



Rethinking Consensus vs. Majoritarian Democracy

Michael Coppedge

October 2018

Working Paper

SERIES 2018:78

THE VARIETIES OF DEMOCRACY INSTITUTE



UNIVERSITY OF GOTHENBURG
DEPT OF POLITICAL SCIENCE

Varieties of Democracy (V-Dem) is a new approach to conceptualization and measurement of democracy. The headquarters – the V-Dem Institute – is based at the University of Gothenburg with 17 staff. The project includes a worldwide team with six Principal Investigators, 14 Project Managers, 30 Regional Managers, 170 Country Coordinators, Research Assistants, and 3,000 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

V-Dem Working Papers are available in electronic format at www.v-dem.net.

Copyright © 2018 by authors. All rights reserved.

Rethinking Consensus vs. Majoritarian Democracy*

Michael Coppedge

Professor

University of Notre Dame

* This research project was supported by Riksbankens Jubileumsfond, Grant M13-0559:1, PI: Staffan I. Lindberg, V-Dem Institute, University of Gothenburg, Sweden; 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. 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, SNIC 2017/1-407 and 2017/1-68. 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

Arend Lijphart's distinction between two dimensions of consensus and majoritarian democracy has been an influential concept. However, several scholars have reported that the distinction does not travel well to other regions or historical periods. This paper argues, more generally, that Lijphart's dimensions can be replicated only when using Lijphart's own data. If one substitutes conceptually similar indicators (in this case, mostly from V-Dem), three or four dimensions emerge, and they are not robust to different samples. Such substitutions would be necessary for anyone wishing to measure Lijphart's dimensions beyond his chosen cases. It is hard to avoid the conclusion that these concepts are not generally useful. However, it is possible to construct a couple of simple indices measuring thinner related concepts.

Introduction

Arend Lijphart's *Patterns of Democracy* (1999) has been a very influential book in comparative politics.¹ Its central claim is that stable democracies vary along two dimensions: an executive-parties dimension and a federal-unitary dimension, each of which has a majoritarian pole and a consensus pole.² This claim rests on a principal components analysis of ten indicators that shows that five of the indicators measure one dimension and five of them measure the other dimension. Lijphart's analysis has elicited quite a few critiques, some of which Lijphart himself has graciously accepted. Nevertheless, the basic distinction between consensus and majoritarian democracy is still widely used. It is a resilient idea.

This paper joins the critical chorus. It confirms some of the findings of past critiques, but analyzes the replication failure in a more comprehensive and systematic way in order to explain *why* others have not been able to find the same two dimensions of democracy that Lijphart discovered. The problem, I argue, is not so much that the world has changed, or that Lijphart's concepts apply only to certain countries. Rather, the main issue is that his analysis is sensitive to the selection and operationalization of indicators. As long as we use Lijphart's indicators, or even subsets of his indicators, we can reproduce his two dimensions. The more we substitute indicators generated by anyone else (which is necessary for replication beyond his small sample), even ones that were designed to measure very similar versions of his concepts, the less possible it becomes to replicate his analysis. This may be the result of the many adjustments Lijphart made to the data he used: much of *Patterns of Democracy* consists of documentation and defenses of these decisions. In this paper, I use mostly Varieties of Democracy (V-Dem) data, which provide a best-case opportunity for replication; yet the attempt fails. Replication is a fundamental principle of science. If Lijphart's dimensions cannot be replicated using good data generated by anyone other than Arend Lijphart, they have limited utility. My conclusion is that the problem is conceptual as well as empirical: it is not useful to speak of two consensus vs. majoritarian dimensions of democracy.³

¹ Analyses in this paper address the 1999 edition. A pending version will address the 2012 second edition.

² *Patterns of Democracy* is also known for the claims in the final chapter that consensus democracy is a "kinder, gentler" form of democracy. However, these claims were empirically weak and have never been considered as persuasive as the existence of the two dimensions of democracy.

³ This paper does not address the hypothetical question of how to combine a consensus-majoritarian dimension (if there were one) with electoral democracy to create indices of consensus democracy and majoritarian democracy. It is an interesting question because consensualism and majoritarianism lie at opposite poles of one dimension, yet both presuppose electoral democracy. However, there is a solution. Suppose we have a Consensus vs. Majoritarian Component Index (ConsMaj). Like V-Dem's other component indices, it does not directly take degrees of polyarchy into account. If the index is on a 0-1 scale with 1=maximum consensualism, majoritarianism can be measured as $1 - \text{ConsMaj}$. This poses no obstacle to using the same formula V-Dem uses for all the other democracy indices:
$$v2x_consensusdem = .25*v2x_polyarchy + .25*ConsMaj + .5*(v2x_polyarchy*ConsMaj)$$
$$v2x_majoritariandem = .25*v2x_polyarchy + .25*(1-ConsMaj) + .5*(v2x_polyarchy*(1-ConsMaj))$$

The distinction is too general to meaningfully capture a more multidimensional reality. We would be better off defining and measuring more specific dimensions and attributes of democracy.

I. The Reception of Patterns of Democracy

Patterns of Democracy has (as of August 2017) about 28,000 citations in Google Scholar and 1,657 in the Web of Science. These statistics understate the far-reaching influence of the concept because Lijphart made a similar (albeit one-dimensional) argument in his earlier book *Democracies* (1984) and in many articles and chapters before and after 1999. Even so, the statistics for this one book compare favorably with Web of Science citations of, for example, Guillermo O'Donnell's *Modernization and Bureaucratic-Authoritarianism* (1973 and 1979) (563 cites) and Adam Przeworski et al.'s *Democracy and Development* (2000) (1,221 cites), a bit fewer than Theda Skocpol's *States and Social Revolutions* (1979) (1,809 cites), although only a fraction of the more than 7,000 citations for Robert Putnam et al.'s *Making Democracy Work* (1993), which probably marks the upper limit in comparative politics. Lijphart's 1999 book is indisputably well-known. Mainwaring called Lijphart's distinction between consensus and majoritarian democracy "the single most influential typology of modern democracies" (2001, 171).

Some other prominent books have mined similar institutional territory. G. Bingham Powell's *Elections as Instruments of Democracy* (2000) made a similar distinction between majoritarian and "proportional" democracies. George Tsebelis's *Veto Players* (2002) and many related articles focuses attention on the same kinds of institutions, such as executives, legislatures, political parties, courts, and federalism. Gerring and Thacker's *A Centripetal Theory of Democratic Governance* (2008) also dealt with the type of electoral system, the type of executive, and federal vs. unitary government. To be sure, these other scholars do not necessarily agree with Lijphart. Tsebelis argues that the kinds of institutions that Lijphart considers consensual are prone to policy paralysis, and Gerring and Thacker see advantages in the rather majoritarian, power-concentrating "centripetal" government. Lijphart had a long-running debate with Donald Horowitz (2001) about the relative advantages of institutions that encourage diversity or unity. But all of these authors are interested in how democratic institutions concentrate or disperse power in ways that correspond closely with Lijphart's distinction between consensus and majoritarian democracy. This is the distinction that I question in this paper.

The result of these formulas is that while Consensus and Majoritarianism have a perfect -1 correlation, indices of consensus and majoritarian democracy are not very correlated. Both increase as polyarchy increases. However, at any fixed level of polyarchy, there is a tradeoff between consensualism and majoritarianism.

There have been more than a dozen published critiques of Lijphart's two dimensions of democracy, making for extended lively debate.⁴ Some critics have conceptual or theoretical objections. Bogaards (2000), for example, argues that over the course of his career, Lijphart shifted from an empirically based concept of consociationalism to a more normative concept of consensus democracy and later to a more neutral notion of "power-sharing democracy"; and that the inconsistencies between how these three types classify cases make for a shaky foundation for testing Lijphart's prescriptive claims. Ganghof (2010 and 2012) argues that Lijphart misconceived the oversized coalitions indicator and presents evidence that when properly conceived, it does not belong on either of the two dimensions.

