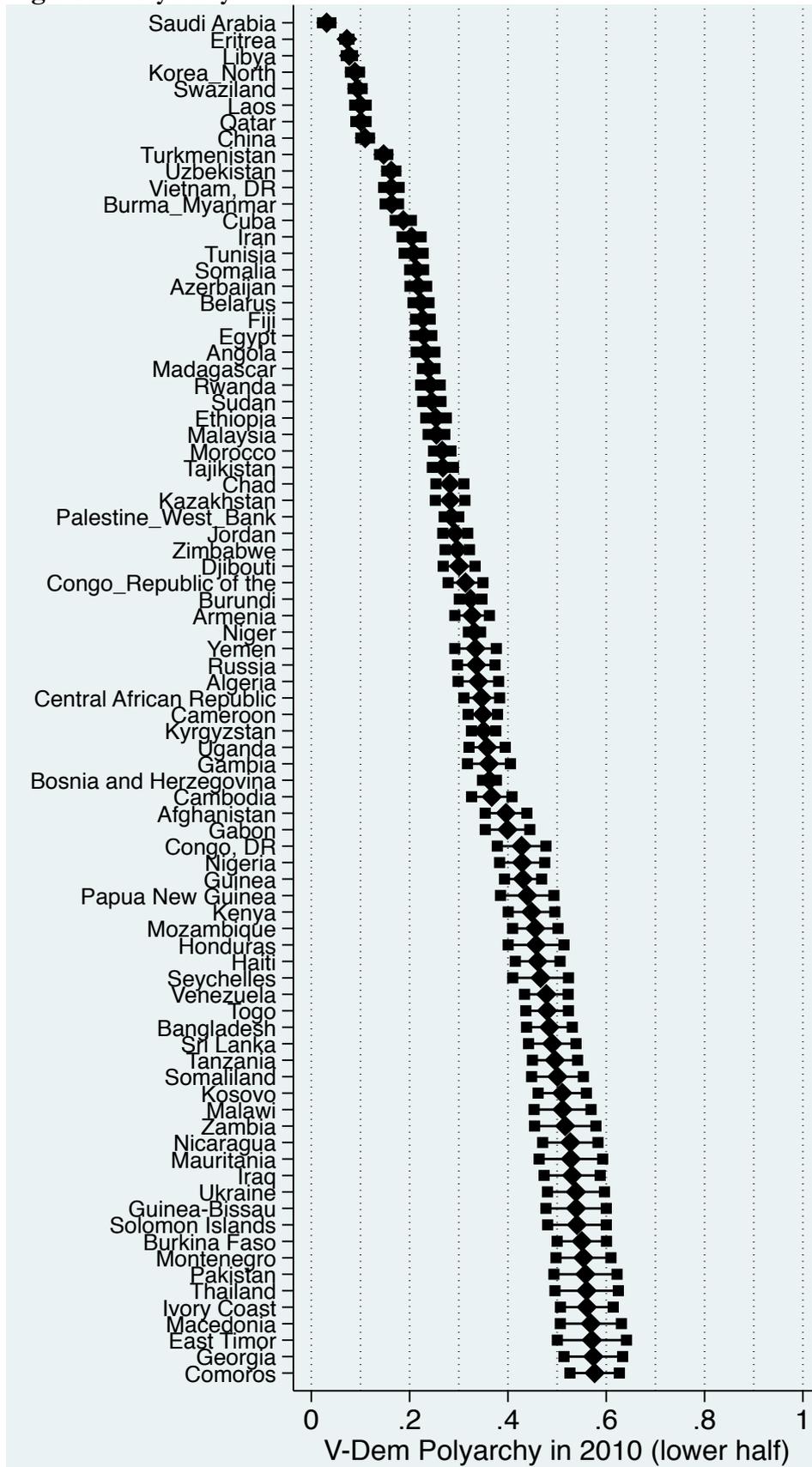
*Note:* Polity and Freedom House have been rescaled to vary from 0-1 by using their theoretical maxima and minima. UDS, lacking such theoretically defined maxima and minima, has been converted using the empirical counterparts.

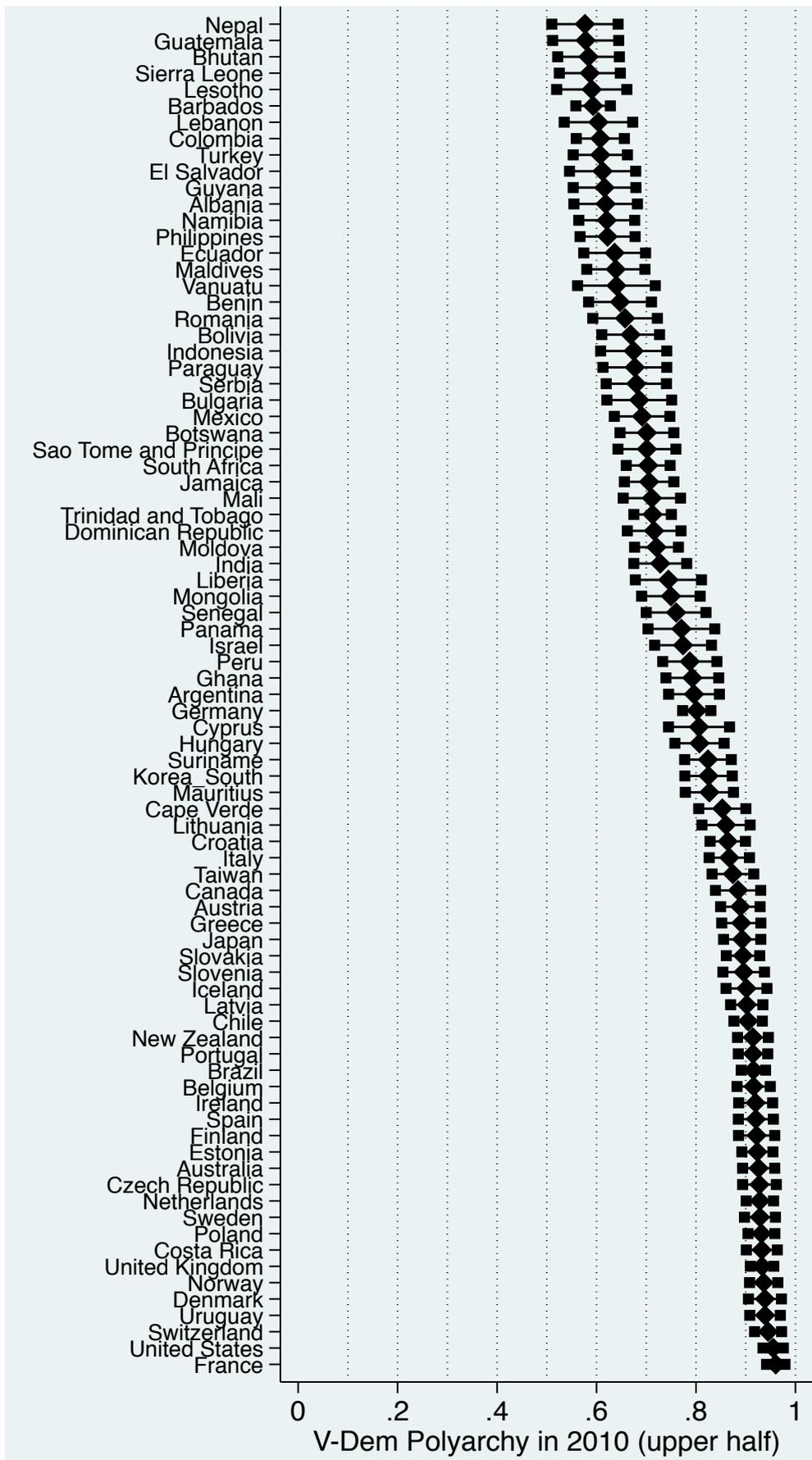
As compared to Polity, the V-Dem polyarchy measure tends to come out as a more conservative measure of electoral democracy. Our inclusion of a suffrage component that is much more demanding most likely explains part of this (as is the case of Boix, Miller and Rosato 2012), yet differences persist even today and seems also affected by the V-Dem index having a higher standard for in particular free and fair elections. As compared to Freedom House, the probably second most used graded indicator of electoral democracy, the most notable difference is that the V-Dem polyarchy index does not display a downward trend in the last decade. The more exact reasons for this also deserves further study.

