Re-examining thresholds of continuous democracy measures

Yuko Kasuya & Kota Mori

Published online: 10 Nov 2021.

Submit your article to this journal ⬚

View related articles ⬚

View Crossmark data ⬚

Routledge
Taylor & Francis Group

Check for updates

Re-examining thresholds of continuous democracy measures

Yuko Kasuya [a,b] and Kota Mori[b]

[a]Department of Political Science, Keio University, Tokyo, Japan; [b]V-Dem East Asia Centre, Tokyo, Japan

ABSTRACT

Scholars frequently dichotomize continuous measures of democracy by setting a regime cut-off. However, such cut-offs often lack theoretical or empirical justifications, making the resulting classifications difficult to interpret conceptually. We investigate this challenge involving three major continuous democracy measures: the Freedom House score (FH), the Polity score, and the Regime of the World (RoW) that is based on the V-Dem's Electoral Democracy Index (EDI). We develop a framework to empirically derive thresholds using categorical democracy measures as benchmarks. Our analyses find that the cut-offs that yield the highest consistency with the classifications of BMR, CGV, and GWF are 3.5 for FH, 5 for Polity and 0.39 for EDI/RoW. These levels are lower than the conventional cut-offs, implying less demanding democratic standards. Consequently, the conventional cut-offs (2.5 for FH, 6 for Polity and 0.5 for EDI/RoW) endeavour to reflect more stringent standards of democracy than what these dichotomous measures employ.

Introduction

Scholars have developed two major types of measurement methods of democracy. One is the dichotomous measure of democracy, reflecting the theoretical understanding that a country can be either a democracy or an autocracy. The other type employs the notion that democracy is a continuum and measures it as such. When the latter type is used in a study that requires regime type classification, scholars often convert the data into dichotomous variables by using a predetermined cut-off point. For example, the Varieties of Democracy (V-Dem) Institute's annual Democracy Report provides the global number of democracies and autocracies based on a threshold imposed on their continuous democracy measure (Alizada et al., 2021). Similarly, the Freedom House's Freedom in the World annual report provides the count of 'free' countries, which are usually interpreted as an equivalent of democracies, based on a cut-off set on its continuous measure (Repucci & Slipowitz, 2021).

Due to the complex nature of continuous measures, however, choosing an appropriate threshold is not a trivial task. In fact, despite decades of effort in conceptualising and measuring democracy, the issue of the democracy threshold has only received minor attention. Although there are some thresholds that are conventionally applied, as we

---

CONTACT Yuko Kasuya ykasuya@keio.jp; ykasuya@law.keio.ac.jp Department of Political Science, Keio University, 2-15-45, Mita, Minato, Tokyo 108-8345 Japan.

discuss below, they often lack theoretical justification, and some are even seen as arbitrary choices. This is a critical oversight because analyses using these cut-offs risk improper interpretation of the research findings.

This article aims to fill this gap in the literature by re-examining the regime cut-offs for three major continuous measures: the Freedom House's global freedom score (hereafter FH), the Polity IV Project's Polity score (hereafter Polity) and the V-Dem Project's Electoral Democracy Index (EDI).[1] For FH and Polity, the providers offer a set of suggested regime cut-offs in their official documents. Regarding the EDI, a 2018 article by Lührmann et al. (2018) introduced a regime classification called the 'Regime of the World' (hereafter RoW), which is based on the EDI. Although this scheme is not officially endorsed by the V-Dem Project, the RoW has been used in many studies, including the V-Dem Project's official publications.[2]

Our approach to the threshold assessment uses dichotomous measures of democracy as benchmarks. Many dichotomous measures, especially the ones we use, document their coding criteria explicitly by defining the set of conditions a country must fulfil to be classified as a democracy. As a result, the empirical thresholds that we derive so as to best approximate the benchmarks can be interpreted as employing a similar level of requirement as the benchmarks. We first employ three dichotomous democracy measures with similar minimalist democracy conception created by the following groups as benchmarks: (1) Boix, Miller and Rosato (hereafter BMR, Boix et al., 2013); (2) Cheibub, Gandhi and Vreeland (hereafter CGV, Cheibub et al., 2010); and (3) Geddes, Wright, and Franz (hereafter GWF, Geddes et al., 2014a).[3] In our analyses, the cut-off points that achieve the highest consistency with the binary measure coding are 3.5 for FH, 5 for Polity and 0.39 for EDI. These register lower standards for democracy than the conventional thresholds (2.5, 6 and 0.5, respectively). This result implies that in classifying a country as a democracy, the conventional cut-offs of FH, Polity and RoW require a more stringent set of conditions than those prescribed in the dichotomous measures. We further conduct analogous analyses using different benchmark variables: BMR's democracy coding including female suffrage and the eight-fold regime classification from Skaaning et al. (2015)'s Lexical Index of Electoral Democracy (LIED).[4] Our results confirm that the conventional cut-off levels of FH, Polity and EDI represent a higher standard than the procedural minimum notion of democracy.

This study's contribution is both theoretical and practical. Theoretically, we address the issue of measurement validity concerning regime cut-offs. Currently, the conventional thresholds of three continuous democracy measures face the problematic situation in which discussion of the concept to be measured is almost absent while the measurements already exist. Our study mitigates this situation by illuminating the nature of democracy these conventional cut-offs denote. Practically, this study provides empirical researchers with guidance to choose reliable and interpretable thresholds that fit particular research contexts. This is particularly important for the study areas where the use of dichotomous classifications is pervasive or even necessary. For example, scholars addressing the recent debate on democratic backsliding often count the number of democracies and non-democracies or that of backsliders and compare them across time and regions of the world (e.g. Alizada et al., 2021). Properly chosen thresholds would enhance our understanding of the state of the world and strengthen statistical findings. In addition, since our method is designed to be generic, it is not limited to a specific set of indices; one may

apply the method we introduce in this study to different variables and benchmarks with little modification.

Our analyses proceed in four steps. First, we review the related literature and demonstrate that the three continuous measures' cut-offs conventionally used in practice lack a clearly defined concept that they are supposed to measure. Second, we discuss the properties of three dichotomous measures that we use as the first set of benchmarks, namely, BMR, CGV, and GWF. Third, using these measures, we empirically derive the cut-off point for each continuous measure that best approximates the dichotomous measures' regime classification. Fourth, we conduct similar optimisation analyses with two alternative sets of benchmark measures: BMR's extension variable that includes female suffrage condition, and Skaaning et al. (2015)'s LIED. In the conclusion section, we summarise the implications of this study.

Using continuous measures as though they are categorical

Depending on how one conceptualises democracy, there are two ways of measuring this 'essentially contested concept' (Collier et al., 2006). Some scholars maintain that countries are either in a state of democracy or non-democracy, meaning that there is no grey zone (e.g. Linz, 1975; Sartori, 1987; Huntington, 1991; Alvarez et al., 1996). Building on this concept of democracy, a number of databases have been developed that code political regimes as a dichotomous variable (Przeworski et al., 2000; Boix et al., 2013; Geddes et al., 2014a). However, other scholars consider the conceptual structure of democracy a continuum. For example, Dahl (1973, p. 241) states that 'every country is located somewhere along the underlying continuum' of being democratic. The continuous measures of democracy reflect this view of democracy (Bollen, 1980; Gastil, 1985; Gastil, 1990; Gurr et al., 1990; Vanhanen, 2000; Coppedge et al., 2011).

