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

Jan Teorell
Department of Political Science, Lund University
jan.teorell@svet.lu.se

Michael Coppedge
Kellogg Institute for International Studies, University of Notre Dame
coppedge.1@nd.edu

John Gerring
Department of Political Science, Boston University
jgerring@bu.edu

Staffan Lindberg
Department of Political Science, University of Gothenburg
staffan.i.lindberg@pol.gu.se

Version November 20, 2014
Word count: 12,698

Paper to be presented at the Carlos III-Juan March Institute of Social Sciences, Madrid, November 28, 2014

WORK IN PROGRESS. PLEASE DO NOT CITE WITHOUT PERMISSION OF THE AUTHORS
Abstract: Extant measures of electoral democracy – such as FH, Polity, ACLP, Vanhanen, etc – suffer from several deficiencies. A first problem concerns the lack of conceptually meaningful disaggregated component or indicator scores of the broader construct, which means that the dimensionality of the phenomenon of electoral democracy is largely unknown. Related to this, the potential disjoint between minimalist and maximalist conceptions of electoral democracy cannot be explored with extant indicators, since they base this measurement strategy on conceptual fiat rather than empirical investigation. Second, rules of aggregation from indicators to component and/or index scores are rarely systematized or defended on conceptual grounds. Third, extant measures lack information on measurement uncertainty. The aim of this paper is to develop a new set of measures of electoral democracy on a global sample from 1900 to the present based on the V-Dem indicators and addressing these concerns.
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), an agreement, by and large, 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 of democracy – such as liberal, participatory, deliberative, egalitarian etc. (see Coppedge et al. 2011) – that do not stress elections as democracy’s core institutional pillar.

Given this conceptual point of departure, we shall propose a new way of measuring electoral democracy. This measure is based on the methodology of the Varieties of Democracy project (www.v-dem.net), which draws on multiple expert perceptions of a huge number of disaggregated indicators from a global sample of countries since 1900. More precisely, we will present country-year level component measures of the six core “institutional guarantees” in Dahl’s (1971, 1989, 1998) concept of “polyarchy,” as well as a proposal for how to aggregate these components into an aggregate measure of electoral democracy, including both point estimates and measurement uncertainty.

We shall argue this new measure has several advantages over the widely used extant measures of electoral democracy with similar coverage, such as Freedom House, Polity, Vanhanen (2000), Przeworski et al. (2000) and Boix et al. (2011). First, we can anchor the overall index of electoral democracy more directly within Dahls (1971) extremely influential theoretical framework. By measuring the six Dahlian (1998, 85) components of “Elected officials”, “Free and fair elections”, “Freedom of expression”, “Alternative sources of information”, “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. Once our data has been made public, our measure also allows for disaggregation, both to the components and to the indicators tapping into them. 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, all the data feeding into it are freely available on the project’s website, and it can thus be cross-examined or revised to fit the more specific needs of each user. Finally, we provide systematic measures of uncertainty (or measurement error).

The paper proceeds as follows. We start by stating our points of departure with respect to the conceptualization of polyarchy. We then outline the V-Dem methodology, followed by one section on how we measure the polyarchy components, and one on how we aggregate these components into the polyarchy index. We then provide some preliminary results to highlight the added value of our new measure.


As is well known, 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), as summarized in Table 1.¹

[Table 1 about here]

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

1 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 because eligibility and suffrage tend to go hand in hand (as argued by Coppedge & Reinicke 1990, 53), the second because it is more of a summary proxy for all the other institutional requirements taken together.
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. Second, and even more unambiguously, 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 fifth Dahlian institutional guarantees only.

What are the theoretical arguments in favor of restricting the domain of democracy to a more minimal set of components? 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. “By not bundling in additional elements of democratic practice, such as civil liberties,” argue Boix et al. (2013, 1527), “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.” We contend, however, that empirical consideration cannot solely 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; Cepeddge & Reinicke 1990; Hadenius 1992; Gasiorowski 1996; Mainwaring et al. 2001; Bowman et al. 2005). That said, an attractive feature of the disaggregated approach to measurement 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.

As against to Dahlian “maximalist” conception, on the other hand, the argument could be made that too many features of the political system are built into the concept. Most

2 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.
importantly, why should non-electoral components such as freedom of expression and alternative sources of information (institutional guarantees #3-#4 in Table 1) 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., Diamond 1999, 8-10; Möller & Skaaning 2010, 268-171; Munck 2009, 55-56). Without denying the need for such diminished subtypes for certain analytical purposes, we still contend 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 expression and the press, in order to ascertain that elections work as intended.3

On a final, more methodological note, it makes a great deal of difference for how to measure an electoral concept of democracy whether one adopts a causal or constitutive model of how indicators and components are related to higher-level concepts (Goertz 2006). In classical measurement theory, concepts are viewed as latent constructs that give rise to, or “cause,” their respective observable 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 more theory-driven or “ontological” approach to measurement and aggregation according to which concepts are defined by their components or indicators, the latter thus being 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 indicators “caused” by it.

3 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), but 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).
Both the causal and the constitutive view of democracy measures have been advocated.\(^4\) Bollen (1990, 12), for example, gives a very lucid presentation of the former view in writing: “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, mush like intelligence, is thus according to this view a latent trait of a country (Treier & Jackman 2008), one that lives a life of its own, as it were, and that causes observable indicators such as the degree of political liberties or rights to move as the underlying level of democracy moves.

To illustrate the opposite view, however, Goertz (2009, 15) writes that “to have competitive elections is not a symptom of democracy, it is not caused by democracy, but rather it constitutes what democracy is.”