An added feature of the V-Dem polyarchy index, as noted above, is that we not only provide point estimates but also uncertainty estimates. In Figure 5, we display the ranking of countries in our sample in the year of 2010, from lowest to highest degree of polyarchy, together with 1 standard deviation confidence intervals. In this particular year, Saudi Arabia, Eritrea, Libya, North Korea, Swaziland, Laos, Qatar and China score at the very bottom, whereas Switzerland, United States and France are ranked highest (in a large group of mostly Western established democracies that by the confidence intervals are indistinguishable in terms of polyarchy scores). The margins of measurement

error tend to be largest in the middle (cf. Treier and Jackman 2008), which both makes intuitive sense (these are more unstable conditions) and in line with the finding reported above that coders disagree more in the “muddled middle”.

Figure 5. Polyarchy in 2010

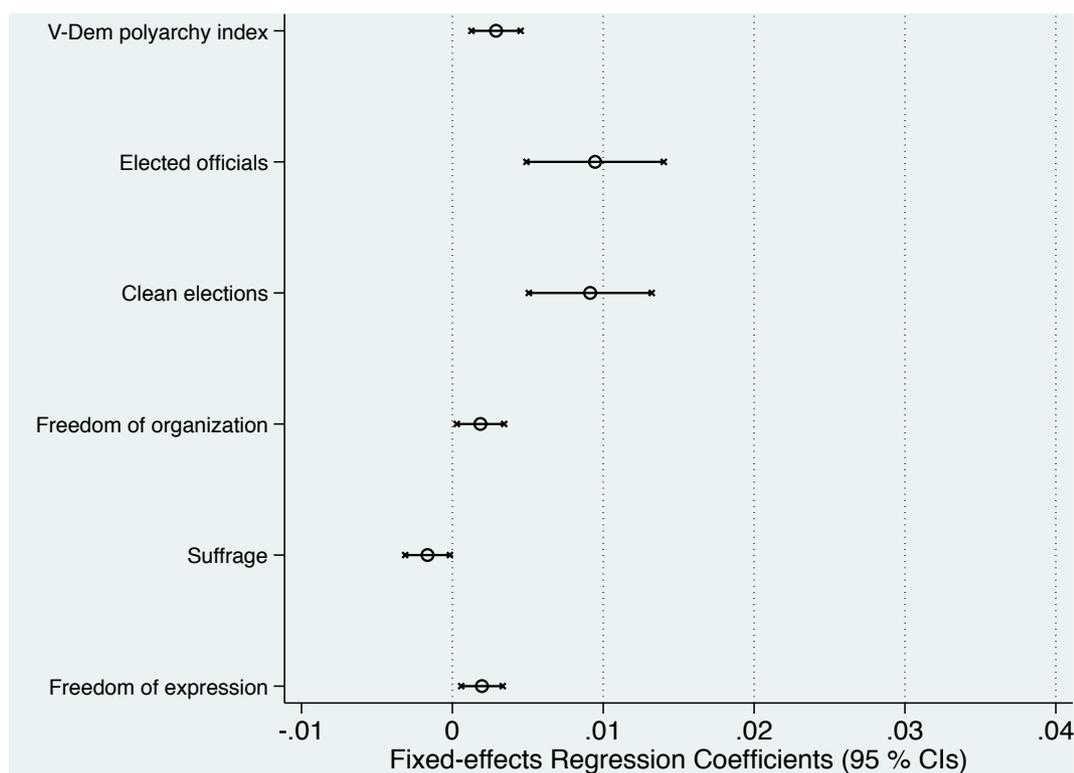




As a final illustration of the added value of our new Polyarchy index, we will highlight the potential insights gained from disaggregation. A long-standing debate in the social sciences concerns whether economic development, as originally hypothesized by Lipset (1959), is a driver of democratic development. Probably due to the lack of appropriate, disaggregated data very little attention has been paid to the possibility that economic development could be driving certain components of electoral democracy, but not others (for a notable exception, see Aidt & Jensen 2014; 2016). If we unbundle the different components of electoral democracy, do we get different results with respect to the overall modernization effect?

In Figure 6, we show the results of a series of dynamic panel regressions with this purpose in mind. To address potential endogeneity issues, each model includes one lag of the dependent variable, as well as the lag of  $\ln(\text{GDP}/\text{capita})$  (data from Madison). We also control for both country- and year-fixed effects to address concerns of omitted variable bias.<sup>25</sup>

**Figure 6. Economic development, polyarchy and its components**



*Note:* Entries are regression coefficients for the effects of lagged  $\ln(\text{GDP}/\text{capita})$ , controlling for year- and country-fixed effects and one lagged dependent variable.

<sup>25</sup> With an average time series components of 55 years per country, Nickel bias should not be of great concern (Beck & Katz 2011).

Concentrating on the short-term effects of  $\ln(\text{GDP}/\text{capita})$ , the results show a significant general effect on polyarchy in this sample of countries. When we “drill down” one level, however, we find that this result is driven by two electoral components of the polyarchy index – elected officials and free and fair elections – whereas the freedom components (freedom of organization and expression) are only weakly linked to economic development, and suffrage even negatively so. In other words, modernization does not seem to be driving “liberalization”, in O’Donnell & Schmitter’s (1986) terms, but rather more full “transitions” to establishing the fuller set of electoral institutions. The more exact reasons for this are discussed in more detail elsewhere (Knutsen et al. 2015).

## Conclusion

This paper presents a fresh index of electoral democracy. We argue that it has several advantages over existing measures typically used in academic as well as policy analyses. The new V-Dem index is firmly anchored in democratic theory capturing all necessary institutions following Dahl’s (1971, 1989, 1998) theory of polyarchy that we argue has almost universal acceptance today. Second, the five components with their respective indices and their constituent, detailed and precise 36 indicators makes it possible to drill down into what exactly drives changes in scores. This transparency and detail also makes it possible for the first time to disaggregate analyses of factors such as economic development’s relationships to different components of democracy, and it also allow end users to instead draw on more minimalist conceptions of electoral democracy should that be more motivated by their research questions.

The V-Dem measure of electoral democracy relies on a unique data generating process. Over 2,600 academic and other experts provided ratings in their special areas of expertise for particular countries according to a fully detailed protocol. Aggregation of country expert ratings is done using state-of-the-art Bayesian ordinal item-response theory modeling addressing many thorny issues regarding cross-country, and cross-rater comparability while also making it possible for V-Dem to provide measures of uncertainty associated with each country-year observation. Bayesian factor analysis is employed also to aggregation of indicators to component indices, preserving adequate information of measurement uncertainty.

Finally, several tests of validity of the V-Dem polyarchy index strengthen our belief that it is by and large not driven by biases stemming from coder characteristics or ideological predilections. There is a fair amount of coder agreement for individual country-year indicators, and deviations from that

pattern can be meaningfully understood. We find strong correlation with other existing measures of electoral democracy, but also meaningful differences, that cannot be easily explained by features of the V-Dem data generating process such as the composition of coders for particular countries.

It is our hope that this new resource will prove useful to scholars to bring forward new and more robust knowledge on the causes and consequences of democratization as well as democracy. It is a tall agenda and one that we as an academic community have been involving ourselves in for 60 years now. We are not so delusional that we think this new measure will lead to the solutions for all our research problems, but we hope it will contribute to clearing some.