A common practice in the literature is to convert the variables originally measured on a continuous scale into dichotomous ones for the purpose of statistical analysis. In our review of ten major journals in comparative politics and international relations published between 2000–2019, 149 articles employed Polity as a discrete variable while nine articles used FH.[5] Indeed, it has become almost an industry standard to use the cut-offs suggested by the providers – namely, 2.5 for FH and 6 for Polity – to the extent that they appear in an influential textbook of comparative politics (Clark et al., 2017, pp. 157–160). The EDI-based RoW was only introduced in 2018. Along with the recent popularity of the V-Dem data in general, many regime scholars have adopted the RoW, and the number of citations has, thus, quickly increased.[6]

Some scholars have criticised that FH and Polity users have picked different cut-off points than what providers have endorsed. For example, Bogaards (2012) found 14 different ways to set FH's cut-off points other than the provider-suggested value and 18 different ways for Polity. Other scholars have also found that studies can obtain different results depending on which cut-off points they use, thus impairing the accumulation of findings (Elkins, 2000; Casper & Tufis, 2003; Bogaards, 2010; Cheibub et al., 2010).

However, few scholars have critically questioned why those cut-off points were set at their specific initial positions. In our review of the related literature, we found that several scholars raised this question in passing. For example, Lueders and Lust (2017, p. 11) noted that 'the regime thresholds created based on Polity IV or Freedom House are arbitrary'.

The *Comparative Politics* textbook by Clark et al. (2017, p. 157, fn.3) mentioned in a foot-note that 'the decision of where to place the cut-off points is rarely, if ever, justified in a theoretical manner'. However, we found little mention of this point in some of the most fully fledged critiques of democracy measures (e.g. Munck & Verkuilen, 2002; Goertz, 2006). Arguably, this is a crucial oversight in the studies of political regimes because – as discussed below – the conventional cut-offs lack theoretical justification for what they are measuring conceptually.

In our review of the codebooks and related documents, none of these continuous measures of democracy provide a theoretical discussion of where they set their cut-off points. In FH, countries are given the status of 'Free' (1.0–2.5), 'Partly Free' (3.0–5.0) or 'Not Free' (5.5–7.0).[7] However, we could not find any clear conceptual explanation of why these values were assigned to each category in the Freedom House's methodology document. Although the Freedom House does not explicitly state that their Freedom in the World score is a measure of democracy, scholars often equate democracy with the 'Free' status (Clark et al., 2017, p. 160).[8]

The Polity website states that country-years with a Polity score of ≥6 are 'democracies', but it provides no rationale for setting this cut-off point.[9] Meanwhile, the terms they intro-duce are used inconsistently. In the codebook's latest version, a Polity score of ≥7 is called 'full democracy', whereas their website labels a score of ten as a 'full democracy' (Marshall et al., 2002, p. 35). We could not find any explanation as to why and how they differ con-ceptually among our consulted sources.[10]

For the RoW, there are three necessary conditions for a country to be classified as a democracy: (1) *de facto* multiparty elections, (2) free and fair elections and (3) a score of >0.5 on the EDI. The first two conditions are coded on an ordinal scale,[11] and the third relies on a cut-off point set for an interval scale (Lührmann et al., 2018). In this study, we are specifically concerned with the 0.5 cut-off, which determines 99.9% of the RoW's coding from 1900 to 2019.[12] Accordingly, the 'RoW cut-off' refers to the 0.5 threshold, and this differs from the 'RoW classification', which denotes the classification scheme that involves all three conditions mentioned above. The EDI is one of the most abstract-level indices within V-Dem indices, and it aggregates about 40 basic variables (Coppedge et al., 2011, p. 39). We find little theoretical justification for why 0.5 should be the threshold, and – unlike the other two conditions – it is difficult to interpret what 0.5 signifies given the EDI's highly abstract nature. The proponents of this cut-off also write that 0.5 is 'admittedly arbitrary' (Lührmann et al., 2018).

In summary, the three major continuous democracy measures' conventional cut-off points lack theoretical grounds to answer why these values are set at that level. This raises the issue of measurement validity; we are uncertain to what extent the suggested cut-off points appropriately measure the boundary of democracy. In reference to Adcock and Collier's concept measurement scheme, we face a problematic situation where dis-cussion of the concept to be measured is almost absent while the measurements already exist (Adcock & Collier, 2001). This situation can result in a conceptual discrepancy between the researchers' intended analyses and the regime classification implied by the cut-offs. Further, there is a danger of falling into a circular explanation when classifying a given country-year as a democracy *because* it passed a certain threshold without under-standing what the threshold means.

Against this backdrop, one can avoid using the democracy cut-off points but employ these continuous measures as-is. Alternatively, Bogaards suggested choosing the Polity components that presumably measure the phenomenon of interest and using them either as-is or combining them (Bogaards, 2012). Also, Skaaning et al. (2015) developed an eight-point electoral democracy scale that allows users to interpret the meaning of each category. In contrast, we tried to preserve the idea of using cut-off points for continuous democracy measures by providing reference points for conceptual evaluation.

Binary measures as a benchmark

In this study, we empirically derive thresholds of continuous democracy measures by using dichotomous democracy measures as benchmarks. While continuous measures are useful for cross-national and/or inter-temporal comparison (e.g. country A is more democratic than country B, or country C has become more democratic in the last decade), their numeric values tend to be difficult to interpret. This is because continuous measures are often computed from a number of variables representing various properties, often by a complex mathematical formula, and thus, an identical value may result from different variable combinations. As a result, finding appropriate thresholds through a theoretical approach (e.g. based on the mathematical properties of the estimation schemes) becomes a nontrivial task. Dichotomous measures, in contrast, have an advantage in interpretability because the coding rules usually define the criteria for a country-year to be classified as a democracy explicitly. Our approach is to utilise dichotomous measures to enhance the interpretability of continuous values and help to derive appropriate thresholds empirically.

In the next section, we present the outcomes when using BMR, CGV and GWF as benchmark measures. These three are chosen mainly for two reasons. First, all of them have documentation that stipulates the set of conditions for a country-year to be classified as a democracy in their coding scheme. This is of crucial importance because without this property, the derived thresholds would also become ambiguous and uninterpretable values. Second, as Table 1 summarises, all three measures employ a similar, if not identical, set of conditions of democracy. This allows us to use them in a combined manner so that we can mitigate potential noise and biases when relying on a single reference variable (Wilson, 2014). In the latest version of BMR released in 2018, a coding of democracy that accounts for female suffrage is added. Because of the similarity with CGV and GWF, however, we employ the original BMR in this section, and we provide analyses using this new variable later.[13]

Table 1. Comparison of binary measures.

	Measurement goals	Democracy conditions	Countries	Years
BMR	Dahl's polyarchy	Elected executive and legislature; adult male suffrage	222	1800–2015
CGV	Minimalist democracy	Elected executive and legislature; multi-party competition; alternation of power	202	1946–2020
GWF	Autocracy type	Leadership group is elected by free and fair election	182	1946–2010

Note: We provide a similar set of information for the three continuous measures under study in Appendix A.

Before proceeding to the statistical analyses, in the remaining part of this section, we review the properties of the three dichotomous measures to demonstrate their appropriateness as benchmarks. BMR, which follows the notion of democracy defined by Dahl, codes a given country-year as a democracy if it meets the following three conditions: (1) the executive is either directly or indirectly elected in popular elections and is responsible either directly to voters or a legislature; (2) the legislature (or the executive if elected directly) is chosen in free and fair elections; and (3) the majority of adult men have the right to vote (Boix et al., 2013).

