



Better Regime Cutoffs for Continuous Democracy Measures

Yuko Kasuya
Kota Mori

October 2019

Users Working Paper

SERIES 2019:25

THE VARIETIES OF DEMOCRACY INSTITUTE



UNIVERSITY OF GOTHENBURG
DEPT OF POLITICAL SCIENCE

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

Please address comments and/or queries for information to:

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

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

Copyright © 2019 by authors. All rights reserved.

Disclaimer: V-Dem does not do quality control and therefore does not endorse the content of the papers, which is the responsibility of the authors only.

Better Regime Cutoffs for Continuous Democracy Measures

Yuko Kasuya* Kota Mori†

October 17, 2019

*Professor, Department of Political Science, Keio University (ykasuya@law.keio.ac.jp)

†Independent scholar (kmori05@gmail.com)

Abstract

Many vital questions in comparative politics and international relations concern the transition to and from democracy. This tradition makes the measurement of democracy, in particular, the threshold between democracy and non-democracy, a crucial issue. However, the field is filled with arbitrary and inconsistent uses of democracy cutoffs. We examine the validity of the cut points suggested by the most frequently used continuous measures of democracy, namely, Freedom House (FH), Polity, and the Varieties of Democracy (V-Dem). Our analyses use the binary measures of democracy as the reference because they reflect a theoretically informed set of conditions of democracy, a feature that is lacking in the continuous measures' cut points. In particular, we presume that the preferred threshold should achieve the highest consistency with the binary measures' coding when the regime is defined based on that threshold. Our analyses find that such cutoffs are 3.7 for FH, 4.5 for Polity, and 0.42 for V-Dem, rather than those suggested by the providers of these measures (2.5, 6, and 0.5 respectively). This finding implies that three continuous measures of democracy set standards for democracy higher than the binary measures.

Keywords: democracy measures, measurement validity, Freedom House, Polity, V-Dem
6907 words.

Acknowledgement

A previous version of this paper was presented at the V-Dem Institute lunch seminar. The authors thank seminar participants, especially Vanessa Boese for her helpful comments.

1 Introduction

The cross-national measurement of democracy is receiving a renewed attention from scholars, journalists, and policy practitioners alike, due to the recent worldwide trend in democratic backsliding (Lührmann and Lindberg, 2019). Efforts to measure the characteristics of political regimes, however, have a long history. Starting in the 1950s, many attempts to measure the characteristics of political regimes have been made (Dahl, 1956; Lipset, 1959; Cutright, 1965). Over the years, scholars have created a number of democracy measurements with extensive country and period coverage (*e.g.*, Gastil, 1990; Gurr, Jagers and Moore, 1990). More recently, these have been executed with increasing technical sophistication (*e.g.*, Pemstein, Meserve and Melton, 2010; Coppedge, Gerring, Altman, Bernhard, Fish, Hicken, Kroenig, Lindberg, McMann, Paxton et al., 2011; Skaaning, Gerring and Bartusevičius, 2015).

Despite the recent advances, however, the study of democracy measures still lacks a set of standards regarding when to use which index and how. One of these lacunae is the question of how to use continuous democracy measures as if they are dichotomous. In other words, the issue here is where to draw the line to distinguish between democracy and non-democracy for continuous democracy measures such as Freedom House (FH), Polity, and the Varieties of Democracy (V-Dem). As we discuss below, the cut points suggested by the providers of these measures lack reasonable justifications. Moreover, scholars have used these cut points in a rather arbitrary manner, which often resulted in inconclusive or inconsistent results (Bogaards, 2012).

In this paper, we examine the measurement validity (Adcock and Collier, 2001) of democracy cut-points used in FH, Polity, and the V-Dem. Our review of the related literature demonstrate that the cutoffs suggested by the providers of these measures lack systematized definition of the borderline between democracy and non-democracy. Against this backdrop, our analyses employ dichotomous measures that are based on a clear-cut definition of democracy and coded with simple coding rules (Boix, Miller and Rosato, 2013; Cheibub, Gandhi and Vreeland, 2010; Geddes, Wright and Frantz, 2014a) as the benchmark. Using these, we identify the cut points for the continuous measures that best approximate the dichotomous measures' regime classification. In our analyses, 3.7 for FH, 4.5 for Polity, and 0.42 for V-Dem are the cut points that reduce the mismatch to the greatest extent. Conceptually, these cut-points are considerably less restrictive than the ones suggested by the provider (2.5, 6, and 0.5 respectively). This finding implies that the existing cutoffs of the three continuous measures of democracy classify a larger

number of country-years as non-democracy than the classification provided by the dichotomous measures.

Our findings have significant implications for the users of FH, Polity, and V-Dem, including scholars, educators, and journalists. First, anyone using the regime classifications according to the providers of these continuous measures should handle them critically. The world maps and time series graphs of regime classifications, especially those using FH and Polity, are so often used in media reports, classroom teaching, and academic analyses for illustration purposes. Those who use, as well as those who are presented with such visualization, appear to consume them at their face value. Our analyses warn such uncritical attitudes and inform that those classifications lack theoretical justifications and set higher standards of democracy threshold than the binary measures.

Second, this paper contributes to the field of comparative politics and international relations, particularly in the context when scholars employ these continuous democracy measures as a dichotomous variable. Many scholars have lamented that depending on the choice of measures and the cutoff points, the results of statistical analyses differ to a considerable degree (Elkins, 2000; Casper and Tufis, 2003; Bogaards, 2010; Cheibub, Gandhi and Vreeland, 2010). Our suggested cut points can be useful for scholars whose research addresses regime-related issues, as ours is based on a set of theoretical justifications while existing cutoffs, are not. With this concern, this paper contributes to the accumulation of knowledge in this research field.

In what follows, we first discuss the conceptualization and measurement of continuous democracy measures to demonstrate the lack of theoretical justification of their democracy cutoffs. The section that follows discusses our analytic strategies and results. We then provide an application to the Philippine case to show how our suggested cutoffs perform vis-a-vis the provider suggested ones. The last section concludes with the discussion of future research avenues.

2 Using Continuous Measures as if They Are Categorical

In the literature of democracy measurements, one of the important debates discusses whether to view democracy as a dichotomous or continuous construct (Collier and Adcock, 1999; Goertz, 2006). On the one hand, some scholars maintain the view that a country is either in a state of democracy or non-democracy; thus there is no grey zone (*e.g.*, Linz, 1975; Sartori, 1987; Hunt-

ington, 1991; Alvarez, Cheibub, Limongi and Przeworski, 1996). Building on this conception of democracy, scholars developed the databases that code political regime as a dichotomous variable (*e.g.*, Przeworski, Alvarez, Alvarez, Cheibub and Limongi, 2000; Cheibub, Gandhi and Vreeland, 2010; Boix, Miller and Rosato, 2013; Geddes, Wright and Frantz, 2014a).

On the other hand, the conceptual structure of democracy is a continuum for some scholars. For example, Dahl (1973, p.241) believes that “every country is located somewhere along the underlying continuum” of being democratic.¹ This view of democracy is reflected in the continuous measures of democracy developed in studies such as those by Bollen (1980); Gastil (1985, 1990); Gurr, Jagers and Moore (1990); Vanhanen (2000); Coppedge et al. (2011) .

Our concern is not this debate,² but the ways in which continuous measures are treated as dichotomous, or in more general terms, categorical variables. This issue concerns many scholars in the fields of comparative politics and international relations fields because for decades, many studies have used the continuous measures as if they are dichotomous (democracy vs autocracy) or categorical (*e.g.*, democracy vs semi-democracy vs autocracy). Enquires concerning “democratic peace,” regime transition, and democratic diffusion are some of the major examples of research contexts where scholars need to use categorically classified regime types derived from continuous democracies. Indeed, it has become almost an industry standard to use the cutoffs suggested by providers (2.5 for FH and 6 for Polity) to the extent that they appear in influential textbooks of comparative politics (Clark, Golder and Golder, 2017, pp.157-160).

Some scholars have criticized that users of FH and Polity have picked different cut-points than what the providers endorsed. For example, Bogaards (2012) finds 14 different ways to set the cut points of FH other than the provider-suggested value, and 18 different ways for Polity. Bogaards (*ibid*) rightly criticizes that those scholars who set their own cutoff have done so without reasonable justification, particularly those who modified the FH cutoff.

However, as far as we are aware, few scholars have questioned why those cut points are set as they are in the first place. In some of the most famous critiques of these democracy measures, there is little mention on this point (*e.g.*, Munck and Verkuilen, 2002; Goertz, 2006). We believe this is a serious oversight in the studies of regime measurement because, as we discuss below, the provider-suggested cutoffs lack theoretical justifications, which makes it almost impossible

¹Goertz (2006, Chapter 4) analyzes that Dahl conceptualizes democracy (in his terminology, pol-yarchy) as a necessary-sufficient condition structure of the concept. If so, according to Goertz, dichotomous measurement should have been more in line with the construct of this theory of democracy.