Another line of criticism questions Lijphart's choices of indicators for these concepts. Taagepera (2003) provided a stern critique of Lijphart's operational decisions. Taagepera discerned little theoretical rationale to expect the federal-unitary indicators to measure the same dimension, especially central bank independence. He also questioned the logic behind connections among the executives-parties indicators (except for disproportionality and the effective number of parties), especially interest-group pluralism: "[I]t feels odd that two-party systems would go with a profusion of interest groups, while multi-party systems require a two-group interest pattern" (7). He also believed that Lijphart made serious mistakes in measuring cabinet life for executive dominance. Mainwaring's 2001 review of the book (alongside Powell's 2000 book) raised questions about the selection of indicators, the rationale for the dimensions, and several of Lijphart's coding decisions.

Still other critics have reported empirical relationships that are inconsistent the claim of two consensus-majority dimensions. Fortin (2008) reports that the two dimensions could not be replicated using similar ratings for ten Central European democracies, and concludes that the pattern is sensitive to the sample of cases. Vatter (2009), after updating Lijphart's data for 23 OECD countries for 1997-2006 and adding a new direct democracy measure, reports that there are now three dimensions of democracy; but central bank independence moves into the executives-parties dimension and oversized cabinets moves into the third dimension with direct democracy. Vatter and Bernauer (2009) use 25 OECD countries for the same period and find four dimensions, with several of the indicators moving around among them. Croissant and Schächter (2009) disconfirm the two-dimensional pattern for Asia countries from the 1980s through 2005.

A final target of criticism has been Lijphart's claim that consensus democracy brings macroeconomic benefits. Anderson (2001) shows that any macroeconomic benefits of consensus democracy were due solely to two of its components – corporatism and central bank independence

⁴ For an excellent summary of many of the critiques, including some not mentioned here, see Bormann (2010).

(which were the only two components added in 1999 compared to the 1984 book). Similarly, Giuliani (2016) argues that corporatism is a separate concept from consensualism, and empirically demonstrates that all of the alleged benefits of consensus democracy for macroeconomic performance are due solely to corporatism.

Lijphart has acknowledged the thrust of some of these critiques. He wrote in 2003 that he regretted averaging two measures of cabinet life, and wished he had logged rather than truncated its extreme values. But defended most of his other choices. Although he agreed that interest-group pluralism and central bank independence did not belong conceptually with their dimensions as clearly as the other indicators, he defended these choices as reasonable and empirically justified. He speculated that the presidential-parliamentary distinction, which was not in his model, would belong on a third dimension – which suggested to him that it was not worth including.

What is interesting in this debate is that, although scholars question whether certain components belong in the model, whether others should be added, which dimensions they should be expected to measure, whether the model applies to Central Europe or Asia or Africa or marginally democratic countries, hardly anyone suggests abandoning the basic distinction between majoritarian democracy and its alternative, whether that alternative is labeled consensus, consociational, proportional, power-sharing, inclusive, or something else. Something about the distinction seems to be powerfully intuitive and compelling, and perhaps normatively appealing (as it certainly is to Lijphart). The distinction *is* compelling at the level of individual indicators. We all know that there are real differences between two-party and multiparty systems; bare-majority governments and oversized cabinets; federal and unitary systems; constitutions that are hard or easy to amend; pluralist and corporatist systems; and so on. These are meaningful distinctions. However, they are not the question. The question is whether these distinctions are all aligned in such a way that countries at the majoritarian pole of the executives-parties dimension on one indicator tend to be at the same pole on the other four indicators of that dimension; and the same for the opposite pole; and the same for the federal-unitary dimension that Lijphart defined. The literature indicates that these alignments are fragile for some reason, and the analyses to be presented below validate that conclusion. Why? Was the phenomenon that Lijphart discovered limited to the 1945-1996 period? Was it limited to the 36 stable democracies he chose? Was it an artifact of the indicators he selected? Did it depend on the way he generated each indicator? This paper makes a case for the crucial importance of the indicators rather than case selection or period effects.

II. Replication with Lijphart's Data

Lijphart proposed that the distinction between consensus and majoritarian democracy could be measured with ten indicators: the effective number of parties (seats), minimal-winning/one-party coalitions, executive dominance, disproportionality, interest-group pluralism (vs. corporatism), federalism, bicameralism, constitutional rigidity, judicial review, and central bank independence. He found that the first five indicators loaded most strongly on one dimension, which he interpreted as an “executives-parties” or horizontal dimension, while the last five indicators loaded most strongly on a distinct “federal-unitary” or vertical dimension. Table 1 reproduces the Lijphart’s principal components analysis using the data from the 1999 book, which consists of one rating for each of 36 countries on ten indicators. There is only one rating for each country-indicator because Lijphart chose what he considered the most typical value for each country – where possible, a value that remained constant for the country’s whole experience (1945-1996 or whenever the country came into existence or first passed a threshold of sufficient democracy). But if there were changes in the value, he sometimes chose the value that the country had for most of its relevant experience, or in a few cases, the most recent value it attained, or an average of values over the whole timespan.

Table 1: Reproduction of Lijphart's Dimensions

Variable	Factor1	Factor2	Uniqueness
Eigenvalue	3.47	2.85	
Cumulative variance explained	0.347	0.632	
Effective number of parties	-0.896	0.029	0.197
Minimal-winning/one-party coalitions	0.927	-0.084	0.133
Executive dominance	0.740	-0.113	0.440
Disproportionality	0.722	0.077	0.473
Interest-group pluralism	0.781	-0.015	0.390
Federalism	-0.272	0.861	0.185
Bicameralism	0.067	0.736	0.454
Constitutional rigidity	-0.038	0.711	0.493
Judicial review	0.209	0.723	0.434
Central bank independence	-0.065	0.715	0.485
N=36, varimax rotation, KMO=.670.			
Cronbach's alpha for separate factors	0.349	0.740	

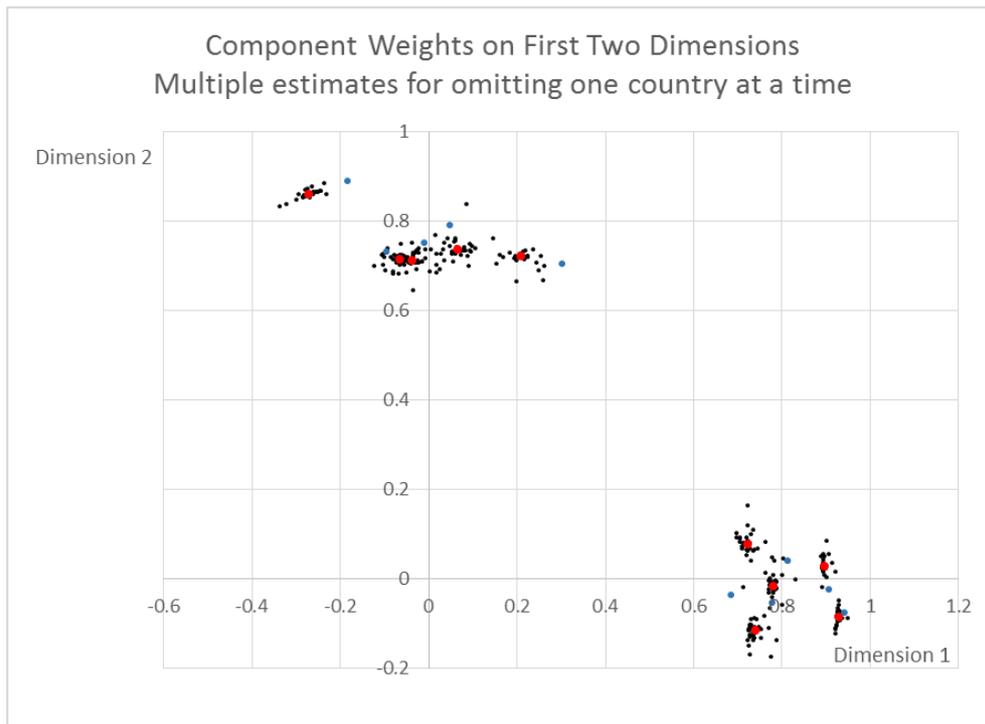
Table 1 confirms a clear two-dimensional pattern: each indicator loads strongly on one component and only weakly on the other, and the indicators are associated with the components