## References

- Aidt, Toke and Peter Sandhold Jensen (2014) "Workers of the world, unite! Franchise extensions and the threat of revolution in Europe, 1820-1938", *European Economic Review* 72: 52-75.
- Aidt, Toke and Peter Sandhold Jensen (2016) "From Open to Secret Ballot: Vote Buying and Modernization", *Comparative Political Studies*, forthcoming.§
- Beck, Nathaniel & Jonathan Katz (1996) "Nuisance vs. Substance: Specifying and Estimating Time-Series Cross-Section Models", *Political Analysis* 6(1): 1–36.
- Beck, Nathaniel & Jonathan Katz (2011) "Modeling Dynamics in Time-Series Cross-Sectional Political Economy Data", *Annual Review of Political Science* 14:331-352.
- Behmer, Markus (2009) "Measuring Media Freedom: Approaches of International Comparison". In Czepek, A., M. Hellwig & Eva Novak (eds.), *Press Freedom in Europe: Concepts and Conditions*. Bristol & Chicago: Intellect, pp. 23-34.
- Bermeo, Nancy (2010) "Interests, Inequality, and the Illusion in the Choice for Fair Elections." *Comparative Political Studies*, 45(8-9), pp. 1119-1147
- Birch, Sarah (2008) "Electoral Systems and Electoral Misconduct." *Comparative Political Studies*, 40(12), 1533-1556.
- Birch, Sarah (2011) *Electoral Malpractice*. New York: Oxford University Press.
- Boix, Carles, Michael Miller, and Sebastian Rosato (2012) "A Complete Data Set of Political Regimes, 1800-2007", *Comparative Political Studies* 46(12): 1523-1554.
- Bollen, Kenneth (1980) "Issues in the Comparative Measurement of Political Democracy", *American Sociological Review*, 370-90.
- Bollen, Kenneth (1989) *Structural Equations with Latent Variables*. New York: John Wiley.
- Bollen, Kenneth (1990) "Political Democracy: Conceptual and Measurement Traps", *Studies in Comparative International Development* 25(1): 7-24
- Bollen, Kenneth (1993) "Liberal Democracy: Validity and Method Factors in Cross-National Measures", *American Journal of Political Science* 37(4): 1207-1230.
- Bollen, Kenneth and Pamela Paxton (2000) "Subjective Measures of Liberal Democracy", *Comparative Political Studies* 33(1): 58-86
- Bowman, Kirk, Fabrice Lehoucq and James Mahoney (2005) "Measuring Political Democracy: Case Expertise, Data Adequacy, and Central America", *Comparative Political Studies* 38(8): 939-970
- Calingaert, Daniel (2006) "Election Rigging and How to Fight It." *Journal of Democracy*, 17(3), 138-151.
- Campbell, Donald T., and Donald W. Fiske (1959) "Convergent and discriminant validation by the multitrait-multimethod matrix", *Psychological Bulletin* 56: 81-105.
- Coma, Ferran Martínez and Carolien van Ham (2015) "Can Experts judge elections? Testing the validity of expert judgments for measuring election integrity" *European Journal of Political Research* 54(2): 305-325.
- Coppedge, Michael and Wolfgang Reinicke (1990) "Measuring Polyarchy", *Studies in Comparative International Development* 25(1): 51–72.
- Coppedge, Michael, Angel Alvarez, and Claudia Maldonado (2008) "Two Persistent Dimensions of

- Democracy: Contestation and Inclusiveness.” *Journal of Politics* 70(3): 335–50.
- Coppedge, Michael, John Gerring, Staffan I. Lindberg, Svend-Erik Skaaning, and Jan Teorell. 2015a. *V-Dem Comparisons and Contrasts with Other Measurement Projects*. Varieties of Democracy (V-Dem) Project.
- Coppedge, Michael, John Gerring, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Carl Henrik Knutsen, Kyle Marquardt, Kelly McMann, Farhad Miri, Pamela Paxton, Daniel Pemstein, Jeffrey Staton, Eitan Tzelgov, Yi-ting Wang, and Brigitte Zimmerman. 2015b. *V-Dem [Country-Year/Country-Date] Dataset v5*. Varieties of Democracy (V-Dem) Project
- Coppedge, Michael, John Gerring, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, with David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Carl Henrik Knutsen, Kelly McMann, Pamela Paxton, Daniel Pemstein, Jeffrey Staton, Brigitte Zimmerman, Frida Andersson, Valeriya Mechkova, and Farhad Miri. 2015c. *V-Dem Codebook v5*. Varieties of Democracy (V-Dem) Project.
- Coppedge, Michael, John Gerring, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, Frida Andersson, Kyle Marquardt, Valeriya Mechkova, Farhad Miri, Daniel Pemstein, Josefine Pernes, Natalia Stepanova, Eitan Tzelgov, and Yi-ting Wang. 2015d. *V-Dem Methodology v5*. Varieties of Democracy (V-Dem) Project.
- Coppedge, Michael, Staffan I. Lindberg, Svend-Erik Skaaning, and Jan Teorell. *n.d.* 2016. “Measuring High Level Democratic Principles Using the V-Dem Data”, *International Political Science Review* 37(2).
- Dahl, Robert (1956) *A Preface to Democratic Theory*. Chicago: The University of Chicago Press
- Dahl, Robert (1971) *Polyarchy*. New Haven & London: Yale University Press.
- Dahl, Robert (1989) *Democracy and Its Critics*. New Haven & London: Yale University Press.
- Dahl, Robert (1998) *On Democracy*. New Haven & London: Yale University Press.
- Diamond, Larry (2002) “Thinking about Hybrid Regimes”, *Journal of Democracy* 13: 21-35
- Donchev, Dilyan and Gergely Ujhelyi (2014) “What do corruption indices measure?”, *Economics & Politics* 26: 309-331
- Donno, Daniela (2013) *Defending Democratic Norms. International Actors and the Politics of Electoral Misconduct*. Oxford: Oxford University Press.
- Downs, Anthony (1957) *An Economic Theory of Democracy*. Boston: Addison-Wesley.
- Elgie, Robert (1998) ”The Classification of democratic regime types: Conceptual ambiguity and contestable assumptions,” *European Journal of Political Research* 33: 219-238
- Gasiorowski, Mark (1996) “An Overview of the Political Regime Change Dataset”, *Comparative Political Studies* 29(4): 469-483.
- Goertz, Gary (2006) *Social Science Concepts: A User’s Guide*. Princeton: Princeton University Press.
- Hadenius, Axel (1992) *Democracy and Development*. Cambridge: Cambridge University Press.
- Hadenius, Axel & Jan Teorell. “Assessing Alternative Indices of Democracy”, *Concepts & Methods Working Papers* 6, IPSA, August 2005.
- Huckfeldt, Robert and John Sprague (1993) “Citizens, contexts, and politics”, in Finifter, A., ed. *Political Science: The State of the Discipline, II*. APSA: The American Political Science Association.