²On this debate, see Collier and Adcock (1999); Goertz (2006).

to evaluate the validity of the suggested cut-points.

In our review of the codebooks and related documents, none of three continuous measures of democracy provides theoretical discussion as to why they set the cut-points as they do. In the case of FH, countries are given the status of “Free” (1.0 to 2.5), “Partly Free” (3.0 to 5.0), or “Not Free” (5.5 to 7.0). Scholars often equate democracy with “Free” status (*e.g.*, Lijphart, 2012), or “Free” and “Partly Free” combined (*e.g.*, Starr and Lindborg, 2003). However, we could not find a clear explanation of why these values are assigned to each category in the Freedom House’s methodology document.³ The Polity website explains that country-years that have a Polity score of 6 and above should be “democracy,” but we could not find any rationale for setting this cut point.⁴ At the same time, the terms they introduce appear to be inconsistent. In the latest version of the codebook (Marshall and Gurr, 2016), a Polity score of 7 or above is called “full democracy” (*ibid.*, p.35), while in their website labels a score of 10 as “full democracy.” We could not find any explanation as to why this difference has arisen or what is the difference between the notions of “full democracy” and “democracy.”⁵

The V-Dem database provides five conceptually different models of democracy, and in this paper we focus on the model called “Electoral Democracy” (Coppedge et al., 2011) because it is the closest model to other democracy measures we investigate in this paper. Here, we call this the V-Dem measure. A recent research article that was written by the providers of V-Dem elaborates on how one can classify political regimes following Robert Dahl’s notion of polyarchy (Lührmann, Tannenberg and Lindberg, 2018). In operationalizing their conceptual scheme, those authors propose that three necessary conditions must be fulfilled for a country to qualify as a democracy: (1) de facto multiparty elections, (2) free and fair elections, and (3) a score higher than 0.5 on the Electoral Democracy Index. The first two are coded through an ordinal scale, and in order to pass these conditions a given country-year needs to satisfy at least having “one real opposition party [that] is allowed to contest” and “substantial competition and freedom of participation”(Coppedge and Ziblatt, 2019, p.55, p.60).⁶ The third condition relies on a cut-point set for an interval scale (Electoral Democracy index), which is one of the “most abstract” (Coppedge and Ziblatt, 2019, p.39) level indices within V-Dem indices. We

³<https://freedomhouse.org/report/methodology-freedom-world-2019> (accessed August 1, 2019).

⁴<http://www.systemicpeace.org/polity/polity4x.htm> (accessed July 26, 2019).

⁵In addition to the codebooks of all versions of the Polity database and the website, we consulted Marshall, Gurr, Davenport and Jagers (2002).

⁶These variables are `v2elmulpar_osp` and `v2elfrfair_osp`.

find little theoretical justification about why 0.5 should be the threshold, and, unlike the other two conditions, it is difficult to interpret what 0.5 signifies given the highly abstract nature of the Electoral Democracy indicator. The authors of the article also write that the cut-point of 0.5 is “admittedly arbitrary”(Lührmann, Tannenberg and Lindberg, 2018, p.5).

In summary, the cut points recommended by the three major continuous democracy measures lack theoretical grounds as to why these cutoff values are suggested. This raises the issue of measurement validity: we are not certain to what extent the suggested cut-point values correspond to the actual transition from or to democracy to non-democracy. In reference to the sequence of concept measurement suggested by Adcock and Collier (2001), the manner in which cutoff points are set in FH, Polity and V-Dem lack the discussion on the concept to be measured (the “systematized concept” in Adcock and Collier’s terminology), and only the measurement (cut-points) are presented. At the same time, few studies have focused on studying the provider-suggested cutoffs’ empirical relevance.⁷

Against this backdrop, some scholars advise against using the democracy cut points but taking out the component’s indicators of Polity. For example, Bogaards (2012) suggests choosing the components that presumably measure the phenomenon of interests, and using them either as is or as a combined alternative measure. As he cites in his article, a number of studies have taken this route. Instead, we try to save the idea of using cut points for the continuous measures of democracy.

3 Finding Better Cut Points

Our strategy employs dichotomous measures as the benchmark of regime classification and identifies a cutoff that achieves the highest consistency between the dichotomous measures and the regime defined with that cutoff. Put differently, we conduct a “criterion validation” (Adcock and Collier, 2001) by taking the dichotomous democracy measures as a reference. As we discuss in detail below, the dichotomous measures create the democracy cutoffs based on a theoretical and empirical examination, which is a lacking feature among the continuous measures discussed above.

⁷One important exception is (Lührmann, Tannenberg and Lindberg, 2018), which provides comparison of various democracy measures vis-a-vis their suggested classifications. They find similar results as ours (V-Dem cutoff is more restrictive than BMR, CGV, and GWF), yet the lack of theoretical grounds of 0.5 cut-points remains as an issue. Boix, Miller and Rosato (2013, p.1536-1537) provides a brief discussion comparing BMR, CGV, FH as well as several other measures.

	Measurement goal	Attributes	Components of attributes	Scale	Cut points	Country	Years
Continuous measures							
Freedom House	freedom	contestation participation governance	political rights (4 items); civil rights (4 items)	1 to 7	1.0-2.5 Free 3.0-5.0 Partly free 5.5 - 7.0 Not free	195	1973 to present
Polity	authority structure	contestation participation constraints	competitiveness/openness of exec. recruitment; compe- titive participation; constraints on executive	-10 to 10	≥ 6 as democracy	167	1800 to present
V-Dem	Dahl's polyarchy	contestation participation	freedom of expression and association; clean election; elected officials; majority suffrage	0 to 1	≥ 0.5 as democracy plus multiparty/free and fair election conditions	202	1900 to present
Dichotomous measures							
BMR	Dahl's polyarchy	contestation participation	elected executive and legislature; adult male suffrage	0 or 1	1 = democracy 0 = autocracy	222	1800 to 2015
CGV	minimalist democracy	contestation	elected exec. and legislature; multi-party competition; alternation of power	0 or 1	1 = democracy 0 = autocracy	202	1946 to 2008
GWF	autocracy type	contestation	leadership group is elected by free and fair election	0 or 1	1 = democracy 0 = autocracy	182	1946 to 2010

Table 1: Comparison of democracy measures

Table 1 summarizes the main features of dichotomous and continuous measures of democracy under our study. While there are many other attempts to measure democracy,⁸ we choose the ones in Table 1 because they are most frequently used in academic research and have extensive coverage in the number of countries and in the time period.

It is worth noting that FH and Polity started out measuring something other than democracy while most critics of FH and Polity appear to overlook this aspect.⁹ Raymond Gastil, the initiator of the Freedom House Survey in 1973, recalled in his 1990 publication that the original intention was to “produce ... an orienting discussion of variation in *level of freedom*” (Gastil, 1990, p.25, emphasis added), and “it was years before that the author understood that the survey was essentially a survey of democracy”(ibid, p.26). In the case of Polity, Gurr’s study of political system persistence published in 1974 was the foundation of this database (Jaggers and Gurr, 1995, p.470). In order to examine the regime duration, he constructed a database of “authority characteristics”¹⁰ of “polities,” by which he meant political systems, as distinguished from nation-states. He goes on to note that his study does not intend to “label a polity as a ‘democracy’ or an ‘autocracy’,” and observes that “most systems in fact have mixed authority patterns” (Gurr, 1974, pp.1487-1488). The Polity III database, launched in 1995, was the first time when the Polity database started to provide “democracy score” and “autocracy score”(Jaggers and Gurr, 1995), and the variable name called “Polity”, which is the combined democracy score and autocracy score in Polity III, starts to appear only in the Polity IV database. These origins and the path-dependent nature of FH and Polity might explain why many scholars found the conceptualization of these measures problematic (*e.g.*, Munck and Verkuilen, 2002). Among the three continuous measures we study, it is only V-Dem that has made explicit attempts to measure the concept of democracy, especially the one proposed by Dahl for “Electoral Democracy”(Dahl, 1973; Coppedge et al., 2011; Teorell, Coppedge, Lindberg and Skaaning, 2019).

We take advantage of dichotomous measures’ two desirable properties. First, in drawing the cutoff line, they follow clear-cut coding procedures that are founded on transparent theoretical considerations: A given country-year is coded as democracy only when it fulfills a certain set of necessary conditions (details discussed below). The second desirable feature is that the

⁸ See, for example, Coppedge, Gerring, Lindberg, Skaaning and Teorell (2017) and Boese (2019) for recent reviews of the existing democracy measures.

⁹An exception is Giannone (2010) and Bush (2017) , who examine the historical evolution of FH, and Gleditsch and Ward (1997) for Polity.