that match the interpretation Lijphart gave them.⁵ The table also includes some statistics not reported in the book: the Eigenvalues, the proportion of the variance explained by each component, each indicator's uniqueness, and Cronbach's alpha, a rating of internal consistency that is estimated separately for the set of indicators associated with each component. There are two and only two Eigenvalues greater than 1, which is consistent with a two-dimensional model. The executives-parties component explains about a third of the variance and the two components together explain about 63 percent. The only indicators with low uniquenesses, signifying a strong relationship with the components, are effective number of parties, oversized cabinets, and federalism; the rest are noisier measures than one would like for these dimensions. The KMO-Bartlett test of sampling adequacy (.670) is only marginally adequate; it would be better to have a larger selection of relevant indicators for these two dimensions. The alpha values are the most disappointing: the value of .740 for the second dimension is conventionally considered "questionable," and the .349 for the first dimension is simply "unacceptable." There is a pattern here, but it is a weak pattern with a lot more measurement error than is typically considered acceptable in social science research.

It is important to acknowledge at the outset that Lijphart's patterns of democracy, weak and noisy though they may be, are robust as long as we use his data. They are not sensitive to dropping a small number of countries from the sample. If we drop one country at a time, the Eigenvalues still range from 3.02 to 3.69 for the first dimension and 2.49 to 2.90 for the second, with the cumulative proportion of variance explained being .335 to .369 and .618 to .649, respectively. Figure 1 plots the component weights for each indicator on each dimension (following varimax rotation, which matters quite a bit here), with one black point for each estimate based on dropping a different country. It is easy to see that the estimates for these different samples are tightly clustered in a small area for each indicator, and that five of the indicators load primarily on dimension 1 and five primarily on dimension 2. The red points are the estimates obtained when dropping three countries simultaneously – Bahamas, Malta, and Papua New Guinea – that are not present in the V-Dem data used below. Because some of the replications below can use only 29 of the 36 cases, this figure also shows blue points for estimates run without seven countries: the three already mentioned plus Barbados, Botswana, Luxembourg, and Mauritius. The pattern is the same. There is thus little reason to expect that replications with subsets of these countries for the same years will produce a pattern other than Lijphart's two dimensions.

⁵ Principal components analysis is necessarily exploratory. A confirmatory factor analysis validates the same pattern fairly well: the factor loadings range from .527 to .972 with a coefficient of determination of .997, CFI=.980, RMSEA=0.051, SRMR=0.100. However, the unique variances of the indicators range from .054 to .722 (judicial review). In short, the commonality is explained well by the two factors, but much of the variance in the indicators is not shared.

Figure 1



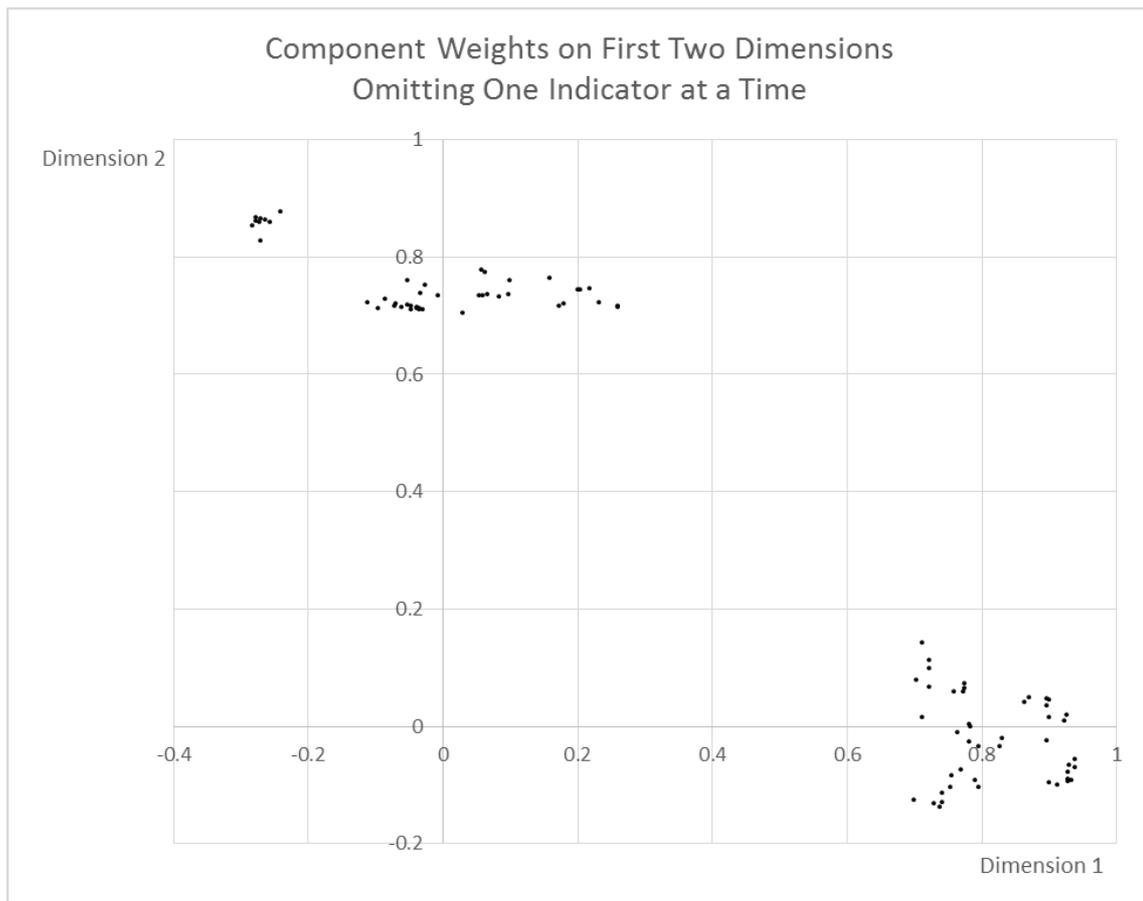
Dimension 1 weight for ENPS is reversed to be positive.

Red points indicate estimates that omit Bahamas, Malta, and Papua New Guinea.

Blue points indicate estimates that omit the above three and Barbados, Botswana, Luxembourg, and Mauritius.

Neither is the two-dimensional pattern sensitive to dropping one or two indicators from the model. If we drop one indicator at a time, the Eigenvalues still range from 2.85 to 3.47 for the first dimension and 2.18 to 2.85 for the second, with the cumulative proportion of variance explained being .317 to .386 and .611 to .657, respectively: always two and only two dimensions, with the same indicators loading on their assigned dimensions. Figure 2 plots the component weights for these estimates. They are clustered a little less tightly than in Figure 1, but it is still easy to appreciate that the pattern is approximately the same regardless of which indicator is dropped. The same pattern holds when dropping any two of Lijphart's indicators at a time: There are always just two dimensions and the same indicators load on each (not shown). We therefore need not fear that a replication will fail if one or two of the indicators is missing. So far, then, the fragility of the two consensus-majoritarian dimensions does not seem to be due to the selection of different subsets of Lijphart's countries or indicators. The culprit is more likely to be the substitution of completely different indicators.