Intending to measure the minimalist notion of democracy, the creators of the CGV defined democracy as a regime in which 'government officials are filled as a consequence of contested elections' (Cheibub et al., 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 a 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 alteration in power must have taken place under the electoral rules, which were identical to those that brought the incumbent into office.[14]

For the GWF, a straightforward definition of a democratic regime is missing in both the article introducing the database and in the codebook(Geddes et al., 2014b). We speculate that this is because the providers' core concern is classifying various types of autocratic regimes, and democracy is measured as a residual category of autocracy. Nevertheless, their codebook defines an 'undemocratic' regime. By inference, a democratic regime can be defined as a regime consisting of '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 an indirect election by a body of which at least 60 percent were elected in direct, reasonably fair competitive elections; or constitutional succession to a democratically elected executive.' (Geddes et al., 2014b, p. 6).

The above discussion informs us that these three are broadly similar but differ in some of the democracy conditions they employ. For example, CGV, in comparison to the other two, does not include suffrage as a condition, but it does include government turnover to distinguish the regime type. Concerning the extent of suffrage, BMR requires a majority of adult male suffrage, whereas 40% suffice for the GWF's condition.

To check the degree of consistency among the three, we compute the fraction of country-years with the same coding for each pair, as shown in Table 2.[15] While the coding criteria of the three variables are not identical, each pair exhibits a high agreement ratio, around 95%.

In the next section, we employ these three measures as benchmarks for assessing the cut-offs of continuous measures. We expect that this exercise allows us to map the criteria that the dichotomous measures employ – the procedural minimum notion of democracy – onto the scale of the continuous measures.

Table 2. Fraction of country-year pairs in agreement.

	BMR	CGV	GWF
BMR	1.000		
CGV	0.951	1.000	
GWF	0.947	0.952	1.000

Cut-off points that approximate three dichotomous measures

In this section, we present empirical derivation of cut-off points for FH, Polity, EDI/RoW using BMR, CGV, and GWF as dichotomous benchmarks. Methodologically, our approach is an application of the grid-search algorithm: For every possible candidate threshold, we calculate the fraction of agreement between the binary benchmarks and the classification implied. The optimal threshold is the one that achieves the highest agreement ratio.[16]

Figure 1 visually presents the idea behind our approach. In this figure, we plot the histograms of continuous democracy measures and the corresponding density estimates by the classification in the dichotomous benchmarks. In general, we observe that the dark (autocracy in the binary classification) and light (democracy) histograms are polarised to each end, implying a strong correlation between binary and continuous measures. Consistently across the panels, the conventional cut-offs, shown as the vertical dashed lines, are located at more 'democratic' points than the intersection of two density plots.[17] From this visual inspection we can infer that – to be classified as a democracy in FH, Polity and EDI/RoW under the conventional cut-offs – a county needs to fulfil higher democratic standards than the dichotomous measures.

Figure 2 shows the optimisation of the thresholds visually. In this figure, we plot the fraction of agreement against the cut-off point levels for each pair of continuous and dichotomous measures. The top row of the graph exhibits the average scores across three dichotomous measures and the graphs in the second row and below show the results using the dichotomous measures separately.[18] Dotted and dashed lines indicate the cut-offs that achieve the highest consistency scores[19] and the conventional cut-offs, respectively. This figure also indicates that, in all settings, the conventional cut-offs
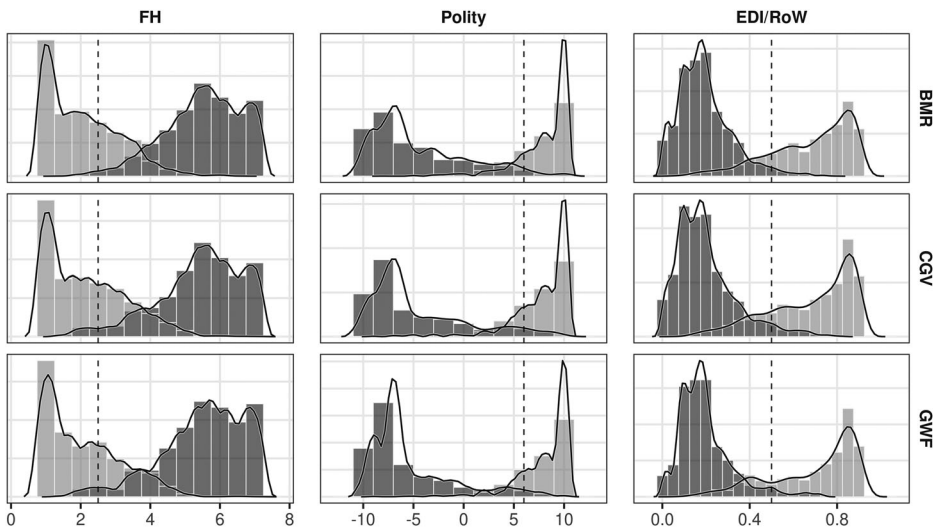

Figure 1. Distribution of continuous democracy measures by binary classification. Note: The histograms of the continuous measures, separately for country-years coded as a democracy (light-grey) and those coded as an autocracy (dark-grey). Solid lines are the corresponding kernel density estimates. The vertical dashed lines indicate the conventional cut-off points.
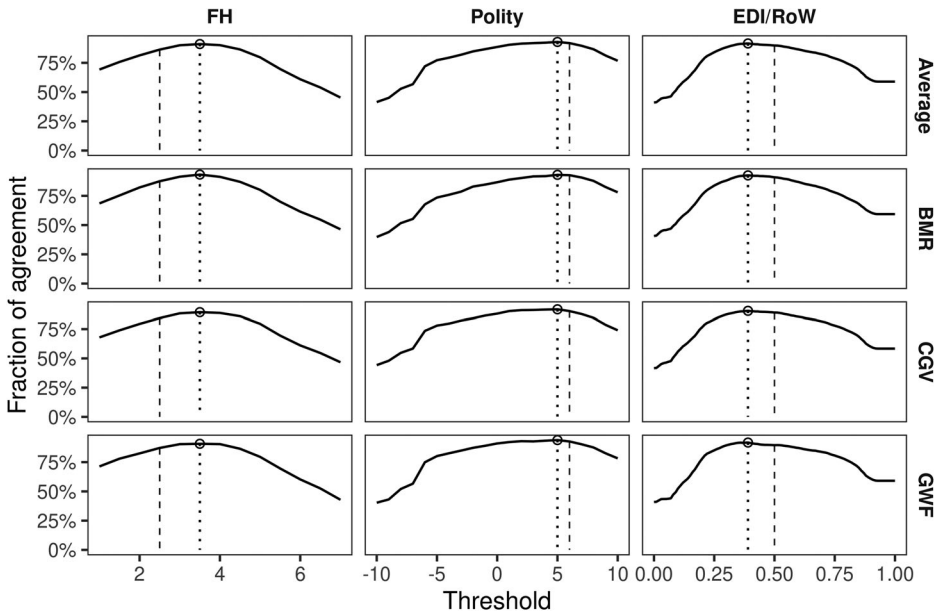
Figure 2. Fraction of agreement with dichotomous benchmarks by threshold levels. Note: The fraction of agreement between dichotomous benchmarks and the classification implied by threshold levels. Columns represent continuous measures and rows correspond to the benchmark variables, with the first row representing the average of the other three. The dotted and dashed vertical lines indicate the conventional and the optimal thresholds.

are located at a more democratic level than the peaks of the curves, thus confirming the same finding as presented in Figure 1.

Table 3 compares the conventional cut-off points with those that achieve the highest average consistency.[20] Although the conventional cut-offs of FH, Polity and EDI/RoW are 2.5, 6 and 0.5, respectively, our analyses indicate that the highest fraction of agreement is achieved at the cut-off points 3.5 for FH, 5 for Polity and 0.39 for EDI/RoW.[21] Hereafter, for the sake of convenience, we will interchangeably call these 'optimal cut-offs'. By examining the scores in the table, one might think that there are only slight differences between the two groups. This is because most of the country-years in the datasets lie in the zones of clearly democratic or clearly autocratic, where threshold adjustments do not alter the classification for most cases. The cut-off scores' differences come mainly from the small fraction of the country-years at the borders of the regime transition that are subject to measurement disagreements (Coppedge et al., 2011, p. 252). In this regard, the optimal

Table 3. Cut-off point estimates using the dichotomous measures as reference.