As pointed out by Munck (2008, 30-31), the choice of viewing lower-level components or indicators as effects or constituent elements of higher-order constructs has important repercussions for aggregation. With the classical “effect indicators” approach, in which each indicator is considered a separate and at least partially substitutable measure “caused” by the higher-level construct, the natural aggregation rule is to take the average (or weighted average, with the weights made up of the strength of the “causal” relationship between each indicator and the higher-level construct). With the constituent approach, by contrast, it is up to the researcher to posit what kind of relation should hold between the lower-level indicators and the higher-level construct, and then adapt the choice of aggregation rule accordingly. If each indicator is viewed as perfectly substitutable for the others and hence a sufficient condition for the higher-level construct, for example, then taking the maximum is the warranted aggregation rule. But if, on the other hand, each indicator is a non-substitutable and hence necessary condition for the construct, then the aggregation rule should be either taking the minimum or drawing on multiplication (Goertz 2009, 39-44, 11; cf. Munck & Verkuilen 2002, 24; Munck 2009, 32).

\(^4\) 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 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.
In our measure of polyarchy, we will argue in favor of a middle-road approach, according to which some parts of the highest-level construct may be measured with effect indicators “caused” by its higher-level component, whereas other parts will on theoretical grounds be treated as being “constituted” by its lower-level elements. We will thus argue in favor of the approach that makes most theoretical sense in each case, not adopt any one-size-fits-all general solution.

3. The V-Dem Methodology

Since the large literature on previous 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 it here to say that in our view no extant measure fulfill all of the following four essential criteria: (1) covering all six components of Dahl’s polyarchy conception (as listed in Table 1), also 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; (2) on a global sample of countries and across long swaths of time; (3) with a transparent data generating process and aggregation rules; and (4) also providing estimates of measurement uncertainty.

To be sure, some extant measures live up to some of these criteria. Coppedge & Reinicke (1990), for example, provide measures of all (or almost all) of Dahl’s components, on a global sample and with 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 crisp coding and aggregation rules and their reliance on strictly observable data, also providing wide geographic and temporal coverage – but again with no measures of uncertainty, and only covering a minimalist subset of the concept of polyarchy. Treier & Jackman (2008), 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, unwarranted aggregation rule and non-transparent coding scheme.
There is thus a need for a new measure of polyarchy that can live up to all these stricter theoretical and methodological criteria simultaneously. We argue that the *Varieties of Democracy* (V-Dem for short) data can accomplish exactly this. Three features in particular underpins this claim (for a fuller description of the V-Dem methodology, see www.v-dem.net). The first feature is radical disaggregation: we translate the highest-level principles of democracy into more than 300 detailed questions with well-defined response categories or measurement scales. Although these questions and indicators have been designed to also tap into other conceptions of democracy than the electoral one, this still means we have assembled the polyarchy scale presented below on the basis of more than 40 individual variables. This not only means we are able to measure all six components of Dahl’s polyarchy concept, but also that they are (with one provisional exception) measured with multiple indicators to enhance reliability and provide tests for dimensionality.

Second, the bulk of the data stems from 177 indicators collected from country experts, mostly academics from each country in question, on a sample of countries covering the entire globe. These experts have been recruited based on their academic or other credentials as field experts in the area for which they code (the 177 questions are subdivided into 11 different areas of expertise, and most experts code a cluster of three such areas), on their seriousness of purpose and impartiality. At least 5 experts per country respond to each question and year going back to 1900. This means that more than 2000 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 our collection of expert ratings (see Pemstein et al. 2014). The underpinnings of these measurement models are straightforward: they use patterns of cross-rater (dis)agreement to estimate variations in reliability and systematic bias across disparate measures of the same, or similar, concepts (i.e. multiple expert ratings). In turn, these techniques make use of the bias and reliability estimates to adjust estimates of the latent – that is, only indirectly observed – concept in question. Although this allows us to provide both point estimates and standard

---

5 A notable feature of the V-Dem data is that we code a “country” throughout its history (since 1900) as a semi-sovereign unit. This implies that most colonies, and also some current non-independent territories such as Kosovo and the Palestines, are also included in the sample.
errors for each expert coded indicator, the uncertainty estimates of the higher-level components and indices presented below rely on the posteriors from a normal Bayesian factor analysis model. To make these aggregated models sensitive also to indicator-level uncertainty induced by expert disagreement is something we will be able to address in later stages of the project.

At present (i.e., November 2014), we have completed data collection on all variables pertinent to the electoral democracy index for 89 “countries” (or semi-sovereign territories). The list of countries covered is displayed in Appendix A. Since this sample is heavily biased toward the developing world, all results reported in this paper should be treated as highly preliminary.

4. Measuring the Parts: Six Components of Polyarchy

Our description of the V-Dem measure of polyarchy proceeds as follows. In this section, we present our measures of the six component indices corresponding to Dahl’s “institutional guarantees” presented in Table 1. In the next section, we present our aggregation rules and resulting higher-level indices.

A. Elected executive

This index attempts to measure 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 as the “effectiveness” (Hadenius 1992, 49) or “decisiveness” (O’Donnell 2001, 13) of elections. A “popular election” is 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.

---

6 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.

7 Another attractive feature of the V-Dem data is that experts are required to report a level of confidence for each coding, an indication of their subjective level of uncertainty. To what extent these map onto the uncertainty estimates produced by the measurement model is also subject to future research within the project.

8 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
The logic behind the construct is schematically portrayed in Figure 1. Unlike most of the following component indices, this index has not been designed as drawing on “effect indicators” thought of as reflections of an underlying, unobserved construct. Rather, 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.