- Huntington, Samuel (1991) *The Third Wave: Democratization in the Late 20th Century*. Norman: University of Oklahoma Press.
- Karl, Terry (1986) "Imposing Consent? Electoralism versus Democratization in El Salvador," in *Elections and Democratization in Latin America, 1980-1985*, eds. Paul Drake and Eduardo Silva (San Diego: Center for Iberian and Latin American Studies, Center for US/Mexican Studies, University of California, San Diego), 9-36.
- Kelley, Judith (2013) *Monitoring Democracy: When international election observation works, and why it often fails*. New Jersey: Princeton University Press.
- Knutsen, Carl Henrik, John Gerring, Svend-Erik Skaaning, Jan Teorell, Matthew Maguire, Michael Coppedge and Staffan I. Lindberg (2015) "[Economic Development and Democracy: An Electoral Connection](#)", V-Dem Working Paper Series 2015:16, The Varieties of Democracy Institute, University of Gothenburg.
- Lehoucq, Fabrice (2003) "Electoral Fraud: Causes, Types, and Consequences." *Annual Review of Political Science*, 6: 233–256.
- Lindberg, Staffan I., Michael Coppedge, John Gerring, and Jan Teorell. 2014. "A New Way to Measuring Democracy". *Journal of Democracy* 25(3): 159-169.
- Mainwaring, Scott, Daniel Brinks and Aníbal Pérez-Liñán (2001) "Classifying Political Regimes in Latin America, 1945-1999" *Studies in Comparative International Development*, 36(1): 37-65
- McMann, Kelly, Daniel Pemstein, Brigitte Seim, Jan Teorell and Staffan Lindberg (2016) "Strategies of Validation: Assessing the Varieties of Democracy Corruption Data", *V-Dem Institute Working Paper Series 2016:23*, The Varieties of Democracy Institute, University of Gothenburg.
- Miller, Michael (2013) "Democratic Pieces: Autocratic Elections and Democratic Development since 1815", *British Journal of Political Science* 45: 501-530.
- Møller, Jørgen and Svend-Erik Skaaning (2010) "Beyond the Radial Delusion: Conceptualizing and Measuring Democracy and Non-Democracy", *International Political Science Review* 31(3): 261-283.
- Munck, Gerardo (2009) *Measuring Democracy: A Bridge Between Scholarship & Politics*. Baltimore: The Johns Hopkins University Press.
- Munck, Gerardo and Jay Verkuilen (2002). "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices", *Comparative Political Studies*, 35(1): 5-34.
- Norris, Pippa (2015) *Why Elections Fail*. New York: Cambridge University Press.
- O'Donnell, Guillermo (2001) "Democracy, Law, and Comparative Politics", *Studies in Comparative International Development* 36(1): 7-36.
- O'Donnell, Guillermo and Philippe Schmitter (1986) *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies*. Baltimore & London: The Johns Hopkins University Press
- Paxton, Pamela, Kenneth A. Bollen, Deborah M. Lee, and Hyojoung Kim (2003) "A Half-Century of Suffrage: New Data and a Comparative Analysis." *Studies in Comparative International Development* 38(1): 93–122.
- Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi-ting Wang and Farhad Miri. 2015. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data." *V-Dem Institute Working Paper Series 2015: 21*, December 2015
- Pemstein, Daniel, Stephan Meserve, and James Melton (2010) "Democratic compromise: A latent variable analysis of ten measures of regime type", *Political Analysis* 18: 426-449.

- Przeworski, Adam, Michael Alvarez, José Antonio Cheibub and Fernando Limongi (2000). *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990*. Cambridge: Cambridge University Press.
- Schedler, Andreas (2002) "The Menu of Manipulation," *Journal of Democracy* 13(2): 36-50
- Schumpeter, Joseph (1942) *Capitalism, Socialism & Democracy*. New York & London: George Allen & Unwin.
- Siaroff, Alan (2003) "Comparative presidencies: The inadequacy of the presidential, semi-presidential and parliamentary distinction," *European Journal of Political Research* 42: 287-312
- Simpser, Alberto (2013) *Why Governments and Parties Manipulate Elections: Theory, Practice, and Implications*. Cambridge: Cambridge University Press.
- Skaaning, Svend-Erik (2009) "Measuring Civil Liberty: An Assessment of Standards-Based Data Sets," *Revista de Ciencia Política* 29(3):721-740.
- Steenbergen, Marco and Gary Marks (2007) "Evaluating expert judgements", *European Journal of Political Research* 46: 347-366
- Treier, Shawn and Simon Jackman (2008) "Democracy as a Latent Variable", *American Journal of Political Science* 52(1): 201-217
- van Ham, Carolinen and Lindberg n.d.(2016) "Choosing from the menu of manipulation: Explaining incumbents' choices of electoral manipulation tactics", University of Gothenburg, V-Dem Institute: V-Dem Working Paper Series, No. 26.
- Vanhanen, Tatu (2000) "A New Dataset for Measuring Democracy, 1810-1998". *Journal of Peace Research*, 37: 251-65

## Appendix A

### The sample of fully completed V-Dem countries and coding periods

Afghanistan, 1900-2012  
Albania, 1912-2012  
Algeria, 1900-2012  
Angola, 1900-2012  
Argentina, 1900-2012  
Armenia, 1990-2012  
Australia, 1901-2012  
Austria, 1918-1938, 1945-2012  
Azerbaijan, 1990-2012  
Bangladesh, 1971-2012  
Belarus, 1990-2012  
Belgium, 1900-2012  
Benin, 1900-2012  
Bhutan, 1900-2012  
Bolivia, 1900-2012  
Bosnia and Herzegovina, 1992-2012  
Botswana, 1900-2012  
Brazil, 1900-2012  
Burkina Faso, 1919-1932, 1947-2012  
Burma/Myanmar, 1900-2012  
Burundi, 1916-2012  
Cambodia, 1900-2012  
Cameroon, 1961-2012  
Canada, 1900-2012  
Cape Verde, 1900-2012  
Central African Republic, 1920-2012  
Chad, 1920-2012  
Chile, 1900-2012  
China, 1900-2012  
Colombia, 1900-2012  
Comoros, 1900-1945, 1991-2012  
Congo, Democratic Republic, 1900-2012  
Congo, Republic, 1903-2012  
Costa Rica, 1900-2012  
Croatia, 1941-1945, 1991-2012  
Cuba, 1902-2012  
Cyprus, 1900-2012  
Czech Republic, 1918-2012  
Denmark, 1900-2012  
Djibouti, 1900-2012  
Dominican Republic, 1900-2012  
Ecuador, 1900-2012  
Egypt, 1900-2012  
El Salvador, 1900-2012  
Eritrea, 1900-2012  
Estonia, 1918-1940, 1990-2012  
Ethiopia, 1900-2012  
Fiji, 1900-2012  
Finland, 1900-2012  
France, 1900-2012  
Gabon, 1910-2012  
Gambia, 1900-2012  
Georgia, 1990-2012  
German Democratic Republic, 1945-1990  
Germany, 1900-1945, 1949-2012  
Ghana, 1902-2012  
Greece, 1900-2012  
Guatemala, 1900-2012  
Guinea, 1900-2012  
Guinea-Bissau, 1900-2012  
Guyana, 1900-2012  
Haiti, 1900-2012  
Honduras, 1900-2012  
Hungary, 1918-2012  
Iceland, 1900-2012  
India, 1900-2012  
Indonesia, 1900-2012  
Iran, 1900-2012  
Iraq, 1920-2012  
Ireland, 1919-2012  
Israel, 1948-2012  
Italy, 1900-2012  
Ivory Coast, 1900-2012  
Jamaica, 1900-2012  
Japan, 1900-2012  
Jordan, 1922-2012  
Kazakhstan, 1990-2012  
Kenya, 1900-2012  
Korea, North, 1945-2012  
Korea, South, 1900-2012  
Kosovo, 1999-2012  
Kyrgyzstan, 1990-2012  
Laos, 1900-2012  
Latvia, 1920-1940, 1990-2012  
Lebanon, 1918-2012  
Lesotho, 1900-2012  
Liberia, 1900-2012  
Libya, 1934-1942, 1951-2012