Figure 2



Dimension 1 weight for ENPS is reversed to be positive. Components are switched for two of the models.

III. Alternative Indicators

It is also possible that different results would turn up if we attempted to replicate the analysis for countries or historical periods that were not in Lijphart's sample. However, we cannot do that with his original data, so any out-of-sample replication requires using alternative indicators.⁶ Obviously, it is important for substitute indicators to measure Lijphart's concepts as closely as possible. V-Dem data are close to ideal for this purpose for three reasons (Coppedge et al. 2017a, 2017b). First, V-Dem utilizes a highly disaggregated conceptual scheme that has led to the creation of indicators of hundreds of specific attributes of democracy. This makes it more likely that measures that capture Lijphart's concepts either already exist in the data or can be constructed from several indicators; and that these measures would not be contaminated by other concepts, which is likely to happen when using measures of more general concepts. Second, the V-Dem project took

⁶ A pending version of the paper will attempt replication with a subset of Lijphart's original years, as reported in appendix B of the 1999 book; with Lijphart's updated data from the 2012 second edition; and with updated data on 25 OECD countries from Vatter and Bernauer (2009).

Lijphart's thinking into account when designing its conceptual scheme. In fact, one of its original goals was to produce indices of consensus and majoritarian democracy, which it has not yet done. These were two of the seven "varieties" of democracy V-Dem set out to measure, and the project's concept paper (Coppedge et al. 2011) cites him as the primary inspiration for these conceptions of democracy. Third, although some V-Dem concepts were written with Lijphart's concepts in mind, those who wrote the concepts did not collect the data (except to a trivial degree).⁷ Nearly 3,000 country experts from all over the world submitted the great bulk of the expert ratings. These coders mostly rated one country each over a long span of time, which discouraged them from considering whether a rating for one period might combine with ratings of other countries to form some kind of pattern. They could not coordinate their ratings with experts in other countries, who were anonymous. Furthermore, the measurement model that aggregates many coders' ratings into point estimates analyzes each indicator in isolation, not a group of indicators that might belong to the same dimension (Pemstein et al. 2016). These procedures help ensure that each indicator is a valid measure of each particular concept, not of some possible underlying dimension. In this respect, V-Dem data may in fact be preferable to data collected by a researcher who may unconsciously have had a predetermined outcome in mind while collecting it.

Six of the ten alternative indicators come from multiple V-Dem experts in each country; one (constitutional rigidity) was coded by one V-Dem coordinator in each country; and one (disproportionality) by V-Dem research assistants in Sweden and Chile. The other two (effective number of parties and central bank independence) come from non-V-Dem sources (Bormann and Golder 2016; Garriga 2016).

Lijphart's indicator of judicial review was designed to capture whether courts can invalidate legislation, which depends on whether there is a legal provision for judicial review; and if so, how activist the courts are. My alternative is V-Dem's v2jureview, Judicial review: "Does any court in the judiciary have the legal authority to invalidate governmental policies (e.g. statutes, regulations, decrees, administrative actions) on the grounds that they violate a constitutional provision?" Responses are binary: 0 for No, 1 for Yes. This indicator captures only the legal authority, as V-Dem has no indicator of judicial activism.

For constitutional rigidity, Lijphart operationalized the difficulty of amending the constitution by taking into account primarily the size of the majority needed to amend it. However, he lessened the rigidity if legislators were elected in plurality or majority systems, because that made it easier to win a supermajority; or if there was a non-legislative alternative procedure that was easier. I

⁷ Some leaders of the V-Dem team did some country coding, but their ratings constitute a miniscule fraction of the millions of ratings the project has collected.

constructed an alternative indicator (vconstrig) using two centrally coded indicators. One is v2lgamend, Legislature amends constitution: “By law, can the legislature (including both chambers of the legislature) change the constitution without the involvement of any other body?” To capture one alternative procedure, I also used v2ddlexor, Constitutional changes popular vote:

Is a popular and direct vote required in order for a constitutional change to be legally binding?

0: No, it is not required.

1: Depends on the content of constitutional change (for some it is required, for others however it is not).

2: Yes, any constitutional must be approved directly by the citizenry.

Because the question wording treats the popular vote as a more difficult extra step that increases rigidity, I calculated the index as $v2lgamend + v2ddlexor / 2$. This doubles the difficulty if there is a second step to amendment. This alternative is conceptually somewhat different in that it lacks information about the size of the special legislative majority and ignores the type of electoral system (which is included in a different indicator). However, it does take into account what I consider the two most important elements: whether the legislature has authority to amend, and whether any non-legislative process is permitted.

The concept that Lijphart intended to measure for disproportionality was the distinction between majority or plurality vs. PR electoral systems. However, he operationalized it with Gallagher’s index of seat-vote gaps: a measure of the degree of disproportionality realized in practice rather than of the institutions that one might expect to produce disproportionality. He received some criticism on this point from, among others, Rein Taagepera (2003). For the alternative (vdisprop), I hewed close to the institutional definition by taking the mean of two centrally coded indicators – v2elparlel, Lower chamber electoral system (centrally coded by V-Dem from various sources):

What was the electoral system used in this election for the lower or unicameral chamber of the legislature?

0: Majoritarian.

1: Proportional. [recoded 2]

2: Mixed. [recoded 1]

3: Other (e.g. single non-transferable voting, limited voting)

and v2elloelsy, Lower chamber electoral system/13 categories, which is International IDEA’s finer-grained indicator of the same concept. If one of these indicators was missing values, I simply used the other. I had to recode and transform these before averaging them, as indicated in the brackets. I recoded the IDEA indicator into four levels: 0 for first past the post, a two-round system, alternative vote, and block vote in multi-member districts; 1 for parallel (SMD/PR) and mixed-member proportional (SMD with PR compensatory seats); 2 for list PR; 3 for STV, SNTV,

Limited vote in multi-member districts, and the Borda Count. These two classifications, when recoded, were in close agreement.

Lijphart's measure of interest-group pluralism was designed to capture "a competitive and uncoordinated pluralism of independent groups in contrast with the coordinated and compromise-oriented system of corporatism that is typical of the consensus model" (Lijphart 1999, 171). For this purpose he used Siaroff's index for 21 countries (Siaroff 1998) combined with "impressionistic" ratings of the twelve developing countries based on advice from country experts.⁸ As the alternative (*vigplur*), I used two V-Dem indicators, *v2csstruc*, "CSO structure," and *v2cscnsult*, "CSO consultation." CSO structure is defined as a series categories, and is measured continuously as a country-year's probability of belonging to that category rather than the others:

Civil societies inevitably involve a mix of larger and smaller organizations. Please characterize the relative influence of large mass constituency civil society organizations (CSOs) versus smaller, more local, or narrowly construed CSOs.

0: The state does not allow autonomous CSOs. (0=No, 1=Yes) [v2csstruc_0]

1: Large encompassing organizations dominate. The government and CSOs are linked formally through a corporatist system of interest intermediation; or, due to historical circumstances, particular large CSOs are highly influential. The voice of such organizations is recognized by the government and is accorded special weight by policymakers. (0=No, 1=Yes) [v2csstruc_1]

2: Neither large encompassing nor small CSOs dominate. Influence is contingent on circumstances. Organizations, both large and small, contend with one another to have their voice considered by policymakers. (0=No, 1=Yes) [v2csstruc_2]

3: Small CSOs dominate. Many small organizations contend with one another to have their voices heard by policymakers. (0=No, 1=Yes) [v2csstruc_3]

Most of the variation in the index I constructed comes from summing the probabilities for categories 2 and 3, which both reflect pluralism, and then subtracting the probability for the most corporatist category (1). However, corporatism requires institutionalized consultation of large encompassing organizations, not merely their existence. CSO consultation supplies this information:

Are major civil society organizations (CSOs) routinely consulted by policymakers on policies relevant to their members?