¹⁰The five indicators of authority characteristics are: (1) recruitment, (2) decision constraints, (3) directiveness (regulation), (4)participation, and (5) complexity of decision-making (Gurr, 1974).

lines are drawn by looking at actual political events in each country they code. In contrast, the continuous democracy measures set a single threshold across all cases without systematically investigating the individual cases.¹¹

Among the three dichotomous measures listed in Table 1, we mainly employ BMR as our benchmark, while using CGV and GWF in our supplementary analyses. This is because BMR's definition of democracy is closer to that of the three continuous measures than CGV and GWF. In creating BMR, Boix, Miller and Rosato (2013) explicitly follow Dahl's concept of polyarchy, and codes a given country-year as democracy, if the case meets the the following three conditions: (1) The executive is directly or indirectly elected in popular elections and is responsible either directly to voters or to a legislature; (2) The legislature (or the executive if elected directly) is chosen in free and fair elections; (3) The majority of adult men have the right to vote (ibid., 1530-1531). The first two conditions concern "contestation" and the third condition addresses the "participation" dimension of polyarchy (Dahl, 1973).

With the goal of measuring "minimalist" notion of democracy, the creators of CGV define democracy as a regime in which "government officials are filled as a consequence of contested elections" (Cheibub, Gandhi and Vreeland, 2010, p.69). According to this definition, they code a country-year as democratic if it meets all four of the following conditions: (1) the chief executive must be chosen by popular election or by a body that was itself popularly elected, (2) the legislature must be popularly elected, (3) there must be more than one party competing in the elections, and (4) an alternation in power under electoral rules identical to the ones that brought the incumbent to office must have taken place¹² (ibid, p.69). In this regard, CGV conceptualizes democracy's main attribute as one of contestation, while participation is not included.¹³

For GWF, we could not find a straightforward definition of democratic regime in either the article that introduces the database (Geddes, Wright and Frantz, 2014a) or in the codebook (Geddes, Wright and Frantz, 2014b). This is probably because the providers' core concern is

¹¹V-Dem's is a partial exception because in addition to 0.5 cutoff, two additional conditions described above are coded by investigating individual county-year observation.

¹²Some scholars argue that CGV's requirement of electoral turnover makes this indicator more restrictive than BMR. For example, Bogaards (2012, p.702) notes that this condition makes it particularly difficult for Africa to be classified democratic because of the prevalence of predominant party systems without government turnover in the region. However, in our empirical analyses, CGV-based cutoffs are less restrictive than BMR-based (see Appendix A).

¹³Although suffrage is not included in CGV, as Cheibub, Gandhi and Vreeland (2010) note, the condition of suffrage (participation criteria) is practically met in CGV because for the period it covers, suffrage was already achieved in more of the countries.

with classifying various types of autocratic regimes, and democracy is measured as a residual category of autocracy. Indeed, their codebook provides a definition of “undemocratic” regime.¹⁴ By inference, one can say that GWF’s definition of democracy is a regime where the government is directly or indirectly elected by reasonably fair competitive election. Thus, similar to CGV, Table 1 lists contestation as its core attribute.¹⁵

In sum, and as summarized in Table 1, BMR’s attributes of democracy have better overlaps with those of three continuous measures than CGV and GWF. An additional advantage of BMR is its long historical coverage: BMR’s coding starts from 1800, while CGV and GWF only code the period since World War II. Using BMR can cover a greater number of years that were coded in Polity and V-Dem that go back to 1800 and 1900 respectively. For these reasons, we use BMR as the main benchmark and employ CGV and GWF in supplementary analyses.

Figure 1 shows the distribution of three continuous democracy measures according to the BMR’s regime classification. The unit of observation is the country-year. The light grey bars indicate the densities for country-year pairs coded as a democracy in BMR, and the dark grey bars are those coded as an autocracy in BMR. Solid lines are the kernel density estimates. Roughly speaking, the intersection of the two density plots is the cut point that achieves the highest consistency with the BMR coding. In other words, if we were to set the threshold at this level and classify the regimes accordingly, then the fraction of matches with the BMR would become larger. As a reference, the cutoffs suggested by data providers are shown as dashed lines in the figure. Consistently across the continuous measures, the provider-suggested cutoffs are more “democratic” than the intersection of the two density plots.¹⁶ This suggests that to be classified as a democracy in FH, Polity, and V-Dem under the provider-suggested cutoffs, a county needs to democratize more than in the case of BMR.

Figure 2 conducts the above analysis more explicitly. We plot the threshold level against the accuracy (fraction of matches with the BMR classification) separately for three continuous

¹⁴Geddes, Wright and Frantz (2014*b*, p.6) states:““Undemocratic” is defined as any means other than a direct, reasonably fair competitive election in which at least ten percent of the total population (equivalent to about 40 percent of the adult male population) was eligible to vote; or indirect election by a body at least 60 percent of which was elected in direct, reasonably fair competitive elections; or constitutional succession to a democratically elected executive.”

¹⁵See Geddes, Wright and Frantz (2014*a*) for the comparison between CGV and GWF. We did not include participation as an attribute of GWF because their voter condition does not require the majority of the population.

¹⁶For FH, the intersection is to the left of the dashed line. Since the FH score is lower for more democratic regimes, the direction is same as the other two measures.

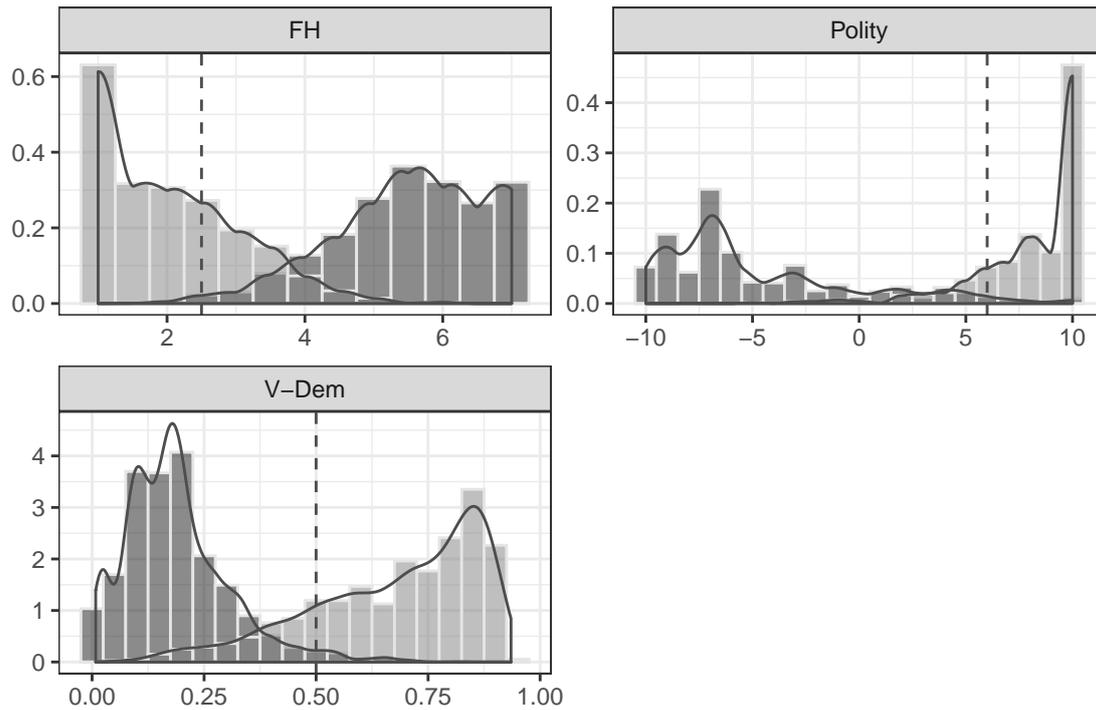


Figure 1: Distribution of continuous democracy measures by BMR classification

measures.¹⁷ Dotted lines indicate the highest accuracy levels and their corresponding cutoffs,¹⁸ and dashed lines the provider-suggested cutoffs. This figure confirms that, in all cases, the provider-suggested cutoffs are located at a more democratic level than the peak of the curves.

One way to interpret the results revealed in Figures 1 and 2 is to go back to the ways in which the components of democracy attributes are set across these measures (see Table 1). In the case of FH, its components of democracy attributes cover issues related to not only political rights but also civil liberties. As often criticized (*e.g.*, Munck and Verkuilen, 2002), some of the items in the civil liberty component go beyond the procedural definition of democracy,¹⁹ while BMR and the rest of the measures in Table 1 adhere to the procedural definition. To qualify as a democracy in this “maximalist” definition of democracy, a country would need to surpass many criteria, although as we reviewed above, a theory-grounded threshold criterion are absent

¹⁷In this paper, thresholds are defined as “inclusive to the democracy.” For example, a threshold 0.4 for V-Dem would mean that a country-year pair is coded as democracy if and only if the score is 0.4 or larger. Similarly, a threshold 3.0 for FH would mean that those with a score of 3.0 or lower are coded as democracy.

¹⁸In case multiple cutoffs achieve the highest accuracy, we chose the median of such cutoffs.