Lithuania, 1918-1940, 1990-2012  
 Macedonia, 1991-2012  
 Madagascar, 1900-2012  
 Malawi, 1900-2012  
 Malaysia, 1900-2012  
 Maldives, 1900-2012  
 Mali, 1900-2012  
 Mauritania, 1904-2012  
 Mauritius, 1900-2012  
 Mexico, 1900-2012  
 Moldova, 1990-2012  
 Mongolia, 1911-2012  
 Montenegro, 1900-1918, 1998-2012  
 Morocco, 1912-2012  
 Mozambique, 1900-2012  
 Namibia, 1900-2012  
 Nepal, 1900-2012  
 Netherlands, 1900-2012  
 New Zealand, 1900-2012  
 Nicaragua, 1900-2012  
 Niger, 1922-2012  
 Nigeria, 1914-2012  
 Norway, 1900-2012  
 Pakistan, 1947-2012  
 Palestine/British Mandate, 1918-1948  
 Palestine/Gaza, 1948-1967, 2007-2012  
 Palestine/West Bank, 1948-1950, 1967-2012  
 Panama, 1903-2012  
 Papua New Guinea, 1900-2012  
 Paraguay, 1900-2012  
 Peru, 1900-2012  
 Philippines, 1900-2012  
 Poland, 1918-1939, 1944-2012  
 Portugal, 1900-2012  
 Qatar, 1900-2012  
 Romania, 1900-2012  
 Russia, 1900-2012  
 Rwanda, 1916-2012  
 Sao Tome and Principe, 1900-2012  
 Saudi Arabia, 1932-2012  
 Senegal, 1904-2012  
 Serbia, 1900-2012  
 Seychelles, 1903-2012  
 Sierra Leone, 1900-2012  
 Slovakia, 1939-1945, 1993-2012  
 Slovenia, 1989-2012  
 Solomon Islands, 1900-2012  
 Somalia, 1900-2012  
 Somaliland, 1900-1960, 1991-2012  
 South Africa, 1900-2012  
 South Sudan, 2011-2012  
 South Yemen, 1900-1990  
 Spain, 1900-2012  
 Sri Lanka, 1900-2012  
 Sudan, 1900-2012  
 Suriname, 1900-2012  
 Swaziland, 1900-2012  
 Sweden, 1900-2012  
 Switzerland, 1900-2012  
 Syria, 1918-1920, 1922-2012  
 Taiwan, 1900-2012  
 Tajikistan, 1990-2012  
 Tanzania, 1900-2012  
 Thailand, 1900-2012  
 Timor-Leste, 1900-2012  
 Togo, 1916-2012  
 Trinidad and Tobago, 1900-2012  
 Tunisia, 1900-2012  
 Turkey, 1900-2012  
 Turkmenistan, 1990-2012  
 Uganda, 1900-2012  
 Ukraine, 1990-2012  
 United Kingdom, 1900-2012  
 United States, 1900-2012  
 Uruguay, 1900-2012  
 Uzbekistan, 1990-2012  
 Vanuatu, 1906-2012  
 Venezuela, 1900-2012  
 Vietnam, Democratic Republic, 1945-2012  
 Vietnam, Republic of, 1902-1975  
 Yemen, 1918-2012  
 Zambia, 1911-2012  
 Zimbabwe, 1900-2012

# Appendix B

## List of V-Dem variables used in operationalizing polyarchy

There are three different types of V-Dem indicators: Type (A) data is more factual and coded by Research Assistants connected to the project under the supervision of Principal Investigators and Project Managers. Type (A+B) data is gathered from country-specific sources by Country Coordinators under the supervision of centrally employed Research Assistants. Type (C) data requires a larger degree of judgment about the state of affairs obtaining in a particular country at a particular point in time, and is therefore coded by multiple Country Experts.

*A. The elected officials index (HOS=head of state, HOG=head of government):*

v2expathhs/hg – **HOS/HOG appointment in practice (A+B)**: How did the head of state/government reach office? (HOS: 0=coup/rebellion, appointed by 1=foreign power, 2=ruling party [in one-party system], 3=royal council, 4=hereditary succession, appointed by 5=the military, 6=legislature, 7=directly elected, 8=other; HOG: 0=coup/rebellion, appointed by 1=foreign power, 2=ruling party [in one-party system], 3=royal council, 4=hereditary succession, appointed by 5=the military, 6=head of state, 7=legislature, 8=directly elected, 9=other)

v2exaphos/hog – **HOS/HOG selection by legislature in practice (A+B)**: Was approval of the legislature necessary for the appointment of the head of state/government? (0=no, 1=yes)

v2exhoshog – **HOS = HOG (A\*)**: Is the head of state (HOS) also head of government (HOG)? (0=no, 1=yes)

v2exdfcbhs\_rec/v2exdjcbhg – **HOS/HOG appoints cabinet in practice (C)**: In practice, does the head of state/government have the power to appoint – or is the approval of the head of state/government necessary for the appointment of – cabinet ministers? (HOS: 0=no, 1=yes, but only with respect to the head of the cabinet, and only with the tacit consent or explicit confirmation by the legislature; or yes, but only with the tacit consent or explicit confirmation by the legislature, 2=yes, without any need for confirmation by the legislature, but only with respect to the head of the cabinet; or yes, without any need for confirmation by the legislature; HOG: 0=no, 1= yes, but only with the tacit consent or explicit confirmation by the legislature, 2=yes, without any need for confirmation by the legislature)

v2exdfdmhs/v2exdfdsgh – **HOS/HOG dismisses ministers in practice (C)**: If the head of state/government took actions to dismiss cabinet ministers, would he/she be likely to succeed? (0=no, 1= yes, but not at his/her own discretion, only when prompted to as a response to specific events, 2=yes, at his/her own discretion, but with restrictions, 3=yes, at his/her own discretion and without restrictions)

v2lgbicam – **Legislature bicameral (A\*)**: How many chambers does the legislature contain? (0=0, 1=1, 2=2, 3=other, 4=unable to determine, 5-6=NA)

v2lgdcmchm – **Legislature dominant chamber (C)**: If the legislature is bicameral, which chamber is dominant? (0=lower clearly more dominant, 1=lower somewhat more on most issues, 2=roughly co-equal, 3=upper somewhat more on most issues, 4=upper clearly more dominant)

v2lgeello/v2lgelecup – **Lower/Upper chamber elected (B/A)**: What percentage of the lower/upper chamber of the legislature is directly elected in popular elections? (%)

*B. The clean elections index:*

v2elembaut – **EMB autonomy (C)**: Does the Election Management Body (EMB) have autonomy from government to apply election laws and administrative rules impartially in national elections? (0=no, controlled by ruling body, 1=some autonomy on some issues but not critical ones, 2=some autonomy but also partial, 3=acts impartially almost all the time, 4=autonomously and impartially applies election laws and administrative rules)

v2elembcap – **EMB capacity (C)**: Does the Election Management Body (EMB) have sufficient staff and resources to administer a well-run national election? (0=glaring deficits in staff, financial, or other resources, 1=deficits not glaring but seriously compromise organization, 2=might be serious deficiencies but could also be product of human error, 3=partial deficits in resources but not serious or widespread, 4=adequate staff and other resources)

v2elrgstry – **Election voter registry (C)**: In this national election, was there a reasonably accurate voter registry in place and was it used? (0=no registry used, 1=fundamentally flawed registry/20% or more eligible voters could have been disenfranchised, 2=uncertain if potential flaws had impact, 3=imperfect registry but less than 10% of eligible voters may have been disenfranchised, 4=reasonably accurate/less than 1% were affected by any flaws)

v2elvotbuy – **Election vote buying (C)**: In this national election, was there evidence of vote

and/or turnout buying? (1=systematic, widespread by almost all parties/candidates, 2=non-systematic but common, 3=money and/or personal gifts distributed by parties, 4=limited use of money and gifts, 5=no evidence of vote/turnout buying)

v2elirreg – **Election other voting irregularities (C)**: In this national election, was there evidence of other intentional irregularities by incumbent and/or opposition parties, and/or vote fraud? (0=systematic and almost nationwide, 1=non-systematic but rather common, 2=limited and sporadic, 3=limited and many probably unintentional, 4=no evidence of intentional irregularities)

v2elintim – **Election government intimidation (C)**: In this national election, were opposition candidates/parties/campaign workers subjected to repression, intimidation, violence, or harassment by the government, the ruling party, or their agents? (0=strong intimidation and repression during election period, 1=systematic, frequent and violent harassment of opposition, 2=periodic, not systematic, but possibly centrally coordinated harassment, 3=sporadic violent harassment directed at only 1 or 2 branches of opposition, 4=no harassment or intimidation)

v2elpeace – **Election other electoral violence(C)**: In this national election, was the campaign period, election day, and post-election process free from other types (not by the government, the ruling party, or their agents) of violence related to the conduct of the election and the campaigns (but not conducted by the government and its agents)? (0=widespread violence occurring throughout election period, or intense period of more than a week resulting in large number of deaths or displaced persons, 1=violence resulting in a few deaths or persons forced to move temporarily, 2=outbursts of violence limited to a day or two, 3=few isolated acts involving few people, 4=no election-related violence)

v2elfrfair – **Election free and fair (C)**: Taking all aspects of the pre-election period, election day, and the post-election process into account, would you consider this national election to be free and fair? (0=not at all, 1=not really/some irregularities may have affected outcome, 2=substantial competition but some significant irregularities, 3=deficiencies and irregularities but did not in the end affect outcome, 4=some human error but largely unintentional without consequences)