	Conventional		Optimal		
	Cut-off point	Fraction of Agreement (%)	Cut-off point	Fraction of Agreement (%)	N. of Obs.
FH	2.5	86.3	3.5	91.0	19,078
Polity	6	91.8	5	92.8	28,735
EDI/RoW	0.5	89.8	0.39	91.6	31,015

Note: "Optimal" denotes the threshold level that achieves the highest fraction of agreement between the dichotomous measures' coding and classifications implied by thresholds.

thresholds are mainly driven by these grey-zone cases and are chosen so as to maximise coding agreements with the dichotomous benchmarks.[22]

In summary, all three conventional cut-offs are more stringent than the cut-offs that achieve the highest consistency with the combined dichotomous measures. This suggests that the conventional cut-offs signify higher standards of democracy than what the dichotomous benchmarks prescribe.[23]

Thresholds from benchmarks with alternative democracy conceptions

To further examine the impact of the benchmark choice, we provide a similar threshold calculation in this section, using two alternative benchmark measures. Firstly, we employ an extension of BMR with a female suffrage requirement. This variable has recently been added to their dataset and employs the same set of conditions as the original BMR, except that the suffrage requirement is not limited to adult males.[24] Accordingly, this benchmark requires a higher level of democracy standard.

Figure 3 plots the fraction of agreement against the cut-off levels for the original BMR and BMR with female suffrage requirement (labelled 'BMR-female' for short). Note that the result for BMR is reproduced from Figure 2 for the purpose of comparison. We omit FH from this analysis since for the period when FH is available, BMR and BMR-female take identical values.[25] The optimal thresholds with BMR-female are 5.5 for Polity (5 with the original BMR) and 0.48 for EDI/RoW (0.39 with the original).[26] The thresholds in both cases have moved to the direction of higher levels of democracy, due to the more stringent nature of democracy coding of this benchmark variable. In fact, when using BMR-female as the benchmark, the optimal thresholds become close to the conventional ones both for Polity and EDI/RoW.

Secondly, we employ the LIED developed by Skaaning et al. (2015). The authors of the LIED evaluate country-years on the properties related to electoral democracy, such as the existence of election and competitiveness. Based on the fulfilment of these properties, they categorise country-years into eight-fold regime classification from non-electoral autocracy (least democratic) to polyarchy (most democratic). These LIED classes, although
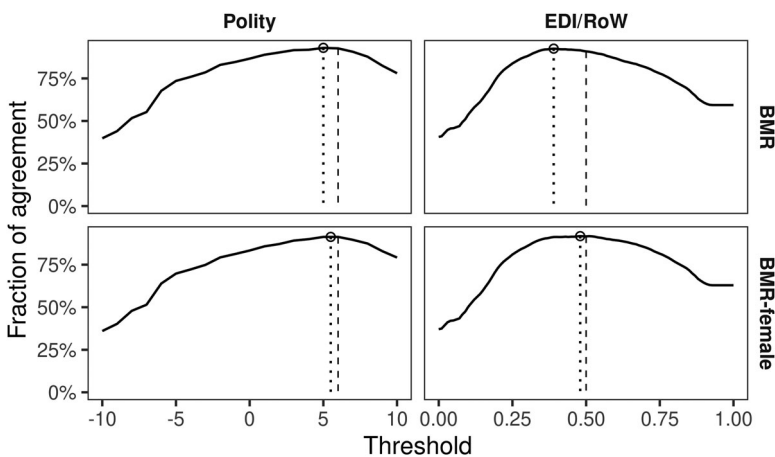

Figure 3. Comparison of the threshold optimisation using BMR and BMR-female as benchmarks.

Table 4. Optimal threshold for the LIED classifications.

	LIED ver.6.0	FH	Polity	EDI/RoW
L0	No election (non-electoral autocracies)	7	-10	0
L1	No-party or one-party elections (one-party autocracies without elected executive)	7	-9	0.10
L2	Multiparty elections for L (multiparty autocracies without elected executive)	4.5	-4	0.21
L3	Multiparty elections for L and E (multiparty autocracies)	4.5	-3	0.25
L4	Minimally competitive, multiparty elections for L & E (exclusive democracies)	3.5	4	0.39
L5	L4 with full male or female suffrage (male democracies)	3.5	6	0.42
L6	L4 with universal suffrage (electoral democracies)	3.5	6	0.43
L7	L4 with universal suffrage and political liberties (polyarchies)	2.5	8	0.66

Note: In the second column, L stands for legislative elections, E stands for executive elections. The regime names in the parenthesis are found in the Codebook for the version 6. The dotted line denotes the distinction between autocracies and democracies according to the terms used by Skaaning et al. (2015).

they are ordered, can be seen as eight distinct dichotomous measures that represent different levels of requirements. Hence, we can apply the same methodology to them to calculate the optimal threshold for each.

Table 4 summarises the results.[27] Each line shows the optimal thresholds for the corresponding LIED class (we call them L0 through L7 to follow the LIED authors' terms). Note that L0, non-electoral autocracies, is the least democratic class among all, and thus, is assigned the smallest possible values as the threshold. For all three continuous measures, we observe a large increase in the thresholds from L1 to L2. This might be because these measures emphasise the existence of multiparty elections when assigning the values. The three dichotomous measures we used as benchmarks in the previous section would be the closest to the L4 (exclusive democracy) in this scale. As expected, the derived thresholds for L4 mostly agree with the ones obtained in the previous section. The conventional thresholds (2.5 for FH, 6 for Polity and 0.5 for EDI/RoW) tend to be located somewhere between L6 and L7 classes. This confirms the finding that they represent a higher standard than the procedural requirement of democracy.

Conclusion

In this paper, we have examined the ambiguity behind the regime classification cut-offs of three major continuous measures of democracy (FH, Polity and EDI/RoW). We developed an analytic scheme to empirically derive optimal cut-offs, using categorical democracy indices as benchmarks. Our analysis employing three dichotomous measures (BMR, CGV, and GWF) as benchmarks revealed that the thresholds that best approximate these benchmarks are lower (less democratic) than the cut-offs conventionally used in the literature: While a common choice of thresholds are 2.5 for FH, 6 for Polity, and 0.5 for EDI/RoW, the optimal ones (i.e. the most consistent with the benchmarks) are 3.5, 5, and 0.39, respectively. Given that these benchmarks employ a procedural minimal view of democracy, we can infer that the conventional cut-offs require higher standards than that. We further conducted additional benchmarking exercises using the BMR with a female suffrage requirement and the LIED as an alternative benchmark variable and similar results. In particular, we found that the conventional cut-offs sit somewhere in-between 'electoral democracy' and 'polyarchy' in the LIED scale, while the cut-offs derived with the three dichotomous indices mostly agree with 'exclusive democracy'.

Given our findings, we suggest that empirical researchers should choose thresholds carefully depending on the research context. For example, when one's theory is concerned with the procedural minimal view of democracy, then the cut-offs we obtained using BMR, CGV and GWF as the benchmarks could be a more appropriate choice than the conventional ones. Since a number of popular databases are not updated every year, empirical researchers often suffer from insufficient data points. In such cases, they can supplement their dataset by discretising the measures from actively updated databases, by using appropriate thresholds estimated by our approach. This would give researchers more opportunities for robustness checks for their analyses with more data with proxy democracy measures.