⁸ I should disclose that I advised Lijphart on several of the indicators he used for Venezuela.

0: No. There is a high degree of insulation of the government from CSO input. The government may sometimes enlist or mobilize CSOs after policies are adopted to sell them to the public at large. But it does not often consult with them in formulating policies.

1: To some degree. CSOs are but one set of voices that policymakers sometimes take into account.

2: Yes. Important CSOs are recognized as stakeholders in important policy areas and given voice on such issues. This can be accomplished through formal corporatist arrangements or through less formal arrangements.

This indicator forms a monotonically increasing 0-2 scale of institutionalized consultation. I divided it by 4 to shrink the range to 0-0.5 and subtracted it from the modified scale of CSO structure to finalize the alternative interest-group pluralism index: $v2csstruc_2 + v2csstruc_3 - v2csstruc_1 - v2cscnsult / 4$. The resulting index ranges (in Lijphart's sample of country-years) from maximum pluralism of 0.91 in Colombia (1992-1996) to maximum corporatism of -1.79 in Sweden (1948-1989).

Lijphart chose to measure executive dominance with cabinet duration, which he believed to be evidence of a powerful executive vis-à-vis the legislature. This has been the most questioned operational decision in part because there is an ongoing debate about how to measure cabinet stability and in part because of the way he adapted this measure for presidential systems (Lijphart 2003, Mainwaring 2001, Taagepera 2003, Tsebelis 2002). Rather than join this debate about cabinet stability, I used V-Dem measures that were designed to directly capture the concept of legislative constraints on the executive. The codebook definition of this index, $v2xlg_legcon$, is

To what extent are the legislature and government agencies (e.g., comptroller general, general prosecutor, or ombudsman) capable of questioning, investigating, and exercising oversight over the executive?

The index is formed by taking the point estimates from a Bayesian factor analysis model of the indicators for legislature questions officials in practice ($v2lgstexp$), executive oversight ($v2lgotovst$), legislature investigates in practice ($v2lginvstp$), and legislature opposition parties ($v2lgoppart$).

Although the definition mentions other bodies such as the ombudsman, this index primarily reflects power relations between the legislature and the executive, which is what is needed.

Lijphart measured cabinet size as the average of the percentage of cabinets that are minimal-winning and the percentage that are one-party, both of which counted as majoritarian arrangements. Comparable measures do not exist for the much more extensive sample of countries and years used in this paper, so I settled for an index that captures the distinction between single-party government and divided-party control of government, with coalition government serving as an intermediate value. The V-Dem codebook defines this index, $v2x_divparctrl$, as

Are the executive and legislature controlled by different political parties?

This variable is a reordered version of the continuous measurement model estimates for indicator v2psnatpar: National party control. After reordering, the positive extreme signifies Divided party control. (A) Different parties or individuals (unconnected to parties) control the executive and the legislature or (B) Executive power is divided between a president/monarch and a prime minister, each of which belongs to different parties; or between a non-partisan monarch and a prime minister. The intermediate values signify Unified coalition control. A single multi-party coalition controls the executive and legislative branches of the national government. (This is true almost by definition in a parliamentary system where a single coalition gathers together a majority of seats.) And the negative extreme signifies “Unified party control. A single party controls the executive and legislative branches of the national government. (This is true almost by definition in a parliamentary system where a single party has a majority of seats.)”

(Note that the poles are flipped: positive values for more consensual divided control, negative values for majoritarian unified control.)⁹ Conceptually, this index measures a significantly different concept than what Lijphart had in mind. Empirically, however, we will see that it is one of the alternatives that is most equivalent to Lijphart’s measures ($r=-.674$).

The indicators of federalism are admittedly distinct in concept, but also acceptably correlated in practice. Lijphart’s concept was a bit unorthodox, as he considered not only strictly federal vs. unitary institutions, but also other forms of decentralization. To capture these, he introduced an intermediate category of semi-federalism that lumped together regional autonomy (as in Spain), "sociological" rather than territorial federalism (as in Belgium, Israel, and the Netherlands); and Papua New Guinea. My alternative lacks the semi-federal category altogether and even departs from conventional notions of federalism by considering only the power of elected regional and local bodies relative to unelected bodies at those levels. The V-Dem codebook defines this “Division of Power” index, v2x_feduni, as

Are there elected local and regional governments, and – if so – to what extent can they operate without interference from unelected bodies at the local level?

This index is an equally weighted average of a local government index and a regional government index. The local government index is the product of a dummy variable for the existence of local government (v2ellocgov), a recoded version of Local government elected (v2ellocelc), and a CDF of Local offices relative power (v2ellocpwr). Local governments are recoded as unelected (0) if they did not exist or if data is missing. They are coded 0.5 if an executive is elected but no assembly, and 1 if an assembly is elected, with or without an executive. The

⁹ Since this index was constructed from some variables that have values only in election years, I carried those values forward to subsequent years as long as the electoral regime was preserved.

regional government index is calculated the same way but using the existence of regional government (v2elreggov), Regional government elected (v2elsrgel), and Regional offices relative power (v2elrgpwr).

Both indicators of federalism are therefore unorthodox; yet they have a correlation of .554 in Lijphart's sample.

Probably the least equivalent indicators are those for bicameralism. Lijphart constructed an indicator to measure the degree to which legislative power is divided between distinctly representative bodies with equal power. To do this, he took into account the number of chambers, their symmetry, and their congruence. I will review his operational decisions, which I consider questionable, below. My alternative, *vbicamerl*, takes into account the number of chambers, the percentage of members who are elected, and which chamber is dominant. For bicameral legislatures, the value is the sum of the percentage of the lower house that is elected and the percentage of the upper house that is elected, with each chamber weighted by their dominance: $v2lgello*(4 - v2lgdomchm)/4 + v2lgelecup*(v2lgdomchm)/4$. For unicameral legislatures, it is simply the percentage of members who are elected. For cases without a legislature, it is zero. This index would seem to capture most of Lijphart's concept – the number of chambers, their symmetry, and the “percent elected” aspect of their congruence – but not the “electoral law” or malapportionment that he also considers as attributes of congruence. Nevertheless, empirically these two indicators are far from equivalent in Lijphart's sample ($r=0.195$).

The last two alternative indicators come from non-V-Dem sources. The effective number of parties, in terms of lower-chamber seats, is a fairly standard index (Laakso and Taagepera 1979). I used extended and update data from Bormann and Golder (2016), compiled in the V-Dem dataset as *e_ellonmpl* and carried forward for non-election years. These two indicators are the most equivalent pair in the replication ($r=.923$). For central bank independence, defined as a combination of a bank with charter authority to have exclusive control over monetary policy and secure tenure of the central bank governor, Lijphart combined three different indices produced by economists. I use a new index, *lvau_garriga* (also carried forward for non-election years) (Garriga 2016), which covers 182 countries from 1970 to 2012 and is most similar to the Cukierman, Webb and Neyapti index that served as Lijphart's main source. It seeks to measure “the central bank's capability of controlling monetary instruments . . . or, inversely, CBI is the set of restrictions to the government's influence on the central bank management of monetary policy” (Garriga 2016, 850). It takes into account personnel, financial independence, and policy independence.

Table 2 shows the correlations between each of Lijphart's indicators and my nearest equivalents for 32-24 of the 36 countries between 1945 and 1996. The average correlation is 0.557, which may be acceptable, but four of the ten are less than 0.5, and they range from 0.923 to 0.195.