¹⁹For example, substantive matters such as “socioeconomic rights,” “freedom from gross socioeconomic inequalities,” “property rights,” and “freedom from war” are part of the indicators of democracy (Munck and Verkuilen, 2002, p.9).

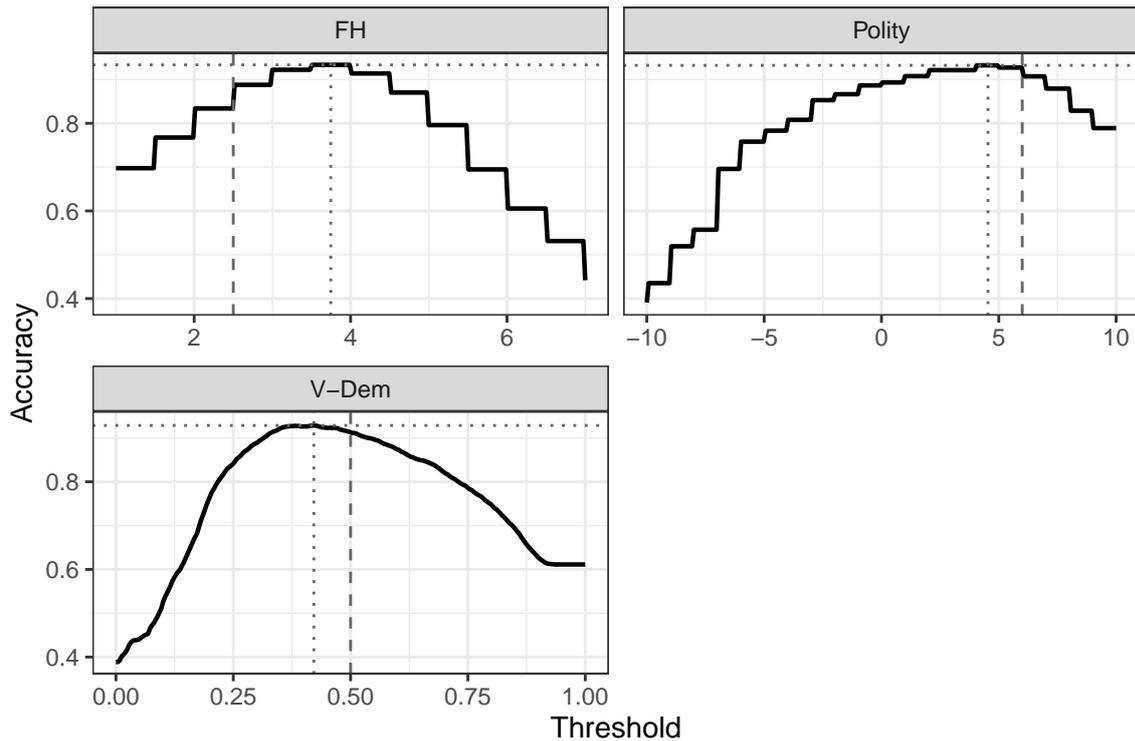


Figure 2: Comparison of cutoffs

in FH's methodology.

In a similar line of reasoning, one can interpret why the gap exists between the provider suggested cut points and the best accuracy estimates of Polity and V-Dem. As we listed in Table 1, BMR on the one hand and Polity and V-Dem on the other have different components of democracy attributes. Although all of them adopt the procedural notion of democracy in principle, BMR employs a bare-bones version of democratic procedure. This is reflected in their components of democracy attributes, which are to have elected executive and legislative branches plus adult male suffrage. In contrast, Polity and V-Dem embrace additional components. In the case of Polity, the constraints on the chief executive during and in-between elections are additional aspects. In V-Dem, the criteria of freedom of expression and association, and clean election are added to the bare-bone conditions stipulated in BMR. The existence of these additional considerations in Polity and V-Dem might explain why the provider-suggested cutoffs are more restrictive than BMR-based cutoffs. Nevertheless, these interpretations are merely speculative, and they should not legitimize the provider-suggested cutoffs as they are, because the fact that the providers of Polity and V-Dem are silent about their cutoffs' theoretical grounds remains unchanged.

Table 2 summarizes the best-accuracy and the provider-suggested cut points with their re-

spective accuracy scores with regards to the BMR coding. While the providers of FH, Polity, and V-Dem suggest democracy cutoffs of 2.5, 6, and 0.5 respectively, our analyses indicate that the best accuracy cut points are 3.7 for FH, 4.5 for Polity, and 0.42 for V-Dem. With regard to the degree of accuracy, one might think that there are only slight differences between the two groups. This is because of the fact that most of the country-years in the databases lie at the zones of clearly democratic or clearly autocratic, while the country-years that are at the borders of regime transition and are subject to measurement disagreements constitute a small fraction of cases. This points to the same observation that was made in Coppedge et al. (2011, p.252): the major democracy measurements have very high levels of correlation overall, but their overlaps decrease significantly when the country-years of extreme values on democracy scores are removed. At the same time, our analyses show that when it comes to setting the democracy threshold, the small number of cases at the border zones make a big impact in setting the cutoffs. In addition, it corroborates what was revealed in figures 1 and 2: the accuracy improvement from the provider’s cut-point to our suggested cut-point is the largest in FH (from 88.7% to 93.3%), followed by V-Dem (from 92.7% to 93.2%), then Polity (from 92.7% to 93.2%).

	Best Accuracy		Provider Suggested		N. of Obs
	Cut Point	Accuracy (%)	Cut Point	Accuracy (%)	
FH	3.7	93.3	2.5	88.7	6,077
Polity	4.5	93.2	6	92.7	10,891
V-Dem	0.42	92.8	0.5	91.2	11,763

Table 2: Cut point estimates using BMR as reference

Table 3 compares confusion matrices between the BMR’s classification and the classification derived by continuous measures with two alternative cutoffs: the one suggested by the data provider and the one that achieves the highest accuracy. Each table takes the BMR’s classification in rows and the “derived” classification in columns. The first number in each cell corresponds to the fraction of matches in the provider-suggested cutoff, and the second number is that of the highest-accuracy cutoff (our suggested threshold). Notice that the fractions are computed by the BMR’s classification, *i.e.*, the number of cases divided by the row totals. The top-left cell in a table is called “true negatives” (TN) in the sense that the derived classification correctly codes the cases as zero (autocracy). The top-right cell is called “false positives” (FP) since the derived classification codes them as one (democracy), which turns out to be incorrect (in BMR). Analogously, we call the bottom-left and bottom-right “false negatives” (FN) and

“true positives” (TP), respectively. The overall accuracy, which we maximize when finding our threshold, is defined as $(TN + TP) / (TN + FN + FP + TP)$, where each term denotes the number of cases that fall in the corresponding cell.

	Autocracy (above cutoff)	Democracy (below cutoff)
Autocracy (BMR=0)	98.6% , 93.0%	1.4%, 7.0%
Democracy (BMR=1)	23.6%, 6.3%	76.4%, 93.7%

FH Provider suggested a cutoff: 2.5; the cutoff that achieves the highest accuracy: 3.7.

	Autocracy (below cutoff)	Democracy (above cutoff)
Autocracy (BMR=0)	96.3% , 94.1%	3.7%, 5.9%
Democracy (BMR=1)	12.8%, 8.1%	87.2%, 91.9%

Polity Provider suggested a cutoff: 6; the cutoff that achieves the highest accuracy: 4.5.

	Autocracy (below cutoff)	Democracy (above cutoff)
Autocracy (BMR=0)	97.6% , 95.6%	2.4%, 4.4%
Democracy (BMR=1)	18.7%, 11.5%	81.2%, 88.5%

V-DEM Provider suggested a cutoff: 0.5; the cutoff that achieves the highest accuracy: 0.42.

Table 3: Confusion matrices with BMR as reference

Note: Cross-tabulation of the BMR's regime classification and the derived classification, under the provider suggested threshold (left), and the threshold that achieves the highest accuracy (right). Each number represents the fraction within the BMR classification; *i.e.*, the number of cases divided by the total cases by rows. The larger numbers in the top-left and bottom-right cells are shown in bold letters.

Table 3 reveals additional details of our analyses. First, the provider-suggested cutoff achieves a near 100 percent match for the country-year pairs coded as zero (autocracy) in the BMR, while there is a relatively low rate for those coded as one (democracy). This happens because, as we already analyzed above, the provider-suggested cutoffs are restrictive and tend to “misclassify” country-year pairs coded as (marginally) democratic by the BMR as an autocracy. Our suggested cutoffs improve overall accuracy by improving the matches for the cases coded as a democracy (BMR=1) with a cost of mismatches for those coded as autocracy (BMR=0). As a result, in all of the three measures, our suggested cutoffs reduce the imbalance between the top-left and bottom-right cells over the provider suggested cutoffs. In this manner, our suggested cutoffs mitigate the downward bias (restrictive standard) we discussed above.