[Figure 1 about here]

There are thus six different links of appointment/selection to take into account in constructing this index, all of which are scaled 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).\(^9\) 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 \(d=0\).

In polities 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 an 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

\(^9\) In terms of the V-Dem variables (see Appendix B): \(b=v2lgello\) (converted to fractions) in unicameral legislatures \((v2lbicam==1)\), and a weighted average of \(v2lgello\) and \(v2lgelecup\) (converted to fractions) in bicameral legislatures \((v2gbicam==2)\), where \(v2lgdchm/4\) defines the weight.
elected.\textsuperscript{10}

As Figure 2 reveals, the resulting index (called $v_{2x\_accex}$) is bimodally distributed, and as could be expected in this sample of developing countries skewed towards scores of 0 (having no elected executive). 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 which at present has a unified executive with a president elected by a parliament, only 75% of the seats of which are directly elected.

![Figure 2 about here]

Two potential limitations of the current version of this index should be noted. First, the extent to which non-elected “accountability groups” (such as the military) may affect executive dismissal and/or veto important domestic policy proposals has not been taken into account, the reason being that these variables have for technical reasons still not been processed through the measurement model. We plan to add information of “tutelary powers” of this kind in future iterations. Secondly, since this index concentrates on the election/appointment of the chief executive, presidential systems with reserved, unelected seats in the legislature is now not being penalized. For example, 8 out of 150 seats in the unicameral national assembly of Zambia are filled by presidential appointment and 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. Since this compromises the extent to which popular elections in these countries are “effective,” one could contemplate introducing a penalty to these countries.

\textbf{B. Clean elections}

The second component, which Dahl (1998, 85) calls “free, fair and frequent” elections, is designed to capture the absence of elections fraud, defined as the “introduction of bias into the administration of elections such that the voting process itself is distorted” (Schedler 2002, 10).

\textsuperscript{10} In technical terms, define $hosw$ as the weight for the head of state. If the head of state is also head of government ($v_{2xhoshog}=1$), $hosw=1$. If the head of state has more power than the head of government over the appointment and dismissal of cabinet ministers, then $hosw=1$; if the reverse is true, $hosw=0$. If they share equal power, $hosw=.5$. Define the weight for the head of government as $hogw=1-hosw$. The index is then assigned the value $hosw*[\max(a1,b*c1)]+hogw*[\max(a1*d,b*c1*d,a2,b*c2)]$. 

———
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 “clean.”

The V-dem questionnaire contains six questions pertaining exactly to this as judged from the country experts perspective on an election-by-election basis (see Appendix B on exact question wording and response categories). They tap into the extent of (a) registration fraud, (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 question whether the election, all things considered, could be considered “free and fair.” Finally, the questionnaire also contained two more general and not election-specific questions about the autonomy and capacity of the election administration body (EMB) to conduct well-run elections.

We view the aggregated country-level responses to these eight questions as partially substitutable effect indicators caused by the unobserved latent trait “clean elections”. To test this measurement proposition, we ran a Bayesian factor analysis model, the results of which are reported in Table 2. As can be seen, two indicators are particularly ill-behaved in terms of having weaker loadings and a larger share of their variance unaccounted for (uniqueness): vote buying and election violence not conducted by the government. This is thus an indication that the construct of free and fair elections could be multidimensional, a notion we will want to explore in future iterations of this paper. For present purposes, however, the fit to a unidimensional model was deemed good enough.

[Table 2 about here]

---

11 If legislative (lower house or constituent assembly) and presidential elections were held concurrently, the questions pertained 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).

12 The questionnaire also includes other election indicators, such as the incidence of election boycotts and freedom of campaign media. These are excluded in part because they tap into the regulations of election campaigns rather than the electoral process itself, and in part due to issues of multidimensionality.

13 We were also able to replicate this model for an additional set of 73 countries for a fuller sample of 2,574 election years. The results are fairly similar, but overall the fit to a unidimensional model improves substantively.
The index of clean elections (called v2x_frefair) is based on the point estimates from this Bayesian factor analysis model, distributed as in Figure 3. Since this index is only observed for election years, however, we face a problem with how to extrapolate its values over time. Although we could in the future 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, 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).

A thornier issue concerns what value to assign to the index during periods of electoral interruptions (such as after coups or violent takeovers of the government etc). Recall that, being based on the scores from the item response theory measurement model, there is no natural zero point of this index, nor is there a theoretically defined minimum score. The interim solution we have therefore implemented is to convert the index score to a probability (0-1) score by using the normal cumulative distribution function (cdf), and then replace all observations with 0 during electoral interruptions. 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 4 shows how this transformed 0-1 index, with 0 replacement for electoral interruptions, map onto the original BFA estimates. As can be seen, apart from the few elections that (within the same year) were aborted by an electoral interruption (these are the observations clustered along the horizontal line at 0 on the transformed scale), 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 Expression

As noted above, Dahl’s concept of polyarchy also includes some non-electoral aspects, most conspicuously having to do with the freedom of expression. If we want to be true to his extremely influential conceptualization of “democracy,” we will thus need to borrow some indicators from the “liberal” sphere and incorporate them into our measure of “electoral” democracy. Arguably, going for a core of indicators on freedom of expression should include indicators of media freedom (Behmer 2009), both censorship of print/broadcast media and (in present times) of the internet, as well as other more informal forms of restrictions on media freedom such as harassment of journalists, media self-censorship and repression of cultural and academic expressions of political dissent. In addition, there should be freedom of discussion in society at large (Skaaning 2009). We again view the aggregated country-level responses to seven (reduced to six) such questions as partially substitutable effect indicators caused by the unobserved latent trait “freedom of expression.” As indicated by Table 3, this measurement proposition clearly fits the data. With the partial exception of the internet censorship indicators, the strong loading on most indicators indicate a unidimensional underlying construct giving rise to this data.