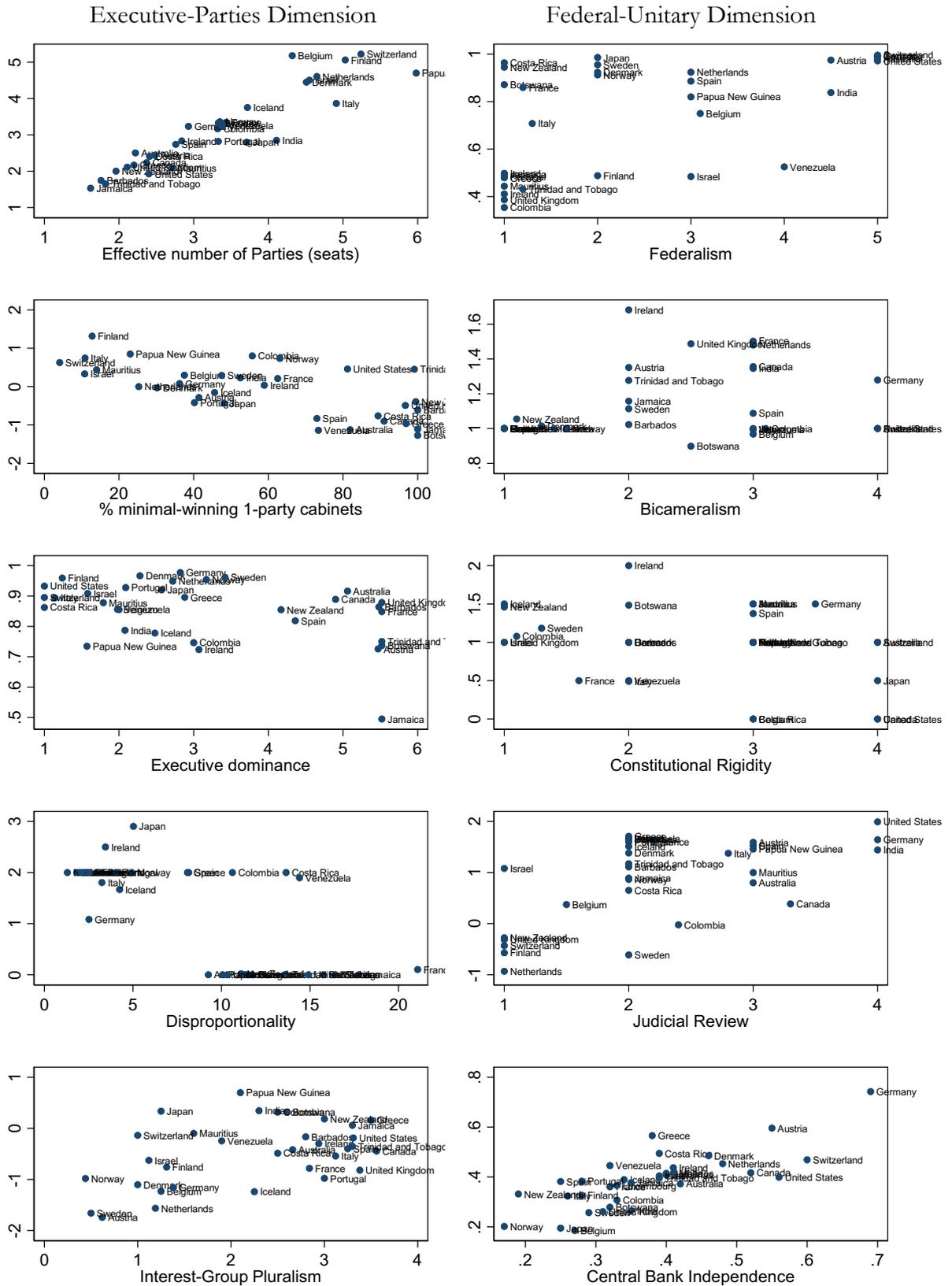
Given the unevenness of the correlations, it would not be surprising if they yield different results in a principal components analysis.

Table 2: Correlations between Lijphart's indicators and their nearest equivalents

<i>Executive-parties dimension</i>	r	n
Effective number of parties	0.923	33
Minimal-winning/one-party coalitions	-0.674	33
Executive dominance	-0.400	33
Disproportionality	-0.716	33
Interest-group pluralism	0.467	33
<i>Federal-unitary dimension</i>		
Federalism	0.554	32
Bicameralism	0.195	33
Constitutional rigidity	-0.312	33
Judicial review	0.563	33
Central bank independence	0.766	34

Figure 3 compares each pair of indicators in a series of scatterplots and labels countries for ease of interpretation. On the executives-parties dimension, three of the relationships look reasonable and strike me as having face validity: effective number of parties, minimal-winning or one-party cabinets, and disproportionality. On the federal-unitary dimension, there is moderately strong agreement on judicial review and central bank independence. For the other pairs, readers can draw their own conclusions about which of the two indicators measures the concept more validly. In the federalism graph, for example, the relationship looks good except for a cluster of nine outliers in the upper-left. The alternative index rates Costa Rica, New Zealand, Botswana, France, Japan, Sweden, Denmark, and Norway as more federal than unitary, which is odd by conventional constitutional criteria.

Figure 3: Lijphart's indicators (X) vs. nearest equivalents (Y)



However, if we regard the existence of elected officials at the regional and local level who are autonomous from unelected officials, these countries are correctly represented; and this may be a more relevant concept for measuring consensus democracy than what the constitution says about the structure of government. In the interest-group pluralism plot, the alternative indicator seems correct to place France, Portugal, and Iceland closer to the corporatist pole than to the pluralist pole, where Siaroff's index placed them.

When assessing validity, it is germane to consider how the data were generated. On the one hand, there is a presumption that Lijphart knew best: these are his concepts and dimensions, so we should defer to his judgment about how to define the concepts, which sources of information to use, and how to adjust the data to measure his concepts better. He was extremely diligent and transparent about all of these decisions. In fact, much of *Patterns of Democracy* is documentation of these decisions. On the other hand, there is also a risk that he unconsciously biased the measurements to support a predetermined conclusion. When the same person gathers the data and analyzes it, there is always a risk of bending the ratings to confirm a hoped-for conclusion. This is a natural human tendency. I think it is rarely intentional, but it is a risk we must take care to avoid. For this reason, many researchers prefer to use measures that were not created to analyze the question for which they are being used.

Lijphart made many operational choices that were wise, necessary, and eminently defensible. For example, when measuring the percentage of minimal-winning or one-party cabinets, he weighted cabinets by their duration. When calculating disproportionality, he excluded elections for uncontested seats when possible and excluded boycotted elections. Both of these choices eliminated some of the seat-vote gaps that were not products of the electoral system. For constitutional rigidity, he began with the size of the majority needed to amend the constitution. For judicial review, rather than limiting his attention to formal legal provisions, he did what he could to take judicial activism into account as well. Other scholars would do well to follow his lead in such choices.

However, Lijphart also made other choices that are puzzling and debatable:

- When calculating the effective number of parties, he treated all factionalized parties (which are hard to classify to begin with) as 1.5 parties, regardless of the number of factions or their degree of institutionalization.
- For the minimal-winning or one-party cabinets indicator, he counted minority cabinets as oversized. There is a rationale for it: minority cabinets have to rely on some parties in the opposition to pass their legislation, so they may in effect act like oversized cabinets. However,

it is also possible that they rely on just enough other party support to act like a minimal-winning coalition.