In summary, our statistical analyses inform that the providers of the three continuous measures drew the democracy cutoff line too restrictively. If we employ the assumption that BMR is a valid dichotomous measure of democracy, better cut points are 3.7 for FH, 4.5 for Polity, and 0.42 for V-Dem rather than those suggested by the providers of these measures (2.5, 6, and 0.5, respectively). In Appendix A, we provide the results of the same analysis using CGV and GWF, and find qualitatively equivalent results.

In the next section, we will examine how these two sets of cut points perform using a country example.

4 Country Study

We employ the Philippines as an example to compare the performance of the providers’ and our suggested cutoffs. The Philippine case was chosen for two reasons. First, because the country has experienced several regime transitions since independence, it has often been on the border zone of democracy and autocracy. This feature makes the Philippines more conducive to examining the relevance of cutoffs than those countries that have remained highly democratic or highly authoritarian over the years. Second, the authors are familiar with the political history of the country, and can, therefore confidently evaluate the regime transition based on the BMR’s coding rules discussed in Section 3. Some might argue that the selection of this case was arbitrary and based on convenience, and we agree with this criticism. However, we do not think that this is a problem because our case illustration is in line with what our statistical analyses have already shown. In other words, we are not picking up an outlier to argue against

the general tendency.

Let us first describe the regime trajectory of the Philippines since independence to the present, as understood by many country specialists. Three periods of different regime types existed in the Philippines: (1) post-independence democracy (1946 July-1972 September), (2) Marcos' dictatorship (1972 September-1986 February), and (3) post-Marcos democracy (1982 February to the present). Since independence from the United States in July 1946, until President Ferdinand Marcos declared Martial Law in September 1972, the Philippines was democratic with relatively free and fair elections with universal suffrage.²⁰ The period of autocratic rule under Marcos continued until February of 1986 when Marcos was ousted through the "People Power Revolution." During this four-day event, thousands of Filipinos went onto the streets to call for the resignation of Marcos in the wake of the presidential election that people believed Marcos stole from his opponent Corazon Aquino. After the ouster of Marcos, technically speaking, a free and fair election of the chief executive did not happen until 1992, thus if one strictly follows the BMR's coding rule, the country was not a democracy until 1992. However, there is almost unanimous consensus among the Philippine experts that democratization happened in 1986, not in 1992. Also in BMR, 1986 is the transition year. Since 1986, a period of democracy continued since the executive and legislature are elected through relatively free and fair elections.²¹ We use the above periodization and the coding of country-years as the benchmark in this paper.

²⁰Our coding of the point of democratic breakdown differs from the BMR's start year of 1966. As a matter of fact, President Marcos was first elected in the 1965 election, and re-elected in 1969. According to the rules set by BMR, the Philippines should be a democracy until 1971, since they state that the time of regime breakdown is "from the point at which electoral contestation becomes uncompetitive or manipulated," but not at the point the ruler who destroyed the competitive elections started his/her tenure as in the CGV's operationalization (Boix et al. 2013, pp.1535-1536). In the BMR's coding, Philippine's autocracy started in 1966, and we believe BMR is not following its own coding procedure for the Philippines from 1966 to 1972.

²¹There is a recent debate about whether the Philippines under President Duterte (2016 to present) is a democracy or not due to his "illiberal" politics. In view of the BMR's definition, however, we maintain that the Duterte era should be classified as a democratic political regime because the executive and the legislature are both elected through relatively free and fair elections.

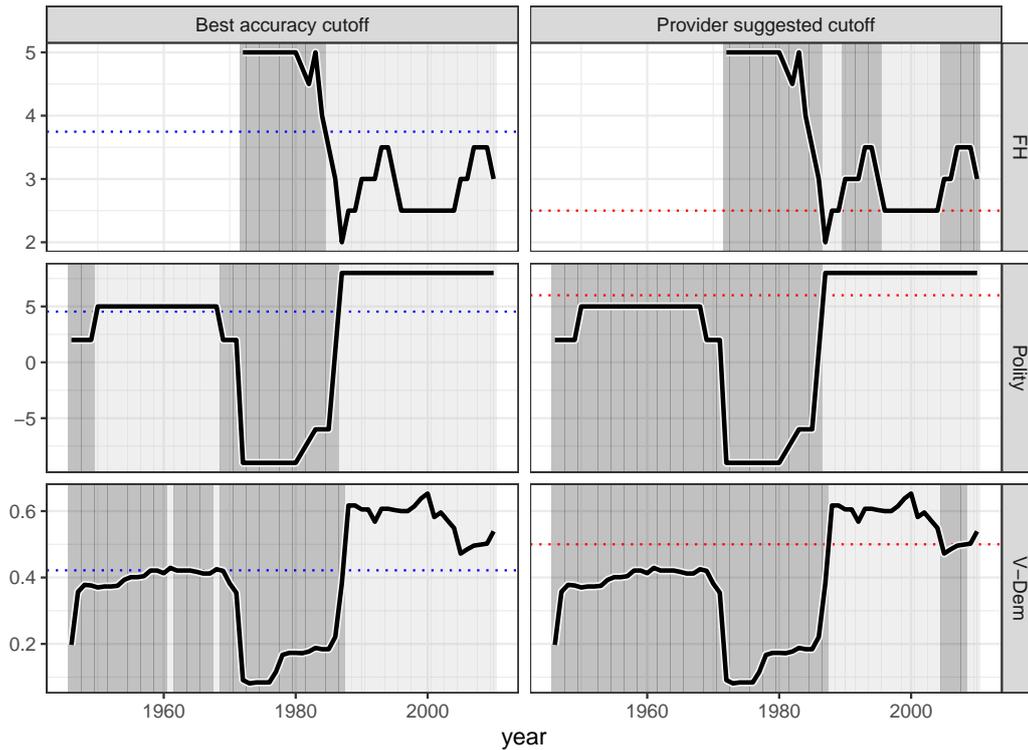


Figure 3: Regime Trajectory of the Philippines

Figure 3 compares the scores of FH, Polity, and V-Dem with two sets of democracy cutoff lines. The horizontal lines in the right column panels are those that the providers suggest, while the lines in the left column panels are those suggested by our analyses. The dark-grey areas indicate autocratic country-years, and the light-gray areas denote democracy according to the given cutoffs. For all three measurements, the provider-suggested cutoffs make the Philippines appear to be authoritarian throughout most of its history, while our suggested cutoffs depict it less so.

Table 4 provides the year-by-year comparison of what Figure 3 illustrated, but adds to our benchmark assessment of regime classification discussed above. The left panel is the classifications according to the provider suggested cutoffs, and the right panel is those according to the best accuracy cutoffs. The grey cells indicate that the year when the regime classification deviates from that of the benchmark's. In this regard, the less the areas with grey color, the more the regime classification is in line with the benchmark assessment. In the table, one can discern that there are greater grey areas in the left panel (provider suggested cutoffs) than the right panel (our suggested cutoffs). This indicates that our suggested cutoffs perform better than the provider-suggested cutoffs in understanding the Philippines' regime trajectory.

The details of the mismatch can be analyzed as follows (readers not interested in the country

details may skip to the conclusion). In the case of FH, the provider's cutoff (2.5) classifies 1986, from 1990 to 1995, and from 2005 to 2018 as autocratic although most country experts consider these years as democratic. Based on what happened during these years in the Philippines, it is difficult to argue that the country during these years should be regarded as an autocracy. Possibly, the attempts of military coups d'état could constitute grounds for the regime category change. However, the major coup attempts happened in the years 1987, 1990, 2001, 2003, 2006, and 2007. Of these, half of them (1987, 2001, 2003) are not coded as autocracy thus this reasoning would be inconsistent. When our suggested cutoff (3.7) is used, the number of inadequate coding significantly declines. Using our cutoff, only one year (1985) differs from the classification of the benchmark: 1985 should have been coded as autocracy, but the FH score is above the democracy cutoff for that year.

For Polity, using the provider's suggested cutoff of 6, 1946 to 1971 are coded as autocracies. Classifying this period as an autocracy is inaccurate in light of the country experts' evaluation and the criteria provided by BMR. When our suggested cutoff (4.5) is used, the amount of disagreement with the benchmark is reduced substantively, although not completely. Years from 1946 to 1949, and from 1969 to 1971 are short of crossing the democracy threshold, but the rest of the years (1950-1968) are now correctly coded. For the years after 1972, both the provider's suggestion and our cutoffs yield the same coding as the benchmark.

The case of V-Dem is similar to Polity. The provider suggested cutoff (0.5) classifies all years from 1946 to 1971 as an autocracy, although they should have been coded as a democracy in view of most experts. Using our suggested cutoff of 0.42, the disagreement with the benchmark is mitigated to some extent: eight years out of 25 coded years (1958-59, 1961-64, 1968-69) are now in line with the baseline assessment. As for the period after the People Power Revolution, both the provider's and our suggested cutoff classify 1986 and 1987 as autocracies, which differs from the baseline. Using the provider-suggested cutoff also evaluates the years from 2005 to 2008 as autocracy, again differing from the benchmark. One could possibly argue that the state of emergency declared by President Gloria Macapagal-Arroyo in 2006 (lasted from February 24 to March 3), and the d'état attempt in 2007 are good enough reasons for these years to be classified as autocracies. In light of BMR's coding rules and the county experts' views, however, these events are short of downgrading the country as an autocracy. Using our suggested measure, these "miscoding" can be avoided.