*C. The freedom of organization index:*

v2psparban – **Party ban (C)**: Are any parties banned? (0=all parties except state-sponsored party, 1=elections are non-partisan/no official parties, 2=many parties, 3=only a few, 4=none)

v2psbars – **Barriers to parties (C)**: How restrictive are the barriers to forming a party?

(0=parties not allowed, 1=impossible for non government-affiliated parties to form, 2=significant obstacles, 3=modest barriers, 4=no substantial barriers)

v2psoppaut – **Opposition parties autonomy (C)**: Are opposition parties independent and autonomous of the ruling regime? (0=not allowed, 1=opposition parties selected or co-opted by ruling regime, 2=some independent, autonomous opposition parties, 3=most significant opposition parties are autonomous, 4=all opposition parties are independent)

v2elmulpar – **Elections multiparty (C)**: Was this national election multiparty? (0=no-party or single-party, 1=no-party or single-party but multiple candidates from same party contest legislative seats or presidency, 2=at least one real opposition party allowed but competition constrained, 3=multiparty but competition prevented for at least one opposition party or by conditions such as civil unrest, 4=multiparty elections, though a few marginal parties may not be permitted)

v2cseeorgs – **CSO entry and exit (C)**: To what extent does the government achieve control over entry and exit by civil society organizations (CSOs) into public life? (0=monopolistic control/only government-sponsored orgs allowed to engage in political activity, repression of those who defy, 1=substantial control/government licenses all CSOs, active repression of those who defy, 2=moderate control/at least some orgs play an active political role, government does not or cannot repress them, 3=minimal control/constitutional provisions to ban anti-democratic movements, 4=unconstrained/government does not impede formation and operation)

v2csreprss – **CSO repression (C)**: Does the government attempt to repress civil society organizations (CSOs)? (0=no/free to organize, 1=weakly/government uses material sanctions, 2=moderately/material sanctions and minor legal harassment, 3=substantially/material sanctions, minor legal harassments, and arrests of oppositional CSO participants acting lawfully, 4=severely/violently and actively pursues all members of CSOs)

#### *D. Suffrage:*

V2x\_suffr – **Suffrage Index (A)**: This question applies to citizens only and does not take into consideration restrictions based on age, residence, citizenship, having been convicted for crime, being legally incompetent, or belonging to particular occupational groups such as the clergy, the armed forces, or election officials. It covers legal (de jure) restrictions, not restrictions that may be operative in practice (de facto). Universal suffrage is coded as 100%. Universal male suffrage is coded as 50%. The absence of an electoral regime in the year (as defined v2elecreg) is coded

0%. If qualifying criteria other than gender apply such as property, tax payments, income, education, region, race, ethnicity, religion, and/or 'economic independence', rough estimates are reported. They are based on the number and character of qualifying criteria that are generally translated into percentages in the following ways (if only male suffrage): property/income/taxes and education = 5%; property/income/taxes = 10%; education or property/income/taxes = 20%; 'economic dependency = 40%. If available, numbers of eligible or registered voters and information on population distribution are used to qualify the estimates.

*C. The freedom of expression index:*

v2mecenefm – **Print/broadcast censorship effort(C)**: Does the government directly or indirectly attempt to censor the print or broadcast media? (0=attempts are routine, 1=attempts are limited to especially sensitive issues, 2=rarely attempts to censor major media in any way, and if attempts are discovered, the responsible officials are usually punished.)

v2meharjrn – **Harassment of journalists(C)**: Are individual journalists harassed - i.e., threatened with libel, arrested, imprisoned, beaten, or killed -- by governmental or powerful nongovernmental actors while engaged in legitimate journalistic activities? (0=No journalists dare engage in journalistic activities that would offend powerful actors, 1=some journalists engage, but most are almost always forced to stop, 2=rare for any journalist to be harassed, and if they were, those responsible for harassment would be identified and punished, 3=journalists are never harassed by governmental or powerful non-governmental actors)

v2meslfcen – **Media self-censorship(C)**: Is there self-censorship among journalists when reporting on issues that the government considers politically sensitive? (0= complete and thorough, 1= common but incomplete, 2= occurs only on a few highly sensitive political issues, 3=little or none)

v2cldiscm/v2cldiscw – **Freedom of discussion for men/women(C)**: Are men/women able to openly discuss political issues in private homes and in public spaces? (0=not respected/harsh and immediate intervention, 1=weakly respected/frequently exposed to intervention or harassment, 2=somewhat respected/occasionally exposed to intervention or harassment, 3=mostly respected/minor restraints on expression in private sphere, linked to few cases or to soft sanctions, 4=fully respected/unrestricted freedom of expression) – these two items are strongly correlated ( $r=.81$ ) and hence averaged into *v2x\_cldisc* before proceeding.

v2clacfree – **Freedom of academic and cultural expression (C)**: Is there academic freedom and freedom of cultural expression related to political issues? (0=not respected/censorship and

intimidation are frequent, 1=weakly respected/practiced occasionally but direct criticism of government mostly repressed, 2=somewhat respected/practiced routinely but strong criticism of government sometimes repressed, 3=mostly respected/few limitations and sanctions infrequent and soft, 4=fully respected/no restrictions)

v2mebias – **Media bias (C)**: Is there media bias against opposition parties or candidates? (0=coverage of only official party or candidates, no coverage, or no opposition to cover, 1=cover more than one official party but all opposition receives only negative coverage, 2=cover opposition, but give only negative or no coverage to at least one newsworthy party or candidate, 3=cover opposition, but give an exaggerated amount of coverage to governing party, 4=cover all newsworthy parties and candidates more or less impartially and in proportion to newsworthiness)