- It was reasonable for Lijphart to classify any cabinet representing 80 percent of the legislature as oversized even if it is minimal-winning. However, he made an exception to that when he classified France's 1993 cabinet as minimal-winning because its >80 percent majority was manufactured by the two-round majoritarian electoral system. If an indicator is based on seats, it is hard to justify arbitrarily using votes to reclassify one case; and it is not clear that votes are more relevant than seats for judging how consensual the cabinet is in the legislative arena.
- Japan had minimal-winning cabinets from 1967 to 1993, but Lijphart classified them as minority cabinets because Japanese politicians behaved according to consensus norms. If behavioral criteria can trump institutional criteria, this rule should have been applied consistently for all the cases.
- For coding executive dominance, he capped extreme values of cabinet stability for six countries at 5.52 on the grounds that no executive could possibly be more dominant in his or her country than the Prime Minister is in the UK.
- For the same indicator, he also "impressionistically" altered the values of cabinet stability for the presidential systems of Costa Rica, the US, Venezuela, and Colombia; and for Switzerland.
- To calculate disproportionality in presidential systems, Lijphart used the geometric mean of legislative and presidential disproportionality, which is not possible in parliamentary systems and arguably inserts a pro-consensus bias for parliamentary systems.
- To construct a scale of bicameralism, Lijphart rank-ordered the combinations of the number of chambers, the symmetry of the chambers, and their congruence. If these three aspects were correlated, they could be aggregated in many ways and lead to similar results. And the number of chambers necessarily has an alignment with symmetry and an alignment with congruence simply because symmetry and congruence are irrelevant unless there are two chambers. But there is no empirical relationship between symmetry and congruence: they are independent of each other. Therefore, there is no atheoretical way to combine them into an index; some theoretical guidance must be supplied, and it is lacking in Lijphart's rationale for this indicator, although it is clear that he considered symmetry more important than congruence.
- To his credit, Lijphart used the best index of corporatism available at the time, even before it was published (Siaroff 1998). However, Siaroff rated 21 countries for 1963-70 and 24 countries for 1983-90; 1945-1962, 1971-1982, and 1991-1996 – 71 percent of the sample for these countries – was missing data. Presumably he averaged the two periods; then he extrapolated to the missing periods, arguing that "These two periods may be considered representative for

the long time span from the late 1940s to 1996 used for the analysis of twenty countries in this study and for France in the period 1958-96" (176) – and, implicitly, for the early democratic years for Spain, Portugal, and Greece before 1983.

- Judicial activism changed dramatically in the period of this study in five cases. Rather than average values or using different values for different years, Lijphart used the value that lasted longest in each case, as though it accurately characterized the whole period.
- Lijphart had to be creative to combine the three indices of central bank independence. Because there was a lot of missing data ($N = 33, 18, \text{ and } 13$), he just averaged whichever indicator or indicators had a score for each country. No country had all three scores; 7 had only 1; 29 had 2. However, since they were on different scales, he rescaled them so that their maximums and minimums matched between 0 and 1, even though there was no way to ensure that rescaled values were commensurable.
- Furthermore, the CBI index with the most coverage was originally measured by decade, 1950s-1980s (and not for the 1990s). Lijphart averaged them and ignored the several cases of CBI reforms in the early 1990s, although he mentioned them. The data for the second and third indices is for unknown periods before 1991. Lijphart justified the extrapolations on the grounds that central bank independence changes little, so it's safe to extrapolate to other years (although he notes elsewhere that this was not true in the 1990s).

Between the impressionistic coding of developing countries, the idiosyncratic rules for counting parties, the debatable recoding of certain categories, the extrapolation of missing data, the choices to either average dynamic ratings or pick the most representative values, the rescaling of incommensurable indicators, and the hand-tweaking of values he considered implausible, Lijphart had ample opportunities to adjust the values of indicators in ways that could have nudged them toward a two-dimensional structure. I do not know that this happened. If it did, I certainly do not believe it was intentional. But it is a possibility that cannot be dismissed. When testing for the existence of one or two consensus-majoritarian dimensions I prefer to trust data that were not gathered for that express purpose.

IV. Substituting Indicators

Whether the original indicators or my alternative indicators are more valid, the choice of indicators makes a big difference for replicating Lijphart's patterns of democracy. Table 3 shows a principal components analysis of all ten alternative indicators, limiting the data to Lijphart's countries and years. Due to missing data, seven countries are not included, but as noted above, the omission of

these seven countries does not destroy the two-dimensional structure when using Lijphart's data, so there is little reason to except their omission to do so here. With the alternative indicators, there are four significant dimensions, with two or three indicators loading most heavily on each component. The first component clearly concerns party-system fragmentation. The second, combining Division of power (federalism), Legislative constraints on the executive, and Constitutional rigidity, could be interpreted as a dimension of constraints of the executive – by other levels of government, the legislature, and the constitution. The third component, combining Central bank independence, Judicial review, and Divided party control, seems uninterpretable, as does the fourth component, which combines Bicameralism and Interest-group pluralism. This analysis is not very meaningful, as the KMO statistic suggests that there are too few indicators to measure each dimension well. But whatever these dimension are, they are not Lijphart's two consensus-majoritarian dimensions.

Table 3: Principal Component Analysis of Alternative Indicators, Lijphart's sample

Factor	Eigenvalue	Proportion	Cumulative		
Factor1	2.171	0.217	0.217		
Factor2	1.759	0.176	0.393		
Factor3	1.619	0.162	0.555		
Factor4	1.515	0.152	0.706		
Component weights					
Variable	Factor1	Factor2	Factor3	Factor4	Uniqueness
Disproportional electoral system	0.846	-0.098	0.178	-0.133	0.225
Effective number of parties	0.810	0.180	-0.339	0.045	0.195
Division of power	-0.087	0.828	0.106	0.083	0.289
Legislative constraints on the executive	0.422	0.721	-0.103	-0.087	0.284
Constitutional rigidity	0.133	-0.589	0.183	0.550	0.300
Central bank independence	0.089	0.105	0.734	0.408	0.276
Judicial review	-0.092	-0.137	0.734	-0.252	0.371
Divided party control	0.495	-0.034	-0.582	0.071	0.411
Bicameralism	-0.365	-0.127	-0.043	0.764	0.265
Interest-group pluralism	-0.449	-0.344	0.006	-0.600	0.320
N=29, KMO=.598					

The two-dimensional solution also usually disappears with less radical substitutions. If we substitute just one of the alternative indicators at a time into the original model (i.e., using

Lijphart's sample and indicators but substituting one alternative indicator), we get a three-dimensional solution for seven out of ten models. The only substitutions that preserve the same two factors are the ones with the effective number of parties, legislative constraints, and disproportional electoral system in any combination of two or three of these indicators. In fact, these models offer slight improvements in fit over the original. However, it is safe to assume that most other substitutions would similarly disrupt the two-dimensional pattern.

Another possibility is that one could replicate Lijphart's finding by dropping a few of the alternative indicators that are least comparable to the original ones. One such model using just five indicators – effective number of parties, minimal-winning or one-party coalitions, disproportionality, judicial review, and central bank independence – reproduces Lijphart's two dimensions, but the indicator sampling is inadequate (KMO=.533). Lijphart's patterns of democracy cannot be successfully replicated in his own sample with most of the alternative indicators, much less several of them (other than the three mentioned) or subsets of them.

In order to find out how important the selection of countries and time periods is for patterns of democracy, we can retest the model in Table 3 using different samples. Table 4 summarizes the findings for samples that move the threshold for electoral democracy ("Polyarchy") above and below thresholds of .3, .5, and .7 on a 0-1 scale; vary the period from 1900-1945 to 1945-2016 to 1970-2016 to 1989-2016 to 1900-2016; and either require at least 19 years of an uninterrupted electoral regime or not. (Note: Samples before 1970 have to drop Garriga's central bank independence indicator; samples before 1945 have to drop Bormann and Golder's ENPS as well.) Table 4 lists the number of the component (1, 2, 3, 4) on which each indicator loads most strongly in each sample. These models destroy any hope of robustness for the four-dimensional model of Table 3. These models *always* yields four dimensions, but the indicators load on the first, second, third, and fourth dimensions in apparently random ways. All of the indicators except legislative constraints on the executive and division of power (federalism) load most strongly on at least three of the four dimensions in one model or another. Interestingly, legislative constraints and division of powers always load on the same dimension, but in Lijphart's original model their corresponding indicators always loaded on different dimensions. There is no rhyme or reason to these patterns of democracy. They are also too weak and noisy to be taken seriously as patterns at all.