To summarize, our application exercise to the Philippines yields the following insights.

First, there are no perfect cut points that can flawlessly draw a line between democracy and autocracy for the continuous measures of democracy. One could possibly tailor a cutoff for a particular country, but such an attempt is incompatible with the current practice of setting a single cutoff value applicable to all cases. Second, our suggested cutoffs perform at least better than the provider suggested cutoffs in view of the country experts' understandings about Philippine regime transition. This exercise at the same time revealed that BMR's coding in some cases deviates from experts' assessment, suggesting a limitation of creating a cross-national database by a small number of coders.²² Third, in line with the confusion matrix we provided above (Table 3), the provider-suggested cutoffs more often code a county-year as an autocracy when it should have been coded as a democracy than the reverse scenario. In other words, the Philippine case study also shows that all three provider- suggested cutoffs have restrictive standards in setting their democracy thresholds.

5 Conclusion

This paper has attempted to find theoretically and empirically grounded democracy cutoffs for three major continuous measures of democracy (FH, Polity, V-Dem). Using dichotomous measures built on a set of conditions based on the theories of democracy, we identified the points that yielded the most accurate matching between these two sets of measures. According to our analyses, the cut-points suggested by the providers of FH, Polity, V-Dem are more restrictive than BMR's classification. This means that to be classified as a democracy in FH, Polity, and V-Dem, a higher set of standards needs to be met. More specifically, when we employ BMR as the main reference, the cut points of 3.7 for FH, 4.5 for Polity, and 0.42 for V-Dem are "better" than the provider-suggested cutoffs in the sense that ours reduce the number of mismatches. It should be noted, however, that our suggested cutoffs by no means perfectly classify the regime types in the world.

In this regard, and although it might be a self-defeating conclusion at the fundamental level, our advice to scholars is this: unless necessary, avoid using continuous measures of democracy as if they are categorical. Analyses in this paper showed that any attempt to draw a line to distinguish regime types almost inevitably leads to some inadequate coding. Our case study of the Philippines demonstrated this point succinctly. Scholars should be at least critically

²²Similar issues are raised in Lührmann, Tannenber and Lindberg (2018).

aware of the limitations and the rather restrictive nature of the existing cut-points of the three measures.

Should scholars choose to use the cutoff points in the continuous measures, there are several issues that need to be addressed in future research. First is the question of contextual specificity (Adcock and Collier, 2001, pp.534-535). Our analyses detected cross-regional and cross-period coding inconsistencies as shown in Appendix B. Put differently, the most accurate cut points in our estimates differ by region and period if we use BMR as the benchmark. This suggests that either the benchmark or continuous measures, or both, have some coding bias depending on the region and/or period.²³ This contradicts the principle of the coding schemes of all measures under study; it is supposedly neutral with regard to region and time. We need further research about where such inconsistency comes from, and about how to solve it.

Second, given this paper's findings, the regime sub-types in the continuous measures need re-examination. For example, the providers of Polity labeled the intermediate category of "anocracy" (5 to -5) and "autocracy" (-6 to -10).²⁴ The V-Dem scholars created quadruple sub-types called "liberal democracy," "electoral democracy," "electoral autocracy," and "closed autocracy" (Lührmann, Tannenber and Lindberg, 2018). As this paper showed, more accurate democracy cutoffs lie somewhere other than what the providers claim, the sub-regime cutoffs, which are based on the existing regime cutoffs, and would require reconsideration. Despite more than half-century of efforts to measure political regimes, there is still a lot of research to be done.

²³With regard to FH on this point, see Giannone (2010).

²⁴<http://www.systemicpeace.org/polityproject.html>, accessed on August 22, 2019.

Appendix A Supplementary Analyses using CGV and GWF

A.1 The Best Matching Points

We provide the analyses using CGV and GWF to estimate the best matching point between these two and the continuous measures of democracy. One might think averaging the three measures (BMR, CGV, and GWF) is a better solution than providing a separate analysis. We did not choose this avenue because the definition of democracy as well as the coding rules among the three are slightly different as discussed in Table 1.

	Cut Point at Best Accuracy	Accuracy(%)	Provider's Cut Point	Accuracy(%)	N. of Obs
CGV					
FH	3.7	90.2	2.5	86.2	5,724
Polity	1.4	91.9	6	90.1	7,806
V-Dem	0.38	90.8	0.5	89.1	8,438
GWF					
FH	3.7	90.1	2.5	87.1	5,118
Polity	4.5	93.2	6	92.4	7,385
V-Dem	0.38	92	0.5	90.3	7,673

Table A1: Cut Point Estimates using CGV and GWF

A.2 Comparison of cutoffs

Figure A1 illustrates the cut points as suggested by the measurement providers (dashed line) and as suggested by our analyses (dotted line).

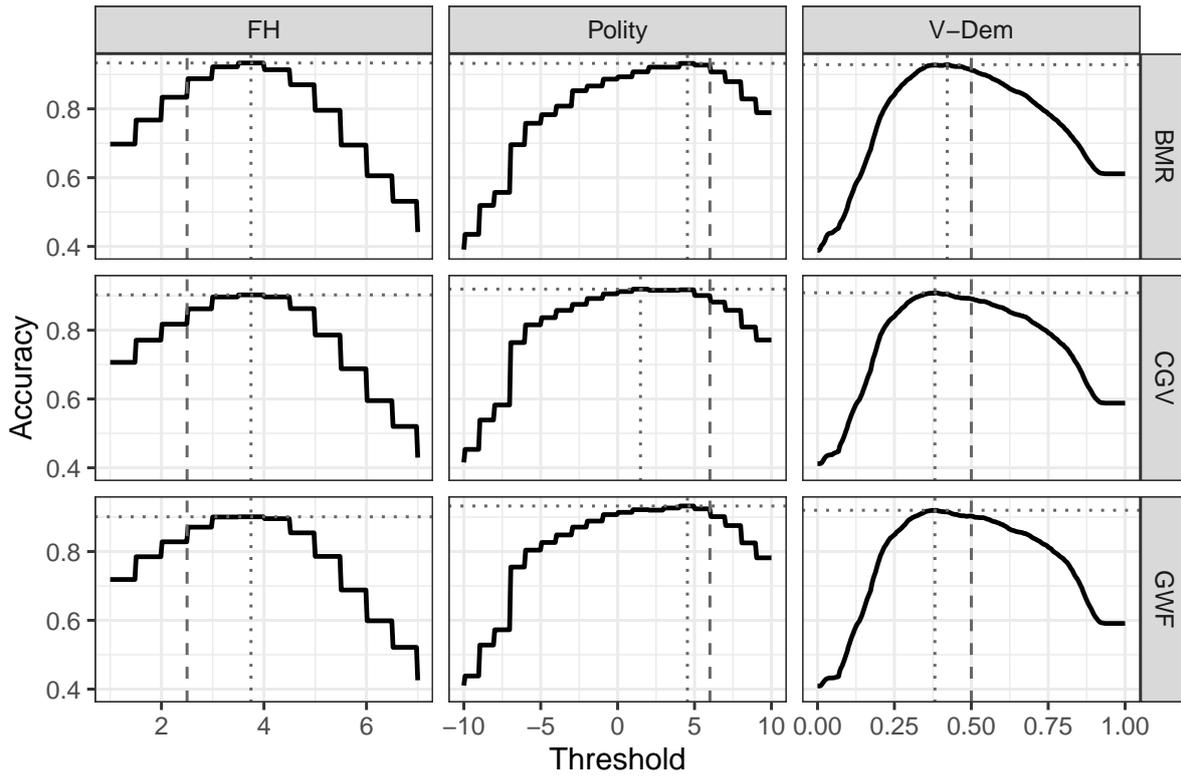


Figure A1: Cutoffs using three dichotomous measures

A.3 Confusion Matrices

The confusion matrices below have the same structure as Table3. The results are also identical to what we described regarding the table.

	Autocracy (above cutoff)	Democracy (below cutoff)
Autocracy (CGV=0)	96.6% , 90.0%	4.4%, 10.0%
Democracy (CGV=1)	26.2%, 9.1%	73.8%, 90.1%

FH. Provider suggested cutoff: 2.5, the cutoff that achieves the highest accuracy: 3.7.

	Autocracy (below cutoff)	Democracy (above cutoff)
Autocracy (CGV=0)	94.2% , 88.5%	5.8%, 11.5%
Democracy (CGV=1)	15.7%, 3.3%	84.3%, 96.7%

Polity. Provider suggested cutoff: 6, the cutoff that achieves the highest accuracy: 1.49.