[Table 3 about here]

Hence, our index of freedom of expression (called v2x_freexp) 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 5.

---

14 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.

15 Note that a complication introduced by this indicator is how to deal with the 0s – having no Internet access, which arguably should not be seen as having less freedom of discussion than code 1, for example, where there is Internet access but this is being blocked by the government (see Appendix B for exact question wording). A proposed solution to this would be to recode all 0s into missing values, and then impute these missing values with the information from the other indicators. This is however not how the current version of the measurement model was implemented.
D. Alternative Sources of Information

This component index, again following Dahl (1971), is based on three other media indicators that are more geared towards measuring media content rather than government censorship: whether the media is biased against opposition parties and candidates, whether major print and broadcast outlets routinely criticize the government, and whether they represent a wide range of political perspectives. The effect indicators model here again seems most appropriate for our purposes, with the results in Table 4 giving strong support for a unidimensional construct. The construction of the index (called $v2x_{altinf}$) follows the same procedures as above, including the 0-1 transformation (see distribution in Figure 6). Not surprisingly, this index is strongly correlated with the freedom of expression index ($r=.78$), and a combined model of the 10 indicators in Tables 3-4 together show strong support for a unidimensional construct underlying both sets of indicators. The reason we have opted to keep them separate is thus purely theoretical: to allow for a replication of all of Dahl’s (1971, 1998) six components of polyarchy.

E. Associational Autonomy

The indices of elected officials and clean elections, as noted above, does not take into account the degree of pluralism in a country. In other words, they could, at least in theory, achieve their maximum 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. The core of this construct, as a component of polyarchy, is party-centered and hence electoral: are political parties free to form, operate independently of the ruling regime, and run in national elections? Yet for associational autonomy in its true
meaning to be fulfilled, there also has to be no barriers of entry and exit or repression of a wider set of civil society organizations, the purpose of which could be completely non-political.\textsuperscript{16} We have thus opted for this thicker set of effect indicators of freedom of organization, as indicated in Table 5, all of which clearly load on a single underlying dimension.\textsuperscript{17} The resulting index (called $v2x\_freorg$), again rescaled to 0-1 through the normal cdf, is – very similarly to both freedom of expression and alternative sources of information – slightly positively skewed for this sample of countries (see Figure 7).

\textit{F. Inclusive Citizenship}

For the purpose of measuring electoral democracy, Dahl’s (1971, 1989, 1998) sixth and final component, “inclusive citizenship,” amounts to the extension of the suffrage. After exploring several not fully satisfying alternative solutions based on ordinal codings of the seriousness of suffrage restrictions (as in Przeworski et al. 2013), or imputed estimates of the number of registered voters as a share of the total population, we have for now opted for an estimate of the proportion of the electorate eligible to vote roughly based on the Paxton \textit{et al.} (2003) methodology (see Appendix B, section F, for further details).\textsuperscript{18} Since this variable has for practical purposes only been collected for election years, its values have been repeated between election years but set to zero during electoral interruptions (as per above). The resulting 0-1 index (called $v2x\_suffr$) is bimodally distributed with a slight tilt towards zero in this particular sample of countries (see Figure 8). In later developments of this paper, we plan

\textsuperscript{16} In this particular respect, however, the dividing line between more minimalist ”Schumpeterian” and a more maximalist ”Dahlian” concept of freedom of organization, where the former would most likely argue in favor of ignoring these additional civil society features, cuts within one of Dahl’s six 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 .98 with the thicker version we opted for.

\textsuperscript{17} Since the multiparty elections indicator is only observed in election years, we face the same extrapolation problem as for $frefair$ above The interim solution (again) implemented is to convert the measurement model score to a probability (0-1) by using the normal cumulative distribution function (cdf), and then replace all observations with 0 during electoral interruptions

\textsuperscript{18} The variable has been hand-coded by Svend-Erik Skaaning, Aarhus University. Since this is a de jure estimate, we would eventually also want to incorporate some information on de facto election-day restrictions to turn out to vote, for which we are still collecting data.
to replace this single-indicator estimate with a latent variable estimate of suffrage based on multiple indicators and sources,

[Figure 8 about here]

5. Measuring the Whole: Aggregating the Components

Remains then the question of how these six separately measured components of polyarchy should best be aggregated together into an index. Before delving in to that question, however, it might prove useful to present some information on how they covary (in this particular sample of country-years). This information is presented in Table 6. Here, in the lower diagonal, the correlation coefficients pertain to the full set of country-year observations. As can be seen, the covariation is moderate to strong in most instances. In particular two clusters of stronger correlations stand out: one between having an elected executive and the extension of the suffrage, the other between freedom of expression, alternative sources of information and freedom of organization. The clean elections index is moderately correlated with both these clusters. One could however argue that a substantial part of these correlations is induced by the fact that periods of electoral interruptions have been replaced by trailing 0s for the clean elections and the suffrage indices (as well as for one indicator of the freedom of organization index). We therefore in the upper diagonal present the correlations for election-years only. Here the picture changes somewhat. Most importantly, almost every single correlation with these two indices becomes significantly weaker. Yet the two clusters of relatively stronger correlations, between an elected executive and suffrage on the one hand, and among the freedom indices on the other, remain.