Table 4: Failure to Replicate in Different Samples

	Most relevant dimension for each indicator										
	by sample										
Indicator	A	B	C	D	E	F	G	H	I	J	K
Effective number of parties	3	1	1	2	4			3	2	2	
Divided party control	4	1	1	4	4	2	3	3	4	4	3
Legislative constraints on the executive	1	2	2	1	2	1	1	1	1	1	1
Disproportional electoral system	2	1	1	3	1	2	3	2	4	4	3
Interest-group pluralism	1	1	2	2	3	4	1	1	1	1	1
Division of power	1	2	2	1	2	1	1	1	1	1	1
Bicameralism	1	4	3	4	1	3	4	4	3	3	2
Constitutional rigidity	2	4	3	3	3	2	2	4	3	3	2
Judicial review	3	3	4	1	4	3	4	3	2	2	4
Central bank independence	2	3	4	3	1			2	4	2	
N	2836	1242	1609	613	614	925	4976	1941	2604	2094	3481
KMO	0.635	0.535	0.513	0.644	0.502	0.429	0.661	0.539	0.578	0.570	0.509

Samples

A: All countries, 1970-2016

B: Lijphart countries, 1970-2016

C: Polyarchy \geq .7, 1970-2016D: Polyarchy \leq .5, 1970-2016E: .5 \leq Polyarchy $<$.7, 1970-2016F: Polyarchy \geq .3, 1900-1945G: Polyarchy \geq .3, 1945-2016H: Polyarchy \geq .3, 1989-2016I: Polyarchy \geq .3, 1970-2016J: Polyarchy \geq .3, 1970-2016, \geq 19 yearsK: Polyarchy \geq .5, 1900-2016, \geq 19 years

Conclusions

When concepts refer to patterns that do not really exist in the observable world, they are not useful and should be abandoned. Consensus and majoritarian democracy, whether on one dimension or two, are such concepts because they are overgeneralizations. Yes, certain aspects of power-sharing tend to occur together in certain contexts, and certain aspects of power-concentration also go together some of the time. But no one has identified even five attributes of majoritarianism that co-occur reliably, especially not in many historical and geographic contexts, and the same is true for consensus-promoting institutions. Maybe with some reformulation of concepts, new data collection, and further analysis, it will be possible to salvage some version of the consensus-majoritarian distinction that has some analytic utility. In the meantime, however, we would do well to theorize about the more specific components of these very general terms, such as separation of powers, division of powers, corporatism, party-system fragmentation, judicial review, and central bank independence. These concepts are general enough.

References

- Anderson, L. 2001. "The Implications of Institutional Design for Macroeconomic Performance: Reassessing the Claims of Consensus Democracy." *Comparative Political Studies* 34(4): 429–52.
- Bogaards, Matthijs. 2000. "The Uneasy Relationship between Empirical and Normative Types in Consociational Theory." *Journal of Theoretical Politics* 12(4): 395-423.
- Bormann, Nils-Christian. 2010. "Patterns of Democracy and Its Critics." *Living Reviews in Democracy*. <http://www.livingreviews.org/lrd-2010-3>.
- Bormann, Nils-Christian; Matt Golder. 2013. "Democratic Electoral Systems around the World, 1946-2011." *Electoral Studies* 32: 360.
- Michael Coppedge and John Gerring, with David Altman, Michael Bernhard, Steven Fish, Allen Hicken, Matthew Kroenig, Staffan I. Lindberg, Kelly McMann, Pamela Paxton, Holli A. Semetko, Svend-Erik Skaaning, Jeffrey Staton, and Jan Teorell, "Conceptualizing and Measuring Democracy: A New Approach," *Perspectives on Politics* 9:2 (June 2011): 247-267.
- 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 L. Marquardt, Kelly McMann, Valeriya Mechkova, Pamela Paxton, Daniel Pemstein, Laura Saxer, Brigitte Seim, Rachel Sigman, and Jeffrey Staton. 2017a. "V-Dem Codebook v7.1." 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, Joshua Krusell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Moa Olin, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Constanza Sanhueza Petrarca, Johannes von Römer, Laura Saxer, Brigitte Seim, Rachel Sigman, Jeffrey Staton, Natalia Stepanova, and Steven Wilson. 2017b. "V-Dem [Country-Year/Country-Date] Dataset v7.1." Varieties of Democracy (V-Dem) Project.
- Croissant, Aurel & Teresa Schächter (2009). Demokratiestrukturen in Asien – Befunde, Determinanten und Konsequenzen. *Zeitschrift für Politikwissenschaft* 19(3): 387-419.
- Fortin, Jessica. 2008. "Patterns of Democracy? Counterevidence from Nineteen Post-Communist Countries." *Zeitschrift für Vergleichende Politikwissenschaft* 2(2): 198-220.
- Ganghof, Steffen. 2010. "Democratic inclusiveness: A Reinterpretation of Lijphart's Patterns of Democracy." *British Journal of Political Science* 40(3): 679-692.
- Ganghof, Steffen. 2012. "Resilient Patterns of Democracy. A Comparative Analysis." *Zeitschrift für Vergleichende Politikwissenschaft* 6(S2): 103.

- Garriga, Ana Carolina. 2016. "Central Bank Independence in the World: A New Dataset." *International Interactions* 42(5): 849-868.
- Gerring, John, and Strom C. Thacker. 2008. *A Centripetal Theory of Democratic Governance*. New York: Cambridge University Press.
- Marco Giuliani. 2016. "Patterns of Democracy Reconsidered: The Ambiguous Relationship between Corporatism and Consensualism." *European Journal of Political Research* 55: 22–42.
- Horowitz, Donald L. 2001. *Ethnic Groups in Conflict*. Berkeley: University of California Press.
- Laakso, Maarku and Rein Taagepera. 1979. "Effective Number of Parties: A Measure with Application to West Europe." *Comparative Political Studies* 12:3-27.
- Lijphart, Arend. 1984. *Democracies: Patterns of Majoritarian and Consensus Government in Twenty-One Countries*. New Haven CT: Yale University Press.
- Lijphart, Arend. 1999. *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. New Haven CT: Yale University Press.
- Lijphart, Arend. 2003. "Debate: Measurement Validity and Institutional Engineering – Reflections on Rein Taagepera's Meta-Study." *Political Studies* 51: 20-25.
- Mainwaring, Scott. 2001. "Two Models of Democracy" (review). *Journal of Democracy* 12(3): 170-175.
- Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi-ting Wang, Joshua Krusell and Farhad Miri. 2017. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data." University of Gothenburg, Varieties of Democracy Institute: Working Paper No. 21, 2nd edition.
- Powell, G.Bingham. 2000. *Elections as Instruments of Democracy: Majoritarian and Proportional Visions*. Yale University Press.
- Siaroff, Alan. 1998. "Corporatism in Twenty-Four Industrial Democracies: Meaning and Measurement" (unpublished manuscript).
- Taagepera, Rein. 2003. "Arend Lijphart's Dimensions of Democracy: Logical Connections and Institutional Design." *Political Studies* 51: 1-19.
- George Tsebelis, *Veto Players: How Political Institutions Work*. Princeton University Press, 2002.
- Adrian Vatter. 2009. "Lijphart Expanded: Three Dimensions of Democracy in Advanced OECD Countries?" *European Political Science Review* 1(1): 125.
- Vatter, Adrian and Julian Bernauer. 2009. "The Missing Dimension of Democracy Institutional Patterns in 25 EU Member States between 1997 and 2006." *European Union Politics* 10(3): 335-359.