	Autocracy (below cutoff)	Democracy (above cutoff)
Autocracy (CGV=0)	95.1% , 90.8%	4.9%, 9.2%
Democracy (CGV=1)	19.4%, 9.0%	80.1%, 91.0%

V-DEM Provider suggested cutoff: 0.5, the cutoff that achieves the highest accuracy: 0.4.

Table A2: Confusion matrices with CGV as reference

Note: Cross-tabulation of the CGV's regime classification and the derived classification, under the provider suggested threshold (left), and the threshold that achieves the highest accuracy (right). Each number represents the fraction within the CGV classification; *i.e.*, the number of cases divided by the total cases by rows. The larger numbers in the top-left and bottom-right cells are shown in bold letters.

	Autocracy (above cutoff)	Democracy (below cutoff)
Autocracy (GWF=0)	96.3% , 89.9%	3.7%, 10.1%
Democracy (GWF=1)	25.3%, 9.6%	74.7%, 90.4%

FH. The provider-suggested cutoff: 2.5; the cutoff that achieves the highest accuracy: 3.7.

	Autocracy (below cutoff)	Democracy (above cutoff)
Autocracy (GWF=0)	95.4% , 93.3%	4.6%, 6.7%
Democracy (GWF=1)	11.9%, 6.9%	88.1%, 93.1%

Polity. The provider-suggested cutoff: 6; the cutoff that achieves the highest accuracy: 4.5.

	Autocracy (below cutoff)	Democracy (above cutoff)
Autocracy (GWF=0)	96.1% , 91.7%	3.9%, 8.3%
Democracy (GWF=1)	18.0%, 7.6%	82.0%, 92.4%

V-DEM The provider-suggested cutoff: 0.5; the cutoff that achieves the highest accuracy: 0.38.

Table A3: Confusion matrices with GWF as reference

Note: Cross-tabulation of the GWF's regime classification and the derived classification, under the provider suggested threshold (left), and the threshold that achieves the highest accuracy (right). Each number represents the fraction within the GWF classification; *i.e.*, the number of cases divided by the total cases by rows. The larger numbers in the top-left and bottom-right cells are shown in bold letters.

Appendix B Inter-Temporal and Inter-Regional Variation in Cut Points

Figure A2 shows the period-separated democracy cut points while Figure A3 displays the region-separated ones. In both figures, the magnitude of the gaps between the provider suggested and our suggested thresholds varies depending on the era and the region. This, in turn, suggests the context-specific nature of regime coding either on the side of BMR, or that of the continuous measures, or both. In the case of FH, the variation is only for the region because it only has one period in our periodization scheme.