[Table 6 about here]

Since this could be interpreted as an indication – but far from conclusive evidence to the effect\(^\text{19}\) – that polyarchy is multidimensional, as argued by Coppedge & Reinicke (1990) and

\(^{19}\) 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.
Coppedge, Alvarez & Maldonado (2008), it raises the bar for what 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 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.

Nevertheless, as argued above, theoretical considerations should take precedence over empirical in the choice of aggregation rules. The fist and most important question is thus whether we consider the six components of polyarchy as being caused by an underlying construct or whether they are to be seen as constitutive of that construct. We will separate our response to this question into two parts, the first based on the four electoral components (A, B, E & F), the other also adding the non-electoral ones (C & D).

Starting with the four electoral components – elected officials, clean elections, freedom of organization, and suffrage – we find a strong rationale in the literature for treating these as constitutive elements. More precisely, as argued by Przeworski et al. (2000) with the exception of the suffrage component, and even more clearly so by Munck (2009) and Boix et al. (2013), these four elements could be considered as non-compensatory and necessary conditions for polyarchy. In other words, the degree of suffrage should not be counted if there is no freedom of organization, if the election results are completely fabricated, or if the highest executive office is not elected. Similarly, the freedom and fairness of elections should not count if only a small minority of the population is enfranchised, and so on and so forth. This theoretical intuition would thus clearly rule out taking the maximum as the aggregation rule, since that is tantamount to treating each component as a substitutable and sufficient condition for defining the higher-level construct (Goertz 2006). It also weighs against the use of simple (or weighted) averaging as the aggregation rule, since averaging means that high score 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 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 not generally the case. Although we agree that the minimum, for the perspective of fuzzy set/Boolean algebra, is the most strict interpretation of what a necessary
condition implies, we believe it has an important drawback: it implies that the value of a single indicator completely determines the aggregated index value. Multiplication has the relative virtue of combining information from all constitutive elements of the overarching concept and hence relies on more information. When all index components are scaled from 0-1, as is here the case, it is also fairly easily interpretable, with each component indicator working as a weight variable, as it were, for the others.

While working on developing a more sophisticated aggregation rule, that could offset the tendency of multiplication to completely ignore variation in the other components whenever one component is scores 0, we have provisionally opted to define this electoral component index of polyarchy (still not counting the non-electoral aspects) as:

\[
1. \text{Electoral component index} = \text{Elected Officials} \times \text{Clean Elections} \times \text{Freedom of Organization} \times \text{Suffrage}
\]

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 Index are multiplied with the corresponding standard errors for the Freedom of Organization Index.

Note that, by the logic of multiplication, a low score on any of the component indices suppresses the value of the overall index. As expected from a sample of countries biased toward the developing and more authoritarian parts of the worlds, the distribution of this index is thus heavily skewed towards zero (see Figure 9).

[Figure 9 about here]

The second part of the aggregation procedure adds the two non-electoral components of freedom of expression and alternative sources of information. We start by averaging these two component indices, since we believe there is a partial substitutability effect between allowing extensive freedom of expression and alternative sources of information. Hence, as a first step:

\[
2. \text{Non-electoral component index} = (\text{Freedom of Expression} + \text{Alternative Sources of Information})/2
\]
The measurement uncertainty is again taken into consideration by the same rule, implying that the standard errors from the BFA estimates of Freedom of Expression are averaged with the standard errors of Alternative Sources of Information.

But how then should the electoral and non-electoral component indices be aggregated. Two main alternatives seem most viable: (1) to again draw on strict multiplication, which would imply that freedom of expression and alternative sources of information only counts to the extent that the other, more purely electoral components (suffrage, freedom of association, elections etc) are fulfilled; or, as an alternative, (2) we could opt for a more compensatory logic allowing countries scoring high on freedom of expression and/or alternative sources of information but low on the electoral component, or vice versa, to fare somewhere in the middle.

It is first of all important to observe the reason why multiplication and averaging in this case produces so different distributional results: by comparing figures 5-6 with figure 9, it is readily apparent that countries more frequently score high on freedom of expression and alternative sources of information than on the multiplicative electoral component index. This is also clearly visible in the scatter plot presented in Figure 10, wherein the index value of the non-electoral component index is almost without exception above the index value of electoral one (in other words, the lower-right diagonal is almost completely empty). The largest difference between a multiplicative and an average-based polyarchy scale will thus arise for countries scoring lowest on the electoral-component, since that is where the difference between to two component indices is the largest. In essence, the question of aggregation rule thus comes down to whether we want to compensate these, in an electoral sense, low-performing countries by their relative higher quality in terms of non-electoral standards?

[Figure 10 about here]

Although future work within our project will hopefully provide a stronger empirical basis for this choice, we will tentatively respond to this question in the positive, and hence opt for the average-based aggregation rule (the measurement uncertainty thus again being propagated into the index by averaging across the two component standard errors), with the slight modification that the original siz Dahlian component should weight equally (hence the electoral component should be weighted twice as heavily as the non-electoral):
(3) Polyarchy=(2 * Electoral component + Non-electoral component)/3

This index is distributed as in Figure 11. Compare this to an alternative, purely multiplicative index:

(3') Multiplicative Polyarchy=Electoral component * Non-electoral component

This index is distributed almost exactly as the electoral component index in Figure 9 above, and the two version map onto each other as in Figure 12. As should be expected from the comparison between the electoral and non-electoral components, the largest discrepancies between the two ($r=.94$) occur at the lower ends of multiplicative index, where countries are allowed to assume modestly positive scores on the additive, by us, preferred version of the polyarchy scale.