We end this article with some future research avenues. First, a similar type of threshold assessment can be conducted for other continuous democracy measures. Promising candidates would include the V-Dem's Liberal Democracy Index and the Economist Intelligence Unit's democracy index.[28] Second, our approach can enhance empirical studies in relatively new areas of inquiry that face data coverage challenges, both in terms of the number of countries included and the length of the data collection period. One example would be the studies on hybrid regimes. There is growing consensus that the less-than-democratic regimes are not transitory but constitute a regime type of its own, namely, a hybrid category between democracy and autocracy (Karl, 1995; Diamond, 2002; Bogaards, 2009; Morlino, 2009). However, most existing classifications for such regimes are either limited to a small set of countries and periods (e.g. Diamond, 2002; Merkel, 2004; Morlino, 2009) or based on cut-offs with little theoretical justification.[29] Using the framework of this study, one may obtain empirical thresholds from existing datasets possibly of small size and use them to expand the dataset both intersectionally and longitudinally. This would provide researchers with opportunities to study topics otherwise impossible to investigate, or at minimum, to test their theories with a larger dataset. Finally, one could improve the fractions of agreement by exploring approximating functions other than simple thresholding. Our analyses found that simple thresholding with a single continuous variable achieved an agreement ratio larger than 90%. This score could further be improved by incorporating multiple predictive variables, possibly in combination. A high agreement ratio would benefit the construction of more reliable proxy variables. In addition, analyses with multiple predictive variables would gain qualitative insights on the target measures, which can reveal the properties of the measures from an empirical standpoint.

Notes

1. Technically speaking, FH and Polity are ordinal measures and EDI is an interval scale. For Polity, we used the polity2 variable. All three indices are included in the V-Dem database, https://www.v-dem.net/en/, and we used version 10).
2. Since 2019, V-Dem's annual reports have presented analyses using RoW.
3. We use Version 3.0 of BMR, and the recently updated version of CGV by Bjørnskov and Rode (2020).
4. We use the version 6.0 of the dataset.
5. We reviewed the following sources: *American Journal of Political Science, American Political Science Review, Comparative Political Studies, International Organization, International Studies, International Studies Quarterly, Journal of Conflict Resolution, Journal of Peace Research,*

Journal of Politics, and *World Politics.* The list of the articles discretizing Polity and FH is available from the authors.

6. As of July, 2021, this article has been cited 134 times (https://scholar.google.co.jp/schhp?hl= ja&as_sdt=0,5 (accessed July 15, 2021)).
7. https://freedomhouse.org/report/methodology-freedom-world-2019 (accessed August 1, 2019; The Freedom House also provides a dichotomous 'electoral democracy' coding, but this is less frequently used in scholarly work; thus, we focus on the global freedom score).
8. In some cases, countries with 'Free' and 'Partly Free' combined were coded as a democracy as in Starr and Lindborg (2003).
9. http://www.systemicpeace.org/polity/polity4x.htm (accessed July 26, 2019).
10. We consulted in addition to the codebooks of all versions of the Polity database and the website.
11. 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 et al., 2019, pp. 55–60).
12. Among 18,649 country-year pairs, the rule on EDI individually reproduces the (binary) RoW coding for 18,626 cases. There are 15 cases where EDI is above 0.5 but coded as electoral autocracy, and 8 cases where EDI less than or equal to 0.5 but coded as electoral or liberal democracy.
13. In Version 3.0, 'democracy_femalesuffrage' variable requires that at least half of adult women have the right to vote in addition to their base democracy conditions. The original BMR, which employs adult male suffrage only, is closer to the definition of democracy used in CGV and GWF, because CGV does not have suffrage conditions, and GWF stipulates male suffrage as one of the conditions for a democracy.
14. Some scholars point out that measuring democracy with election outcomes, in this context the turnover requirement, can be problematic (e.g., Bogaards, 2007).
15. Boix et al. (2013) report analogous statistics calculated with older versions of datasets in Table 2.
16. Wahman et al. (2013) conducts a similar analysis with a different calculation scheme. They find the optimal cut-offs for the average of (re-scaled) FH and Polity from the data of country-years of right after regime changes (i.e. years before breakdown and after transition). This data selection, however, may potentially result in unnecessary loss of information since data points where 'regime did not change' are omitted and not incorporated into the calculation. In contrast, our approach utilizes all relevant information for the calculation. In addition, the comparability of re-scaled continuous measures has not been established by the theoretical works yet. As a result, we calculate the threshold for each continuous measure separately as opposed to re-scaling and averaging them.
17. For FH, the intersection is to the left of the dashed line. Because the FH score is lower for more democratic regimes, the direction is the same as the other two measures.
18. In this paper, regime classification thresholds are defined as 'inclusive to the democracy.' For example, a threshold 0.39 for EDI/RoW would mean that a country-year pair is coded as a democracy if and only if the score is ≥ 0.39. Similarly, a threshold of 3.5 for FH would mean that those with a score of ≤ 3.5 would be coded as a democracy.
19. In case multiple cut-offs achieve the highest accuracy, we chose the median of such cut-offs.
20. Appendix B provides the results when using each dichotomous measure separately. The results are, however, mostly identical to the ones provided in Table 3.
21. Despite the methodological difference, this result is mostly consistent with the results by Wahman et al. The cut-offs 3.5 for FH and 5 for Polity would be re-scaled to 5.83 and 7.5 respectively, thus the average is 6.67. This is close to their overall threshold of 7.03. The same calculation applied to the conventional cut-offs (2.5 for FH and 6 for Polity) would be mapped to 7.75 (the average of 7.5 and 8).
22. See also Appendix C. The full list of country-years where classifications change between the conventional and optimal cut-offs are available from the authors upon request. In total, we have 1,928 country-years with 140 distinct countries.

23. To examine the implications of the different thresholds with an external dataset, Appendix D compares the conventional and optimal cut-offs to the classifications coded by area experts for Lain American countries.
24. This variable, named as democracy_femalesuffrage, was added in the Version 3 of BMR released in 2018.
25. In the current version, FH is available from 1972 to 2018. The last year when BMR and BMR-female take different values is 1970 for Switzerland, followed by 1962 (Guatemala and the Republic of Congo).
26. For Polity, the agreement fraction is exactly equal for threshold 5 and 6, and thus, we take the median of the two.
27. We provide visual inspection of the LIED analysis in Appendix E.
28. https://www.eiu.com/topic/democracy-index (accessed on September 10, 2020).
29. For example, Polity labels the intermediate category as 'anocracy' and uses the cut-off value of five. The RoW authors created quadruple sub-types called 'liberal democracy,' 'electoral democracy,' 'electoral autocracy' and 'closed autocracy' and placed a set of predetermined cut-off values.

Notes on contributors

Yuko Kasuya is a professor of Comparative Politics at Keio University. Her research interests include regime transition, comparative political institutions, electoral systems, party politics, and East and Southeast Asian politics.

Kota Mori is a research fellow of the V-Dem East Asia Regional Center. He also works as an industry data scientist, primarily focused on statistical consultation and database development. He received a Ph.D.in economics from Yale University and his research interest is the application of machine learning in social science disciplines.

ORCID

Yuko Kasuya 🔟 http://orcid.org/0000-0002-5324-3204

Bibliography

Alizada, N., Cole, R., Gastaldi, L., Grahn, S., Hellmeier, S., Kolvani, P., Lachapelle, J., Lührmann, A., Maerz, S. M., Pillai, S., & Lindberg, S. I. (2021). Autocratization turns viral. *Democracy Report 2021*. V-Dem Institute.