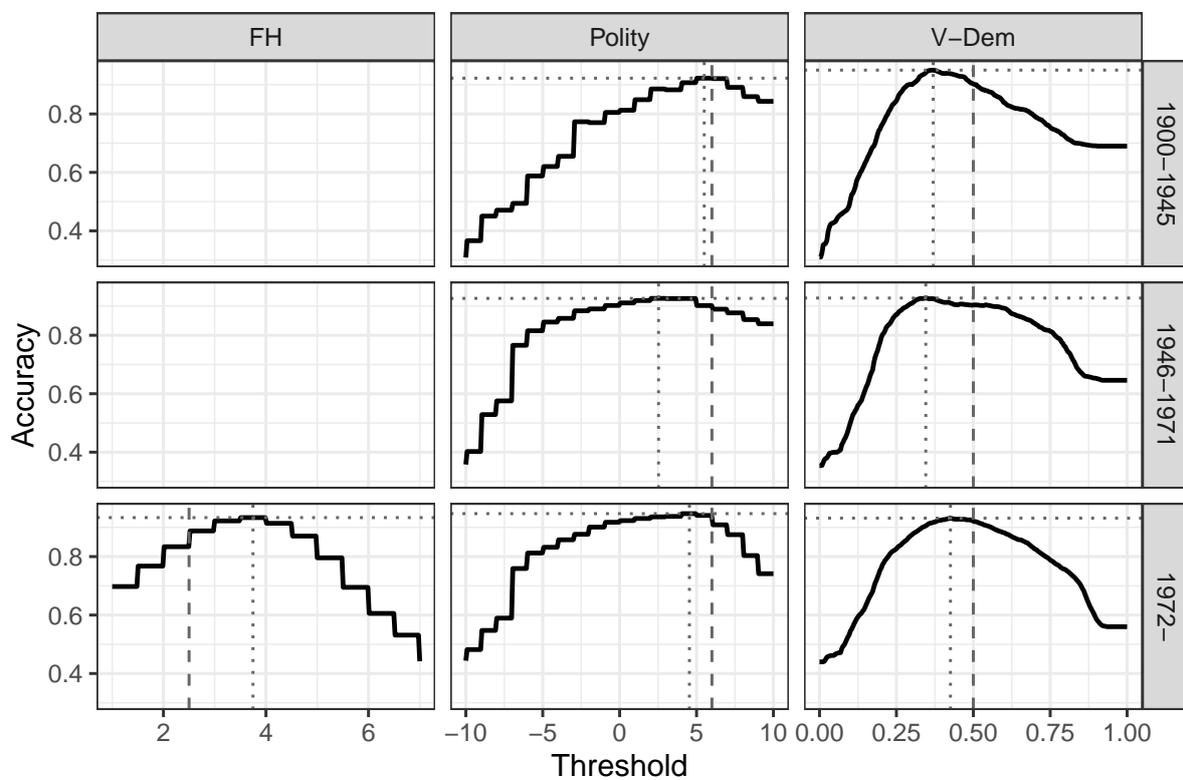


Figure A2: Optimal cutoffs by period using BMR as reference

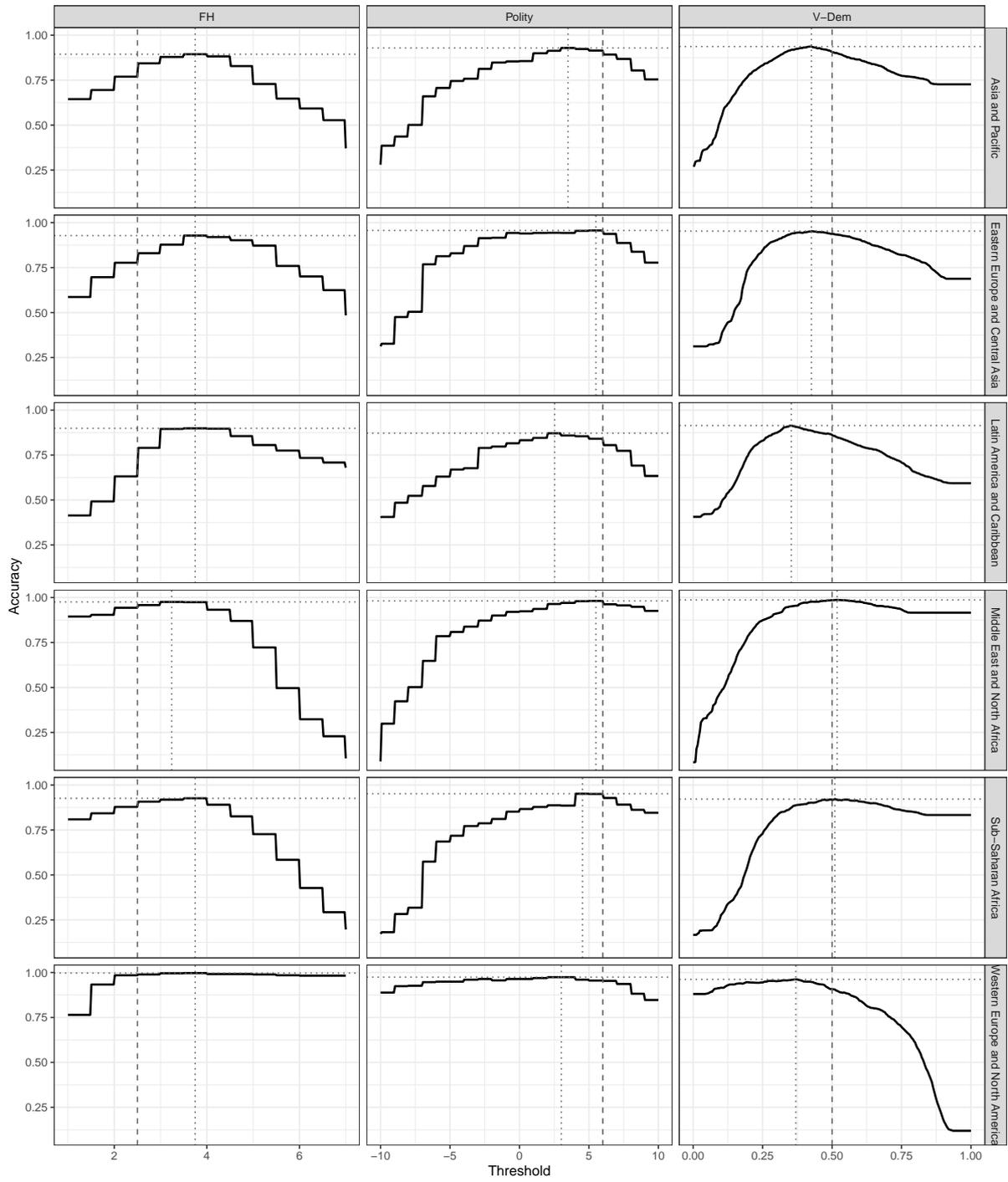


Figure A3: Optimal cutoffs by region using BMR as reference

References

- Adcock, Robert and David Collier. 2001. "Measurement validity: A shared standard for qualitative and quantitative research." *American political science review* 95(3):529–546.
- Alvarez, Mike, José Antonio Cheibub, Fernando Limongi and Adam Przeworski. 1996. "Classifying political regimes." *Studies in Comparative International Development* 31(2):3–36.
- Boese, Vanessa A. 2019. "How (not) to measure democracy." *International Area Studies Review* 22(2):95–127.
- Bogaards, Matthijs. 2010. "Measures of democratization: From degree to type to war." *Political research quarterly* 63(2):475–488.
- Bogaards, Matthijs. 2012. "Where to draw the line? From degree to dichotomy in measures of democracy." *Democratization* 19(4):690–712.
- Boix, Carles, Michael Miller and Sebastian Rosato. 2013. "A complete data set of political regimes, 1800–2007." *Comparative Political Studies* 46(12):1523–1554.
- Bollen, Kenneth A. 1980. "Issues in the comparative measurement of political democracy." *American Sociological Review* pp. 370–390.
- Bush, Sarah Sunn. 2017. "The politics of rating freedom: Ideological affinity, private authority, and the Freedom in the World ratings." *Perspectives on Politics* 15(3):711–731.
- Casper, Gretchen and Claudiu Tufis. 2003. "Correlation versus interchangeability: The limited robustness of empirical findings on democracy using highly correlated data sets." *Political Analysis* 11(2):196–203.
- Cheibub, José Antonio, Jennifer Gandhi and James Raymond Vreeland. 2010. "Democracy and dictatorship revisited." *Public Choice* 143(1):67–101.
URL: <https://doi.org/10.1007/s11127-009-9491-2>
- Clark, William Roberts, Matt Golder and Sona Nadenichek Golder. 2017. *Principles of comparative politics*. CQ Press.
- Collier, David and Robert Adcock. 1999. "Democracy and dichotomies: A pragmatic approach to choices about concepts." *Annual review of political science* 2(1):537–565.

- Coppedge, Michael, John Gerring Carl Henrik Knutsen Staffan I. Lindberg Jan Teorell David Altman Michael Bernhard M. Steven Fish Adam Glynn Allen Hicken Anna Lührmann Kyle L. Marquardt Kelly McMann Pamela Paxton Daniel Pemstein Brigitte Seim Rachel Sigman Svend-Erik Skaaning Jeffrey Staton Agnes Cornell Lisa Gastaldi Haakon Gjerlow Valeriya Mechkova Johannes von Römer Aksel Sunöm Eitan Tzelgov Luca Uberti Yi-ting Wang Tore Wig and Daniel Ziblatt. 2019. “V-Dem Codebook v9.” Varieties of Democracy (V-Dem) Project.
- Coppedge, Michael, John Gerring, David Altman, Michael Bernhard, Steven Fish, Allen Hicken, Matthew Kroenig, Staffan I Lindberg, Kelly McMann, Pamela Paxton et al. 2011. “Conceptualizing and measuring democracy: A new approach.” *Perspectives on Politics* 9(2):247–267.
- Coppedge, Michael, John Gerring, Staffan I Lindberg, Svend-Erik Skaaning and Jan Teorell. 2017. “V-Dem comparisons and contrasts with other measurement projects.” *V-Dem Working Paper* 45.
- Cutright, Phillips. 1965. “Political Structure, Economic Development, and National Social Security Programs.” *American Journal of Sociology* 70(5):537–550.
URL: <https://doi.org/10.1086/223930>
- Dahl, Robert A. 1956. *A preface to democratic theory*. Vol. 115 University of Chicago Press.
- Dahl, Robert Alan. 1973. *Polyarchy: Participation and opposition*. Yale University Press.
- Elkins, Zachary. 2000. “Gradations of democracy? Empirical tests of alternative conceptualizations.” *American Journal of Political Science* .
- Gastil, Raymond Duncan. 1985. “The past, present and future of democracy.” *Journal of International Affairs* pp. 161–179.
- Gastil, Raymond Duncan. 1990. “The comparative survey of freedom: Experiences and suggestions.” *Studies in Comparative International Development* 25(1):25–50.
- Geddes, Barbara, Joseph Wright and Erica Frantz. 2014a. “Autocratic Breakdown and Regime Transitions: A New Data Set.” *Perspectives on Politics* 12(2):313–331.

- Geddes, Barbara, Joseph Wright and Erica Frantz. 2014b. "Autocratic Regimes Code Book, Version 1.2."
- Giannone, Diego. 2010. "Political and ideological aspects in the measurement of democracy: the Freedom House case." *Democratization* 17(1):68–97.
- Gleditsch, Kristian S and Michael D Ward. 1997. "Double take: A reexamination of democracy and autocracy in modern polities." *Journal of Conflict Resolution* 41(3):361–383.
- Goertz, Gary. 2006. *Social science concepts: A user's guide*. Princeton University Press.
- Gurr, Ted Robert. 1974. "Persistence and change in political systems, 1800–1971." *American political science review* 68(4):1482–1504.
- Gurr, Ted Robert, Keith Jagers and Will H. Moore. 1990. "The Transformation of the Western State: The Growth of Democracy, Autocracy, and State Power Since 1800." *Studies in Comparative International Development* 25(1):73–108.
- Huntington, Samuel P. 1991. "Democracy's third wave." *Journal of democracy* 2(2):12–34.
- Jagers, Keith and Ted Robert Gurr. 1995. "Tracking democracy's third wave with the Polity III data." *Journal of peace research* 32(4):469–482.
- Lijphart, Arend. 2012. *Patterns of democracy: Government forms and performance in thirty-six countries*. Yale University Press.
- Linz, Juan José. 1975. *Totalitarian and authoritarian regimes*. Lynne Rienner Publishers.
- Lipset, Seymour Martin. 1959. "Some social requisites of democracy: Economic development and political legitimacy." *American political science review* 53(1):69–105.
- Lührmann, Anna, Marcus Tannenberg and Staffan I Lindberg. 2018. "Regimes of the World (RoW): Opening New Avenues for the Comparative Study of Political Regimes." *Politics & Governance* 6(1).
- Lührmann, Anna and Staffan I Lindberg. 2019. "A third wave of autocratization is here: what is new about it?" *Democratization* pp. 1–19.

- Marshall, Monty G and Ted Robert Gurr. 2016. "Polity IV project: Political regime characteristics and transitions, 1800-2015." <http://www.systemicpeace.org/polity/polity4.htm>. Retrieved on August 1, 2016.
- Marshall, Monty G, Ted Robert Gurr, Christian Davenport and Keith Jagers. 2002. "Polity IV, 1800-1999: Comments on Munck and Verkuilen." *Comparative Political Studies* 35(1):40–45.
- Munck, Gerardo L and Jay Verkuilen. 2002. "Conceptualizing and measuring democracy: Evaluating alternative indices." *Comparative political studies* 35(1):5–34.
- Pemstein, Daniel, Stephen A. Meserve and James Melton. 2010. "Democratic compromise: A latent variable analysis of ten measures of regime type." *Political Analysis* 18(4):426–449.
- Przeworski, Adam, R Michael Alvarez, Michael E Alvarez, Jose Antonio Cheibub and Fernando Limongi. 2000. *Democracy and development: Political institutions and well-being in the world, 1950-1990*. Vol. 3 Cambridge University Press.
- Sartori, Giovanni. 1987. *The theory of democracy revisited*. Oxford University Press.
- Skaaning, Svend-Erik, John Gerring and Henrikas Bartusevičius. 2015. "A lexical index of electoral democracy." *Comparative Political Studies* 48(12):1491–1525.
- Starr, Harvey and Christina Lindborg. 2003. "Democratic dominoes revisited: The hazards of governmental transitions, 1974-1996." *Journal of Conflict Resolution* 47(4):490–519.
- Teorell, Jan, Michael Coppedge, Staffan Lindberg and Svend-Erik Skaaning. 2019. "Measuring polyarchy across the globe, 1900–2017." *Studies in Comparative International Development* 54(1):71–95.
- Vanhanen, Tatu. 2000. "A new dataset for measuring democracy, 1810-1998." *Journal of Peace Research* 37(2):251–265.