[Figure 11 about here]

[Figure 12 about here]

However, we do not arrive at this proposed solution out of purely distributional concerns. There actually is a strand in the literature on how to measure electoral democracy going back to Bollen (1980), Coppedge & Reinicke (1990) and Hadenius (1992), which is based on the logic of averaging electoral components with components based on “freedoms” of different sorts. What would seem to make this practice justifiable is that it fits nicely with O’Donnell & Schmitter’s (1986) concept of “liberalization” – the “phase” in a country’s transition to democracy in which the first opening of the authoritarian regime has occurred, such as a lifting of media censure and wider acceptance for expressions of popular discontent, but before the “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 averaging logic, however, allows such openings to be counted in and of themselves.
6. The Added Value: Some Empirical Illustrations

We start our exploration of this new polyarchy index by showcasing its 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 13, the over-time development in level of electoral democracy over the 20th and early 21st century is well covered by our measure. With the exception of electoral democracy’s rise and authoritarian backlash in Europe and the Americas in the interwar period – a development we should not be expecting to find any traces of in this sample of developing countries ²⁰ – we track both the second and third of Huntington’s (1991) three waves of democracy. As compared to Polity, the V-Dem polyarchy measure tends to come out as a more conservative measure of electoral democracy, a feature that historically could most likely be explained by our explicit inclusion of the suffrage component (as is the case of Boix, Miller and Rosato 2012), but the reasons for which in the last 20 years of development need yet to be fully explored. 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 – only in the last year of measurement, in 2012. The more exact reasons for this also deserves further inquiry.

[Figure 13 about here]

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 14, we display the ranking of countries in our sample, from lowest to highest degree of polyarchy, together with 90% confidence intervals. ²¹ In this sample of countries, Turkmenistan, Laos, Eritrea and Uzbekistan score at the very bottom, whereas Taiwan, South Korea and Brazil are ranked highest. As makes intuitive sense, the margins of measurement error tend to be largest in the middle. It should be noted however that these confidence intervals do not (yet) take cross-coder disagreement within countries and years into account. Instead, the measurement uncertainty reflected in these margins of error reflect cross-indicator variation within the four

²⁰ More precisely, we do not track the first wave in this sample due to the fact that we treat semi-sovereign territories such as colonies on par with independent states. If we restrict our attention to the 11 countries in our sample independent prior to WWII, also the V-Dem polyarchy index shows sign of a rise and decline in electoral democracy in the interwar period.

²¹ These confidence intervals have been arrived at by taking +/- 1.645 times the standard deviation of the posterior distributions. Instead looking at the 6th to 95th percentile arrives at almost exactly the same estimates.
components measured as latent traits (clean elections, freedom of expression, alternative sources of information, and freedom of organization). By implication, they tend to underestimate the true amount of measurement uncertainty.

[Figure 14 about here]

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 within 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 however 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 2011; 2014). That is, if we unbundle the different components of electoral democracy, will other and different results with respect to the overall modernization effect evolve?

In Figure 15, we show the results of a series of dynamic panel regressions with this purpose in mind. To address potential endogeneity issues, each model includes two lags of the dependent variable, as well as the lag of ln(GDP/capita) and yearly GDP/capital growth (data from Madison). We also control for both country- and year-fixed effects to address concerns of omitted variable bias.23

[Figure 15 about here]

Concentrating on the short-term effects of ln(GDP/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 the electoral components of the polyarchy index, whereas the non-electoral component (freedom of expression and alternative sources of information) are not significantly linked to economic development. In other words, at least

22 Based on the lagrange multiplier test recommended by Beck & Katz (1996), two lags of the dependent variable are necessary to clean the error terms from serial autocorrelation.
23 With an average time series components of 55 years per country, Nickel bias should not be of great concern (Beck & Katz 2011).
within this restricted sample of countries, 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.

Drilling down one level further, we also find variation among the four components of the electoral component index. Modernization is most strongly related to the extension of the suffrage, followed by attaining clean elections. The one electoral component that is most weakly connected to level of economic development is freedom of organization.

References


Figure 1. The Logical Structure of the Elected Officials Index
Figure 2. The Elected Officials Index ($n=8,150$ country years)
Figure 3. The Clean Elections Index ($n=1,313$ election years)
Figure 4. Transforming the Clean Elections Index
Figure 5. The Transformed Freedom of Expression Index ($n=7,131$)
Figure 6. The Transformed Alternative Sources of Information Index (n=7,111)
Figure 7. The Transformed Freedom of Organization Index \( (n=7,646) \)
Figure 8. The Suffrage Indicator \((n=8,275)\)
Figure 9. The Electoral Component Index \((n=7,562)\)
Figure 10. The Electoral and Non-Electoral Component Indices (n=6,974)
Figure 11. The V-Dem Polyarchy Index ($n=6,974$)
Figure 12. The Multiplicative and Additive Polyarchy Indices (n=6,974)
Figure 13. Aggregate Trends in the Major Democracy Indices

Note: Results for the 89 completed V-Dem countries only. Freedom House, Polity and Vahanen have been rescaled to 0-1. BMR (Boix, Miller & Rosato 2012), and ACLP (Przeworski et al. 2000; Cheibub & Ghandi 2010) are binary measures.
Figure 14. Polyarchy in 2010
Figure 15. Economic development, polyarchy and its components