Adcock, R., & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review*, *95*(3), 529–546. https://doi.org/10.1017/S0003055401003100

Alvarez, M., Cheibub, J. A., Limongi, F. & Przeworski, A. (1996). Classifying political regimes. *Studies in Comparative International Development,* 31(2), 3–36. https://doi.org/10.1007/BF02719326

Bjørnskov, C., & Rode, M. (2020). Regime types and regime change: A new dataset on democracy, coups, and political institutions. *Review of International Organizations*, *15*(2), 531–551. https://doi.org/10.1007/s11558-019-09345-1

Bogaards, M. (2007). Measuring democracy through election outcomes: A critique with African data. *Comparative Political Studies*, *40*(10), 1211–1237. https://doi.org/10.1177/0010414006288968

Bogaards, M. (2009). How to classify hybrid regimes? Defective democracy and electoral authoritarianism. *Democratization*, *16*(2), 399–423. https://doi.org/10.1080/13510340902777800

Bogaards, M. (2010), Measures of democratization: From degree to type to war. *Political Research Quarterly,* 63(2), 475–488. https://doi.org/10.1177/1065912909358578

Bogaards, M. (2012). Where to draw the line? From degree to dichotomy in measures of democracy. *Democratization,* 19(4), 690–712. https://doi.org/10.1080/13510347.2011.613563

Boix, C., Miller, M., & Rosato, S. (2013). A complete data set of Political regimes, 1800–2007. *Comparative Political Studies*, *46*(12), 1523–1554. https://doi.org/10.1177/0010414012463905

Bollen, K. A. (1980). Issues in the comparative measurement of political democracy. *American Sociological Review,* 45(3), 370–390. https://doi.org/10.2307/2095172

Casper, G. & Tufis, C. (2003). Correlation versus interchangeability: The limited robustness of empirical findings on democracy using highly correlated data sets. *Political Analysis,* 11(2), 196–203. https://doi.org/10.1093/pan/mpg009

Cheibub, J. A., Gandhi, J., & Vreeland, J. R. (2010). Democracy and dictatorship revisited. *Public Choice*, *143*(1), 67–101. https://doi.org/10.1007/s11127-009-9491-2

Clark, W. R., Golder, M., & Golder, S. N. (2017). *Principles of comparative politics*. CQ Press.

Collier, D., Hidalgo, F. D. & Maciuceanu, A. O. (2006) Essentially contested concepts: Debates and applications. *Journal of Political Ideologies,* 11(3), 211–246. https://doi.org/10.1080/13569310600923782

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S.I., Teorell, J., Altman, D., Bernhard, M., Fish, M.S., Glynn, A., Hicken, A.,Lührmann, A., Marquardt, K.L., McMann, K., Paxton, P., Pemstein, D., Seim, B., Sigman, R., Skaaning, S., Staton…., & Ziblatt, D. (2019). "V-Dem Codebook v9" Varieties of Democracy (V-Dem) Project.

Coppedge, M., Gerring, J., Altman, D., Bernhard, M., Fish, S., Hicken, A., Kroenig, M., Lindberg, S.I., McMann, K., Paxton, P., Semetko, H.A., Skaaning, S., Staton, S., & Teorell, J. (2011) Conceptualizing and measuring democracy: A new approach. *Perspectives on Politics*, *9*(2), 247–267. https://doi.org/10.1017/S1537592711000880

Dahl, R. A. (1973). *Polyarchy: Participation and opposition*. Yale University Press.

Diamond, L. (2002). Elections without democracy: Thinking about hybrid regimes. *Journal of Democracy*, *13*(2), 21–35. https://doi.org/10.1353/jod.2002.0025

Elkins, Z. (2000). Gradations of democracy? Empirical tests of alternative conceptualizations. *American Journal of Political Science,* 44(2), 293-300. https://doi.org/10.2307/2669311

Gastil, R. D. (1985). The past, present, and future of democracy. *Journal of International Affairs,* 38(2), 161–179.

Gastil, R.D. (1990). The comparative survey of freedom: Experiences and suggestions. *Studies in Comparative International Development,* 25(1), 25–50. https://doi.org/10.1007/BF02716904

Geddes, B., Wright, J. & Frantz, E. (2014a). Autocratic breakdown and regime transitions: A new data set. *Perspectives on Politics,* 12(2), 313–331. https://doi.org/10.1017/S1537592714000851

Geddes, B., Wright, J., & Frantz, E. (2014b). Autocratic regimes code book, version 1.2. https://sites.psu.edu/dictators/wp-content/uploads/sites/12570/2016/05/GWF-Codebook.pdf

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

Gurr, T. R., Jaggers, K., & Moore, W.H. (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. https://doi.org/10.1007/BF02716906

Huntington, S. P. (1991). Democracy's third wave. *Journal of Democracy*, *2*(2), 12–34. https://doi.org/10.1353/jod.1991.0016

Karl, T. L. (1995). The hybrid regimes of Central America. *Journal of Democracy*, *6*(3), 72–86. https://doi.org/10.1353/jod.1995.0049

Linz, J. J. (1975). *Totalitarian and authoritarian regimes*. Lynne Rienner Publishers.

Lueders, H., & Lust, E. (2017). Multiple measurements, elusive agreement, and unstable outcomes in the study of regime change. Varieties of Democracy (V-Dem) Project Working Paper, 52.

Lührmann, A., Tannenberg, M., & Lindberg, S. I. (2018). Regimes of the world (RoW): Opening new avenues for the comparative study of political regimes. *Politics & Governance*, *6*(1), 60–77. https://doi.org/10.17645/pag.v6i1.1214

Marshall, M. G., Gurr, T. R., Davenport, C., & Jaggers, K. (2002). Polity IV, 1800–1999: Comments on Munck and Verkuilen. *Comparative Political Studies*, *35*(1), 40–45.

Merkel, W. (2004). Embedded and defective democracies. *Democratization*, *11*(5), 33–58. https://doi.org/10.1080/13510340412331304598

Morlino, L. (2009). Are there hybrid regimes? Or are they just an optical illusion? *European Political Science Review*, *1*(2), 273–296. https://doi.org/10.1017/S1755773909000198

Munck, G. L. & Verkuilen, J. (2002). Conceptualizing and measuring democracy: Evaluating alternative indices. *Comparative Political Studies*, *35*(1), 5–34.

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

Repucci, S., & Slipowitz, A. (2021). *Freedom in the world 2021: Democracy under siege*. Freedom House.

Sartori, G. (1987). *The theory of democracy revisited*. Oxford University Press.

Skaaning, S., Gerring, J. & Bartusevicius, H. (2015). A lexical index of electoral democracy. *Comparative Political Studies*, *48*(12), 1491–1525. https://doi.org/10.1177/0010414015581050

Starr, H., & Lindborg, C. (2003). Democratic dominoes revisited: The hazards of governmental transitions, 1974–1996. *Journal of Conflict Resolution*, *47*(4), 490–519. https://doi.org/10.1177/0022002703252373

Vanhanen, T. (2000). A new dataset for measuring democracy, 1810–1998. *Journal of Peace Research*, *37*(2), 251–265. https://doi.org/10.1177/0022343300037002008

Wahman, M., Teorell, J. & Hadenius, A. (2013). Authoritarian regime types revisited: Updated data in comparative perspective. *Contemporary Politics*, *19*(1), 19-34. https://doi.org/10.1080/13569775.2013.773200

Wilson, M. C. (2014). A discreet critique of discrete regime type data. *Comparative Political Studies*, *47*(5), 689–714. https://doi.org/10.1177/0010414013488546

Appendix A. Comparison of Three Continuous Measures

Table A1 presents the three continuous measures we examine with regards to their democracy cut-offs. Their cut-off settings are described in the main text.