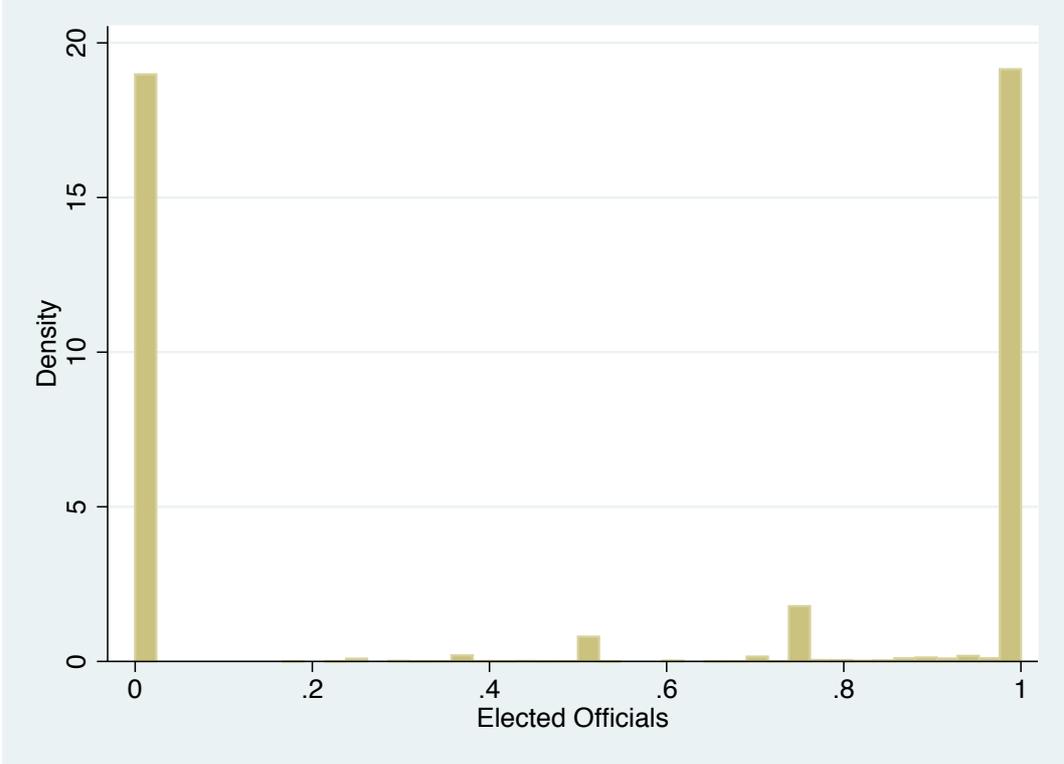
v2mecrit – **Print/broadcast media critical (C)**: Of the major print and broadcast outlets, how many routinely criticize the government? (0=none, 1=a few marginal outlets, 2=some important outlets but some important outlets that never do, 3=all major media outlets at least occasionally)

v2merange – **Print/broadcast media perspectives (C)**: Do the major print and broadcast media represent a wide range of political perspectives? (0=only the government's perspective, 1=represents only perspectives of government and government-approved, semi-official opposition party, 2=a variety of political perspectives represented but systematically ignore at least one that is important to society, 3=all perspectives that are important in this society are represented in at least one major media)

# Appendix C

## Distributions of component indices\*

Figure C1. The Elected Officials Index ( $n=16,376$  country years)



---

\* The following graphs displaying distributions were produced using v4 of the V-Dem dataset, but there would be no substantive differences using v5.

Figure C2. The Clean Elections Index ( $n=3,281$  election years)

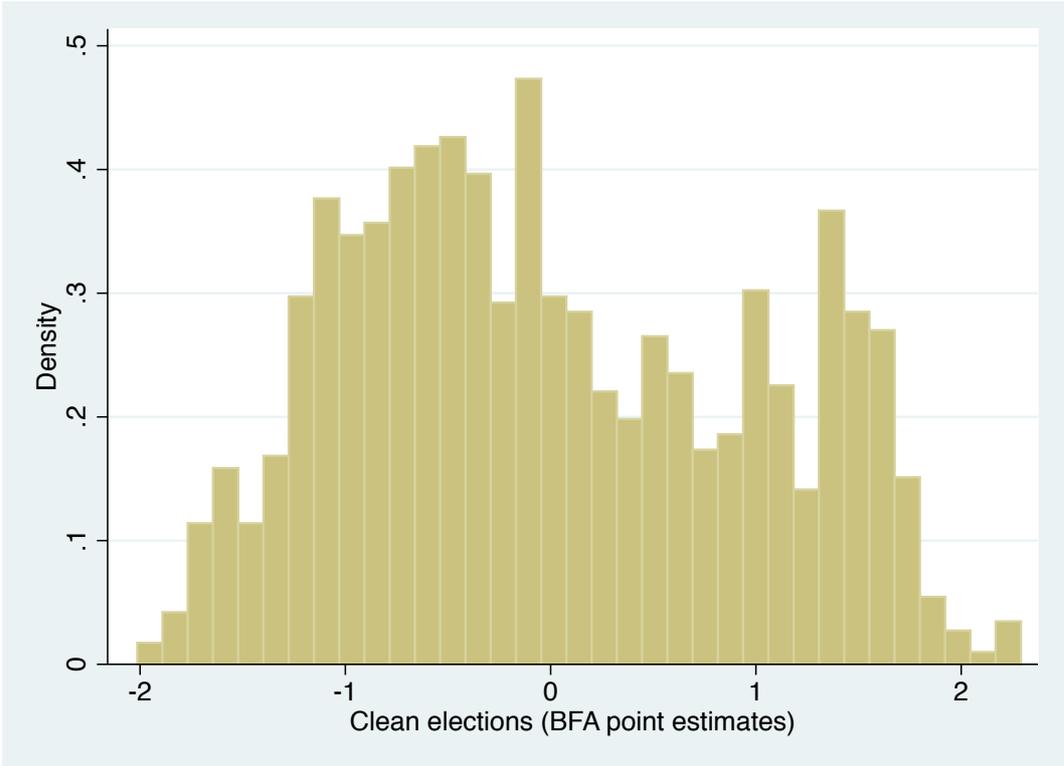


Figure C3. Transforming the Clean Elections Index

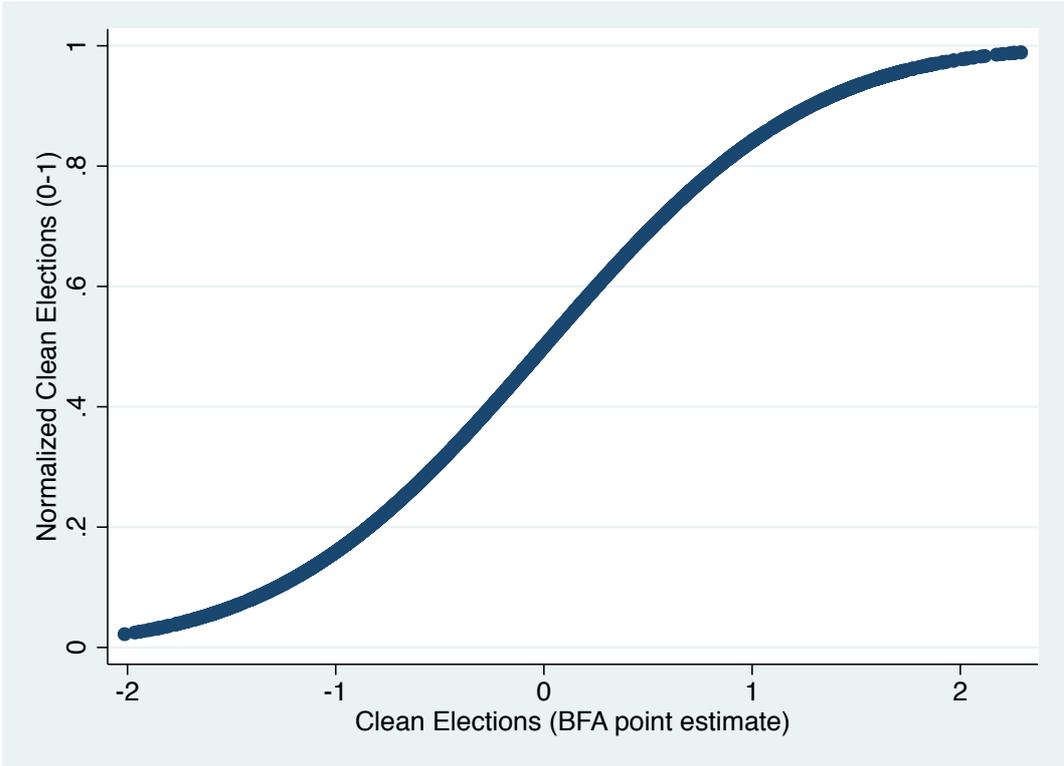


Figure C4. The Normalized Clean Elections Index ( $n=16,317$ )