V-Dem polyarchy index

Electoral component

Non-electoral component

Elected officials

Clean elections

Freedom of organization

Suffrage

Note: Entries are regression coefficients for the effects of lagged ln(GDP/capita), controlling for year- and country-fixed effects, two lagged dependent variables, and lagged growth in GDP/capita.
Table 1. The Six Institutional Guarantees of Polarchy

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A.</td>
<td>Elected officials</td>
</tr>
<tr>
<td>B.</td>
<td>Free, fair and frequent Elections</td>
</tr>
<tr>
<td>C.</td>
<td>Freedom of Expression</td>
</tr>
<tr>
<td>D.</td>
<td>Alternative Sources of Information</td>
</tr>
<tr>
<td>E.</td>
<td>Associational Autonomy</td>
</tr>
<tr>
<td>F.</td>
<td>Inclusive Citizenship</td>
</tr>
</tbody>
</table>

Table 2. Measuring Clean Elections (BFA estimates)

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Loadings ($\Lambda$)</th>
<th>Uniqueness ($\Psi$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proper voter registry (v2elrgstry)</td>
<td>.487</td>
<td>.767</td>
</tr>
<tr>
<td>Vote buying (v2elvotbuy)</td>
<td>.421</td>
<td>.826</td>
</tr>
<tr>
<td>Ballot fraud/vote irregularity (v2elirreg)</td>
<td>.784</td>
<td>.390</td>
</tr>
<tr>
<td>Government intimidation (v2elintim)</td>
<td>.836</td>
<td>.307</td>
</tr>
<tr>
<td>Other electoral violence (v2elpeace)</td>
<td>.323</td>
<td>.898</td>
</tr>
<tr>
<td>Overall judgment “free &amp; fair” (v2elfrfair)</td>
<td>.799</td>
<td>.368</td>
</tr>
<tr>
<td>EMB autonomy (v2elembaut)</td>
<td>.652</td>
<td>.580</td>
</tr>
<tr>
<td>EMB capacity (v2elembcap)</td>
<td>.525</td>
<td>.729</td>
</tr>
</tbody>
</table>

Note: Entries are factor loadings and uniqueness from a normal theory Bayesian factor analysis model; $n=941$ election years.
Table 3. Measuring Freedom of Expression (BFA estimates)

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Loadings (Λ)</th>
<th>Uniqueness (Ψ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Print/broadcast censorship (v2mecenefm)</td>
<td>.797</td>
<td>.366</td>
</tr>
<tr>
<td>Internet censorship (v2mecenefi)</td>
<td>.435</td>
<td>.812</td>
</tr>
<tr>
<td>Harassment of journalists (v2meharjrm)</td>
<td>.670</td>
<td>.551</td>
</tr>
<tr>
<td>Media self-censorship (v2meslfccen)</td>
<td>.780</td>
<td>.392</td>
</tr>
<tr>
<td>Freedom of academic/cultural expr. (v2clacfree)</td>
<td>.787</td>
<td>.382</td>
</tr>
<tr>
<td>Freedom of discussion (v2cldiscm+v2cldiscw)</td>
<td>.878</td>
<td>.231</td>
</tr>
</tbody>
</table>

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=5,656 country-years.
Table 4. Measuring Alternative Sources of Information (BFA estimates)

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Loadings ($\Lambda$)</th>
<th>Uniqueness ($\Psi$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Media bias (v2mebias)</td>
<td>.826</td>
<td>.318</td>
</tr>
<tr>
<td>Print/broadcast media critical (v2mecrit)</td>
<td>.870</td>
<td>.243</td>
</tr>
<tr>
<td>Print/broadcast media perspectives (v2merange)</td>
<td>.907</td>
<td>.177</td>
</tr>
</tbody>
</table>

*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; $n=5,718$ country-years.
Table 5. Measuring Freedom of Organization (BFA estimates)

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Loadings (Λ)</th>
<th>Uniqueness (Ψ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Party ban (v2psparban)</td>
<td>.853</td>
<td>.273</td>
</tr>
<tr>
<td>Barriers to parties (v2psbars)</td>
<td>.885</td>
<td>.219</td>
</tr>
<tr>
<td>Opposition parties autonomy (v2psoppaut)</td>
<td>.845</td>
<td>.287</td>
</tr>
<tr>
<td>Elections multiparty (v2elmulpar)</td>
<td>.770</td>
<td>.407</td>
</tr>
<tr>
<td>CSO entry and exit (v2cseeorgs)</td>
<td>.836</td>
<td>.303</td>
</tr>
<tr>
<td>CSO repression (v2csreprss)</td>
<td>–.757</td>
<td>.428</td>
</tr>
</tbody>
</table>

Note: Entries are factor loadings and uniqueness from a normal theory Bayesian factor analysis model; n=5,860 country-years.
Table 6. Correlations among Polyarchy Components ($r$)

<table>
<thead>
<tr>
<th></th>
<th>Elected Officials</th>
<th>Clean Elections</th>
<th>Freedom of Expression</th>
<th>Alternative Sources of Information</th>
<th>Freedom of Organization</th>
<th>Suffrage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elected Officials</td>
<td>.009</td>
<td>.232</td>
<td>.266</td>
<td>.217</td>
<td>.324</td>
<td></td>
</tr>
<tr>
<td>Clean Elections</td>
<td>.489</td>
<td>.578</td>
<td>.457</td>
<td>.589</td>
<td>.032</td>
<td></td>
</tr>
<tr>
<td>Freedom of Expression</td>
<td>.419</td>
<td>.583</td>
<td>.811</td>
<td>.838</td>
<td>.223</td>
<td></td>
</tr>
<tr>
<td>Alternative Information</td>
<td>.469</td>
<td>.528</td>
<td>.776</td>
<td>.810</td>
<td>.187</td>
<td></td>
</tr>
<tr>
<td>Freedom of Organization</td>
<td>.519</td>
<td>.689</td>
<td>.804</td>
<td>.819</td>
<td>.156</td>
<td></td>
</tr>
<tr>
<td>Suffrage</td>
<td>.735</td>
<td>.587</td>
<td>.430</td>
<td>.446</td>
<td>.526</td>
<td></td>
</tr>
</tbody>
</table>