Appendix B. Analyses Using Three Binary Measures Separately

Table A2 reports the optimal cut-off points and associated fraction of agreement calculated using three binary benchmarks separately. Despite the slight differences across the benchmarks, we obtain the same cut-off points for all combinations.

Appendix C. Fraction of Agreement by the Benchmark Classifications

Figure A1 compares the fraction of agreement of the conventional (light-grey bars) and optimal (dark-grey) thresholds for each pair of continuous and binary measures. To illustrate the different impacts of threshold adjustment, we calculate the fraction of agreement for democracies and auto-cracies in the benchmark classification separately and for all country-years (overall). Note that the optimal cut-offs have been chosen to maximise the overall agreement, so they achieve a higher overall agreement in each panel. This figure reveals that the conventional cut-offs produce less agreement for democracies and higher agreement for autocracies in all panels. The conventional cut-offs tend to judge "democracies" under the benchmark classification as "autocracies," which implies that they require a higher standard of democracy than the benchmark measures. We also observe that the optimal thresholds tend to achieve about the same levels of agreement for democ-racies and autocracies. These thresholds imply that the optimisation process reduces the imbalance between the two groups, thereby attaining higher overall scores.

Appendix D. Latin America

We compare the implied classifications with an external dataset to evaluate the conventional and optimal (calculated with BMR, CGV and GWF benchmarks) thresholds. We employ the Latin America regime classification for 1945–1999 created by Mainwaring, Brinks, and Pérez-Liñán (2001). In this dataset, country-years are categorised into trichotomous regime classifications: "democratic," "semi-democratic," and "authoritarian." The coding is based on four criteria: (1) free and fair com-petitive elections, (2) inclusive franchise, (3) civil liberties and political rights and (4) elected govern-ments governing with effective civilian control over the military. In the first step, each component is evaluated as "no violation," "partial violation" or "major violation." A regime is then coded as "democracy" if all components have no violation, "semi-democracy" if no components have major violation, but some partial violation, and "authoritarian" if some components have major vio-lation. Their definition focuses on procedure but adds a concern for civil liberties and effective gov-erning power. In addition, note that they overlooked the disenfranchisement of women and the illiterate for the early part of the period. In sum, the requirement level is close to the procedural notion of democracy that the three dichotomous benchmarks employ.

In Table A3, we present the cross tabulation between Mainwaring et al. (2001) coding and the classifications implied by two thresholds, conventional and optimal. The percentage in the parenth-esis is the fraction in the row. In general, the conventional cut-offs tend to agree to classify author-itarian regimes (top-left cell), while a lower agreement ratio is found for democratic regimes (bottom-right cell). Note that the cases with "Semi-democratic" coding by Mainwaring et al. (2001) are not coherently coded as democracy or autocracy, suggesting this level of granularity has been lost in the binarisation process.

From this table, we can also calculate the fraction of agreement. Because the treatment of the semi-democratic regimes is ambiguous, we omit them from the calculation. For example, the FH with the conventional cut-off yields 90.5% $(192 + 179)/(192 + 7 + 32 + 179)$, whereas the FH with the optimal cut-off obtains 90.7% $(162 + 210)/(162 + 37 + 1 + 210)$. A similar calculation yields the scores for Polity (91.6% vs 94.8%) and EDI (89.1% vs 95.5%). Overall, the optimal thresholds tend to agree more with the coding by Mainwaring et al. (2001) than the conventional cut-offs. This implies that the requirement of the benchmark variables (BMR, CGV and GWF) is close to the criteria of Mainwaring et al., which mainly focus on the procedural notion and overlook the disenfranchisement of women for the early part of the period.

Appendix E. Optimisation of Threshold Using LIED as Benchmarks

Figure A2 shows the fraction of agreement between LIED benchmarks and the classification implied by threshold levels. Columns represent continuous measures, and rows correspond to the benchmark variables. The dotted and dashed vertical lines indicate the conventional and optimal thresholds.

Appendix F Optimal thresholds for pre- and post-World War Two periods

Table A4 shows the optimal thresholds calculated for 1900–1945 and 1946–2020. The results for the full period are reproduced for comparison purposes. Figure A3 shows the relationship between the threshold and the fraction of agreement more in detail. Due to data availability, we limit the continuous measures to Polity and EDI/RoW and the dichotomous benchmarks to BMR and LIED. Overall, the results for 1946–2020 are similar to the full period results due to the larger fraction of observations for that period (approximately 70 to 80% of the total data points are from 1946 to 2020). From 1900 to 1945, a notable difference is observed for the benchmarks requiring relatively high standards of democracy. For example, the optimal threshold with the benchmark of BMR-female is much higher for Polity and EDI/RoW. Similarly, the thresholds with the LIED6 of class 5 (male democracy) or higher are higher for 1946–2020. In the pre-WWII era, equality in political participation was not fully accepted in most countries, and only highly democratic countries adopted female suffrage.