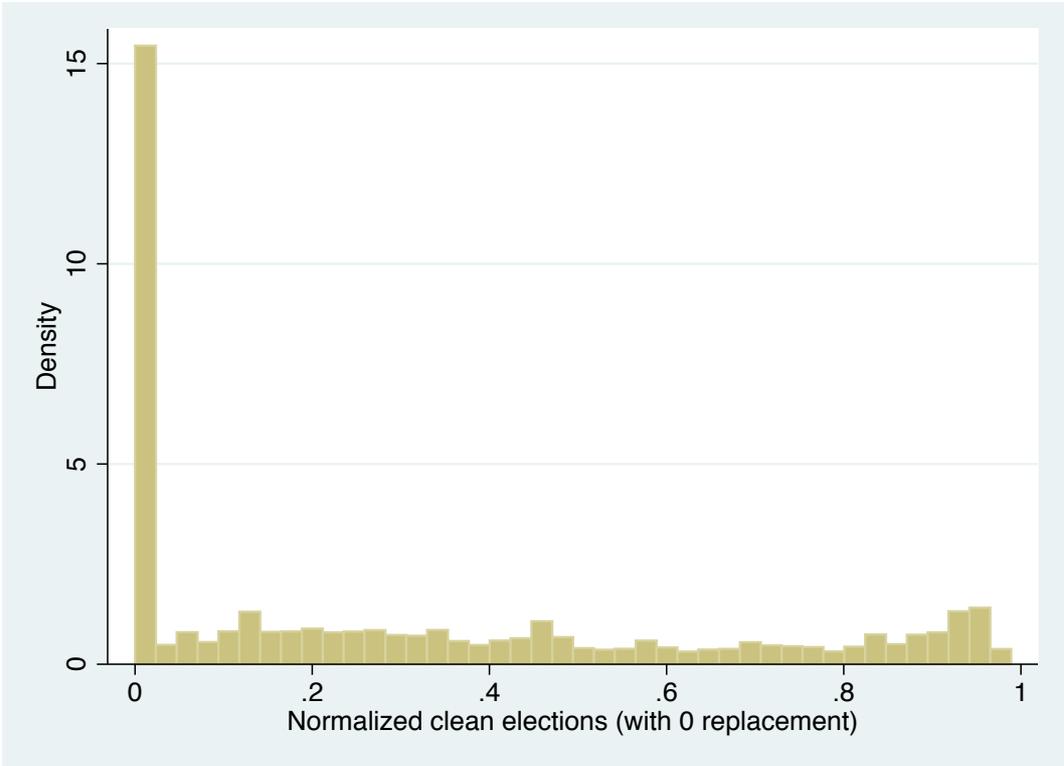


Figure C5. The Normalized Freedom of Organization Index ( $n=16,172$ )

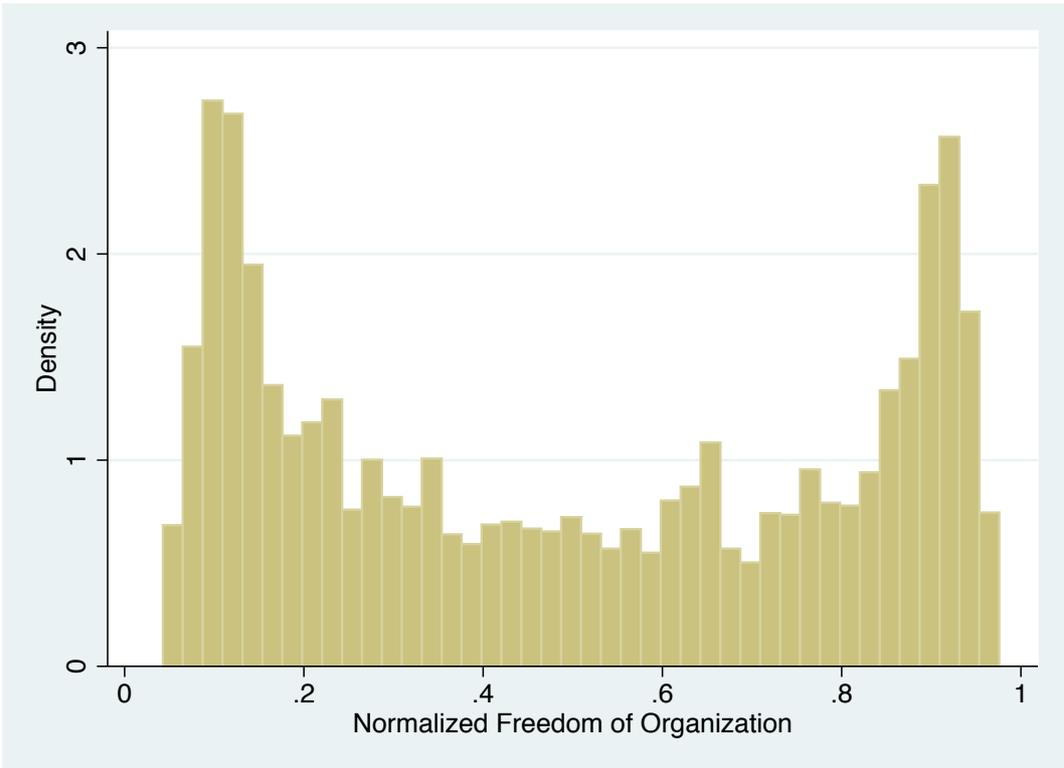


Figure C6. The Suffrage Indicator ( $n=16,474$ )

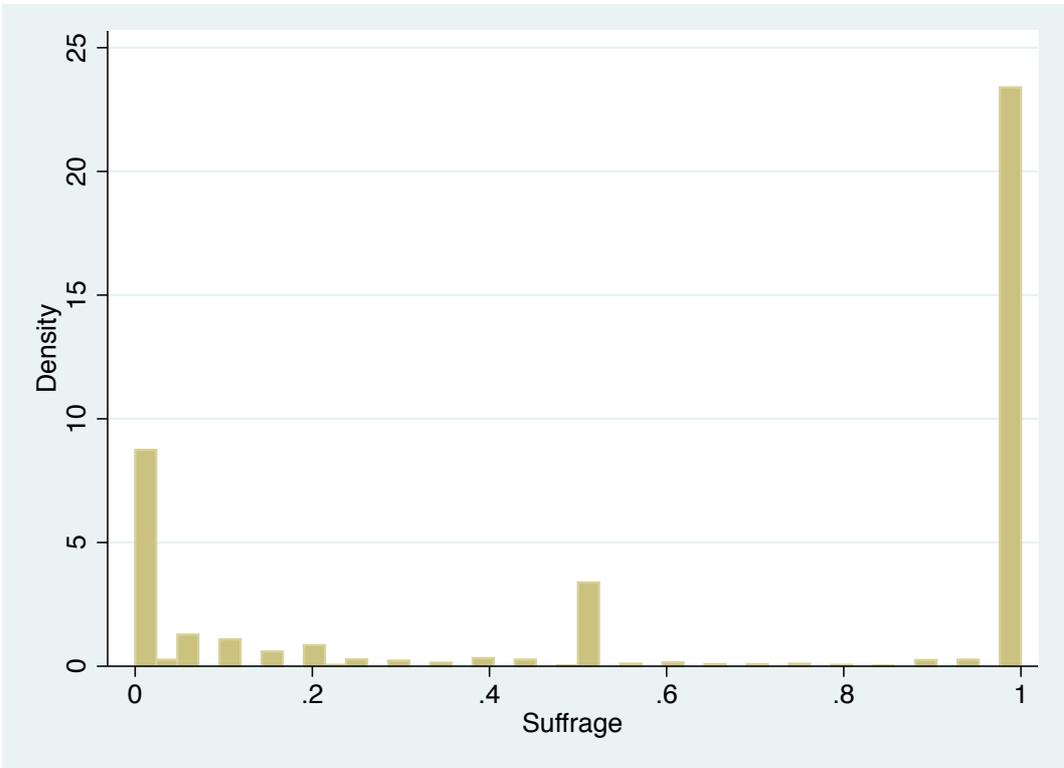


Figure C7. The Normalized Freedom of Expression Index ( $n=15,969$ )