*Note: Entries are correlation coefficients, in the lower diagonal for $n=6,668$ country-years, in the upper diagonal for $n=1,213$ election years.*
Appendix A: The sample of fully completed V-Dem countries and coding periods

Afghanistan, 1900-2012  
Algeria, 1900-2012  
Angola, 1900-2012  
Argentina, 1900-2012  
Armenia, 1990-2012  
Azerbaijan, 1990-2012  
Bangladesh, 1971-2012  
Belarus, 1990-2012  
Benin, 1900-2012  
Bhutan, 1900-2012  
Bolivia, 1900-2012  
Brazil, 1900-2012  
Burkina Faso, 1919-1932, 1947-2012  
Burma/Myanmar, 1900-2012  
Cambodia, 1900-2012  
Cameroon, 1961-2012  
Chad, 1920-2012  
China, 1900-2012  
Colombia, 1900-2012  
Congo, Democratic Republic, 1900-2012  
Congo, Republic, 1903-2012  
Djibouti, 1900-2012  
Dominican Republic, 1900-2012  
El Salvador, 1900-2012  
Eritrea, 1900-2012  
Ethiopia, 1900-2012  
Gabon, 1910-2012  
Gambia, 1900-2012  
Georgia, 1990-2012  
Ghana, 1902-2012  
Guinea, 1900-2012  
Guinea-Bissau, 1900-2012  
Haiti, 1900-2012  
Honduras, 1900-2012  
India, 1900-2012  
Indonesia, 1900-2012  
Jamaica, 1900-2012  
Jordan, 1922-2012  
Kazakhstan, 1990-2012  
Kenya, 1900-2012  
Kosovo, 1999-2012  
Kyrgyzstan, 1990-2012  
Laos, 1900-2012  
Lebanon, 1918-2012  
Libya, 1934-1942, 1951-2012  
Madagascar, 1900-2012  
Mali, 1900-2012  
Mauritania, 1904-2012  
Moldova, 1990-2012  
Morocco, 1900-2012  
Mozambique, 1900-2012  
Namibia, 1900-2012  
Nepal, 1900-2012  
Nicaragua, 1900-2012  
Niger, 1922-2012  
Nigeria, 1914-2012  
North Korea, 1900-2012  
Pakistan, 1947-2012  
Peru, 1900-2012  
Philippines, 1900-2012  
Rwanda, 1916-2012  
Senegal, 1904-2012  
Sierra Leone, 1900-2012  
Somalia, 1900-2012  
Somaliland, 1900-1960, 1991-2012  
South Sudan, 2011-2012  
Sri Lanka, 1900-2012  
Sudan, 1900-2012  
Swaziland, 1900-2012  
Taiwan, 1900-2012  
Tajikistan, 1990-2012  
Tanzania, 1900-2012  
Thailand, 1900-2012  
Togo, 1916-2012  
Trinidad and Tobago, 1900-2012  
Tunisia, 1900-2012  
Turkmenistan, 1990-2012  
Uganda, 1900-2012  
Uzbekistan, 1990-2012  
Venezuela, 1900-2012  
Vietnam, Democratic Republic, 1945-2012  
Vietnam, Republic, 1902-1975  
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 gathered from extant sources (other datasets or secondary sources). The collection of this data is supervised by the Principal Investigators and Project Managers and carried out by Research Assistants connected to the project. Selected A-data are also cross-checked and vetted by V-Dem’s Country Coordinators. Type (B) data is gathered from country-specific sources by Country Coordinators under the supervision of Regional Managers. 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 (B/A): How did the head of state/government reach office? (HOS: 0=coup/rebellion, appointed by 1=foreign power, 2=ruled
ning 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=ruled
ning 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 (B/A): Was approval of the legislature necessary for the appointment of the head of state/government? (0=no, 1=yes)

v2exremhsp/hog – HOS/HOG removal by legislature in practice (C): If the legislature, or either chamber of the legislature, took actions to remove the head of state/government from office, would it be likely to succeed even without having to level accusations of unlawful activity and without the involvement of any other agency? (0=no, under no circumstances, 1=no, unlikely, 2=yes, probably, 3=yes, most likely)

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

v2exdfcbhs/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, 2=yes, but only with the tacit consent or explicit confirmation by the legislature, 3=yes, without any need for confirmation by the legislature, but only with respect to the head of the cabinet, 4=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/v2exdfdshg – 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)

v2lgdomchm – 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)

v2lgello/v2lgelecup – Lower/Upper chamber elected (B/A): What percentage of the lower/upper chamber of the legislature is directly elected in popular elections? (%)
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)
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 expression index:

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.)

Internet censorship effort (C): Does the government attempt to censor information (text, audio, or visuals) on the Internet? (0=country has no Internet access, 1=successfully blocks Internet access except to pro-government or non-political sites, 2=attempts to block access, but many users able to circumvent, 3=allows access, but blocks some sites, 4=government allows unrestricted Internet access)

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)

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)

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)

D. The alternative sources of information index:

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)

E. 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)

**F. 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.