Reference

Mainwaring, S., Brinks, D. & Pérez-Liñán, A. (2001). Classifying political regimes in Latin America. *Studies in Comparative International Development*, *36*(1), 37–65.
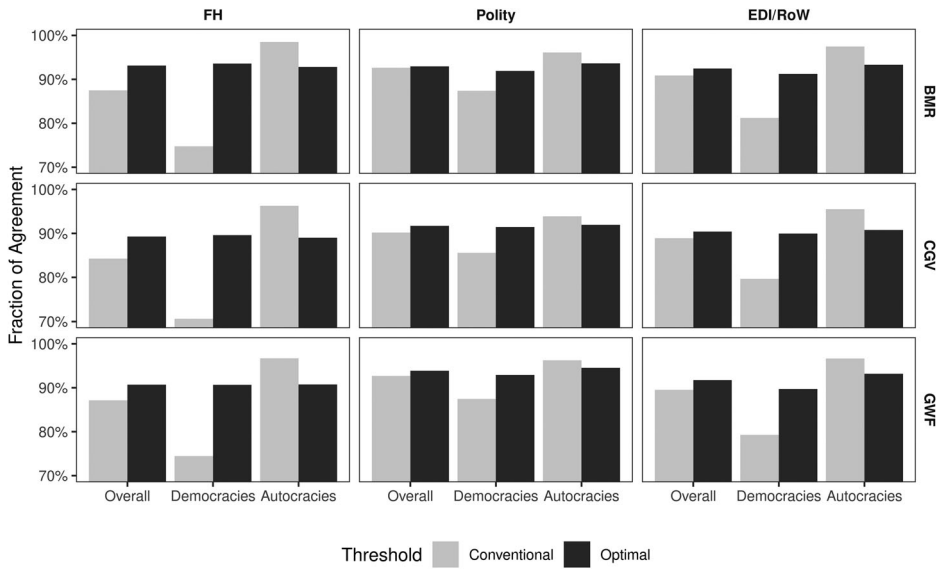
Appendices
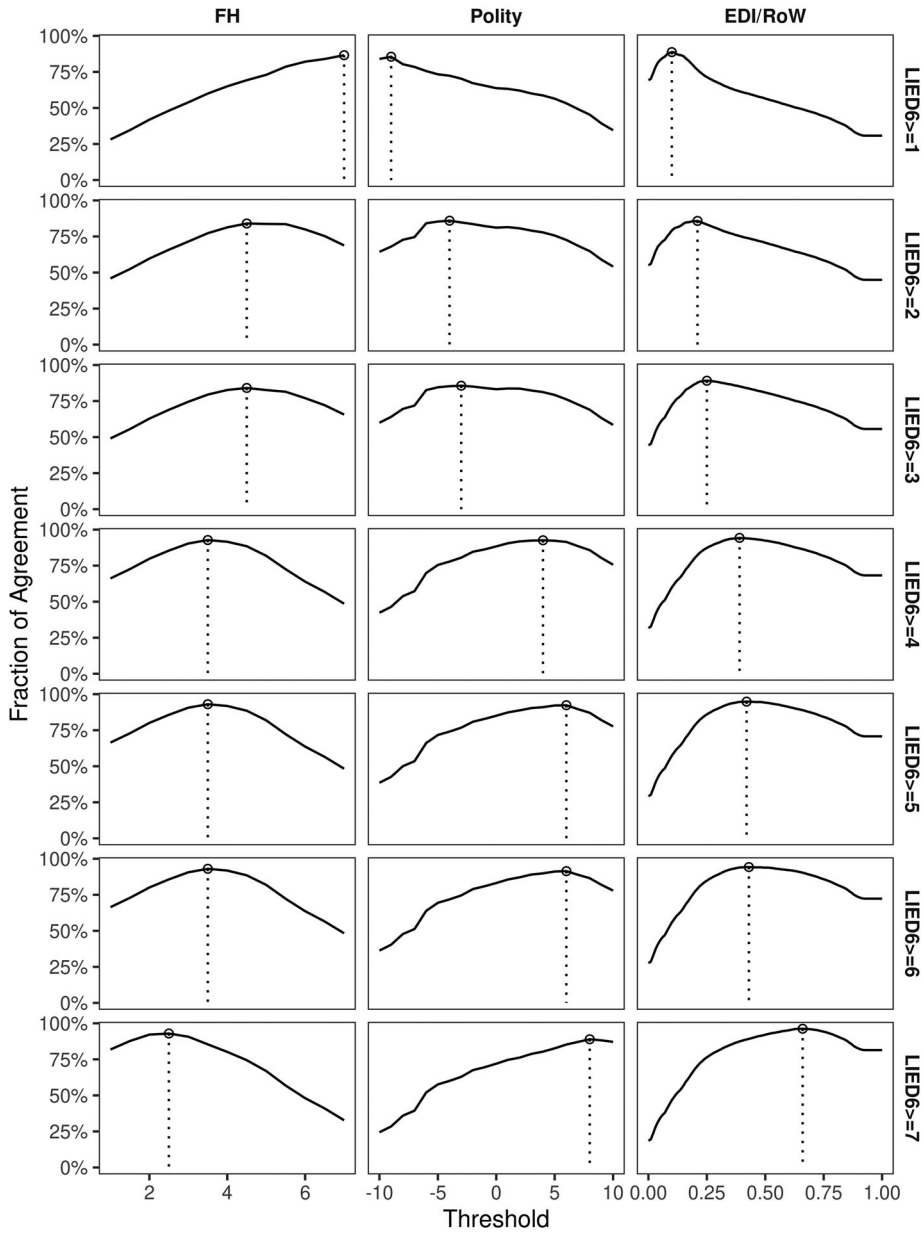

Figure A1. Comparison of fraction of agreement by thresholds

Figure A2. Threshold optimisation with LIED benchmarks

Table A1. Comparison of continuous democracy measures.

	Measurement goals	Democracy conditions	Scale	Cut-off points	Countries	Years
FH	Freedom	political rights (four items); civil rights (four items)	1–7	1.0–2.5 Free 3.0–5.0 Partly free 5.5–7.0 Not free	195	1973–present
Polity	Authority structure	competitiveness/openness of executive recruitments; competitive participation; constraints on executive	−10–10	≥6 Democracy	167	1800–present
EDI/RoW	Dahl's polyarchy	freedom of expression and association; clean election; elected officials; majority suffrage	0–1	≥0.5 Democracy with multiparty/free and fair elections	202	1900–present

Source: Compiled by the authors based on the codebooks of these measures.

Table A2. Optimal cut-off points derived with binary benchmarks separately.

		Conventional		Optimal		
	Benchmark	Cut-off Point	Fraction of Agreement (%)	Cut-off Point	Fraction of Agreement (%)	N. of Obs.
FH	BMR	2.5	87.5	3.5	93.2	6,839
	CGV		84.3	3.5	89.8	7,248
	GWF		87.2	3.5	90.7	4,991
Polity	BMR	6	92.6	5	92.9	11,931
	CGV		90.2	5	91.7	9,205
	GWF		92.7	5	93.9	7,599
EDI/RoW	BMR	0.5	90.9	0.39	92.5	12,433
	CGV		88.9	0.39	90.5	10,879
	GWF		89.5	0.39	91.8	7,703

Table A3. Comparison of thresholds with the classifications by Mainwaring et al.

		Conventional		Optimal	
	Mainwaring et al.'s Coding	Autocracy	Democracy	Autocracy	Democracy
FH	Authoritarian	192 (96.5%)	7 (3.5%)	162 (81.4%)	37 (18.6%)
	Semi-democratic	87 (83.7%)	17 (16.3%)	29 (27.9%)	75 (72.1%)
	Democratic	32 (15.2%)	179 (84.8%)	1 (0.5%)	210 (99.5%)
Polity	Authoritarian	486 (98%)	10 (2%)	486 (98%)	10 (2%)
	Semi-democratic	122 (59.8%)	82 (40.2%)	100 (49%)	104 (51%)
	Democratic	61 (17.6%)	285 (82.4%)	34 (9.8%)	312 (90.2%)
EDI/RoW	Authoritarian	496 (100%)	0 (0%)	491 (99%)	5 (1%)
	Semi-democratic	158 (77.5%)	46 (22.5%)	102 (50%)	102 (50%)
	Democratic	92 (26.6%)	254 (73.4%)	33 (9.5%)	313 (90.5%)

Table A4. Optimal thresholds calculated for period 1900-1945 and 1946-2020. Results for the full period are reproduced for a comparison.

		Optimal Threshold			Fraction of Agreement		
Continuous Measures	Benchmark	1900-2020	1900-1945	1946-2020	1900-2020	1900-1945	1946-2020
Polity	BMR	5	6	5	92.9%	91.8%	93.7%
	BMR-female	5.5	10	5	91.3%	88.6%	93.5%
	LIED6 $\geq$ 1	-9	-9	-9	85.5%	83.7%	86.0%
	LIED6 $\geq$ 2	-4	-4	-4	86.0%	80.0%	87.7%
	LIED6 $\geq$ 3	-3	-3	-4	85.6%	77.7%	88.4%
	LIED6 $\geq$ 4	4	6	4	92.6%	87.9%	94.2%
	LIED6 $\geq$ 5	6	8	5	92.3%	88.0%	93.9%
	LIED6 $\geq$ 6	6	10	5	91.4%	88.0%	93.6%
	LIED6 $\geq$ 7	8	10	8	88.9%	86.8%	90.6%
EDI/RoW	BMR	0.39	0.36	0.39	92.5%	94.3%	92.0%
	BMR-female	0.48	0.62	0.39	91.8%	95.4%	92.0%
	LIED6 $\geq$ 1	0.1	0.1	0.1	88.7%	88.8%	88.7%
	LIED6 $\geq$ 2	0.21	0.15	0.22	85.8%	89.5%	86.5%
	LIED6 $\geq$ 3	0.25	0.26	0.25	89.1%	89.5%	88.9%
	LIED6 $\geq$ 4	0.39	0.405	0.39	94.2%	95.8%	93.5%
	LIED6 $\geq$ 5	0.42	0.485	0.42	94.8%	97.4%	93.8%
	LIED6 $\geq$ 6	0.43	0.62	0.42	94.2%	98.5%	93.8%
	LIED6 $\geq$ 7	0.66	0.73	0.65	96.2%	98.4%	95.4%