Politics and Governance
Open Access Journal | ISSN: 2183-2463

Volume 6, Issue 1 (2018)

Why Choice Matters: Revisiting and Comparing Measures of Democracy

Editors

Heiko Giebler, Saskia P. Ruth and Dag Tanneberg

COGITATIO

Politics and Governance, 2018, Volume 6, Issue 1
Why Choice Matters: Revisiting and Comparing Measures of Democracy

Published by Cogitatio Press
Rua Fialho de Almeida 14, 2º Esq.,
1070-129 Lisbon
Portugal

Academic Editors
Heiko Giebler (WZB Berlin Social Science Center, Germany)
Saskia P. Ruth (German Institute of Global and Area Studies, Germany)
Dag Tanneberg (University of Potsdam, Germany)

Available online at: www.cogitatiopress.com/politicsandgovernance

This issue is licensed under a Creative Commons Attribution 4.0 International License (CC BY). Articles may be reproduced provided that credit is given to the original and *Politics and Governance* is acknowledged as the original venue of publication.

Table of Contents

Why Choice Matters: Revisiting and Comparing Measures of Democracy
Heiko Giebler, Saskia P. Ruth and Dag Tanneberg 1–10

Four Parameters for Measuring Democratic Deliberation: Theoretical and Methodological Challenges and How to Respond
Dannica Fleuß, Karoline Helbig and Gary S. Schaal 11–21

Conceptualizing and Measuring the Quality of Democracy: The Citizens' Perspective
Dieter Fuchs and Edeltraud Roller 22–32

Don't Good Democracies Need "Good" Citizens? Citizen Dispositions and the Study of Democratic Quality
Quinton Mayne and Brigitte Geißel 33–47

Democracy and Human Rights: Concepts, Measures, and Relationships
Todd Landman 48–59

Regimes of the World (RoW): Opening New Avenues for the Comparative Study of Political Regimes
Anna Lührmann, Marcus Tannenberg and Staffan I. Lindberg 60–77

Making Trade-Offs Visible: Theoretical and Methodological Considerations about the Relationship between Dimensions and Institutions of Democracy and Empirical Findings
Hans-Joachim Lauth and Oliver Schlenkrich 78–91

Method Factors in Democracy Indicators
Martin Elff and Sebastian Ziaja 92–104

Different Types of Data and the Validity of Democracy Measures
Svend-Erik Skaaning 105–116

Does the Conceptualization and Measurement of Democracy Quality Matter in Comparative Climate Policy Research?
Romy Escher and Melanie Walter-Rogg 117–144

Politics and Governance (ISSN: 2183–2463)
2018, Volume 6, Issue 1, Pages 1–10
DOI: 10.17645/pag.v6i1.1428

Editorial

Why Choice Matters: Revisiting and Comparing Measures of Democracy

Heiko Giebler [1,]*, Saskia P. Ruth [2] and Dag Tanneberg [3]

[1] Research Unit 'Democracy & Democratization', WZB Berlin Social Science Center, 10785 Berlin, Germany; E-Mail: heiko.giebler@wzb.eu
[2] Institute of Latin American Studies, German Institute of Global and Area Studies, 20354 Hamburg, Germany; E-Mail: saskia.ruth@giga-hamburg.de
[3] Chair of Comparative Politics, University of Potsdam, 14482 Potsdam, Germany; E-Mail: dag.tanneberg@uni-potsdam.de

* Corresponding author

Submitted: 19 February 2018 | Published: 19 March 2018

Abstract
Measures of democracy are in high demand. Scientific and public audiences use them to describe political realities and to substantiate causal claims about those realities. This introduction to the thematic issue reviews the history of democracy measurement since the 1950s. It identifies four development phases of the field, which are characterized by three recurrent topics of debate: (1) what is democracy, (2) what is a good measure of democracy, and (3) do our measurements of democracy register real-world developments? As the answers to those questions have been changing over time, the field of democracy measurement has adapted and reached higher levels of theoretical and methodological sophistication. In effect, the challenges facing contemporary social scientists are not only limited to the challenge of constructing a sound index of democracy. Today, they also need a profound understanding of the differences between various measures of democracy and their implications for empirical applications. The introduction outlines how the contributions to this thematic issue help scholars cope with the recurrent issues of conceptualization, measurement, and application, and concludes by identifying avenues for future research.

Keywords
application; conceptualization; democracy; democratic quality; measurement

Issue
This editorial is part of the issue "Why Choice Matters: Revisiting and Comparing Measures of Democracy", edited by Heiko Giebler (WZB Berlin Social Science Center, Germany), Saskia P. Ruth (German Institute of Global and Area Studies, Germany), and Dag Tanneberg (University of Potsdam, Germany).

© 2018 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Over the past decades, the field of democracy measurement has grown tremendously. The continuous scientific and public demand for measures of democracy has generated an unprecedented wealth of measurement instruments aiming to capture democracy. Yet, having reviewed the development of the field since the 1960s, Bollen (1991, p. 4) found scant evidence for a "smooth evolution towards clear theoretical definitions and finely calibrated instruments". One decade later, Munck and Verkuilen (2002, p. 28) still concluded that "no single index offers a satisfactory response to all three challenges of conceptualization, measurement, and aggregation". But certainly, all is not lost for measuring democracy. Rather, scholars have incorporated much of the critique, which resulted in major improvements. As a result, social sciences today enjoy a vast supply of high-quality approaches to measuring democracy. Now, the challenge is not so much to select a sound index of democracy but rather to understand the theoretical and methodological differences between various indices as well as the consequences of their application. Nevertheless, several recurring topics and issues in the literature on democracy measurement clearly indicate there is still the need for further improvement.

This thematic issue focuses mainly on three aspects to improve research on democracy measurement and research using measures of democracy: (1) *Conceptualization*: what do differences in theoretical grounding and conceptualization of democracy measures mean for empirical analyses? While some measures follow a minimalistic definition of democracy, others go as far as including political outcomes. Moreover, the conceptual differences between graded measures of democracy are seldom in the focus of research. However, they can be quite substantial. Which measures can and should be used for which substantive research questions? (2) *Measurement*: much of the debate on measuring democracy revolves around the nature and scaling of appropriate indicators. Do observables make better or merely different data? Conversely, do expert judgments or public opinion data achieve higher validity or are they just biased in different ways? At the same time, existing measures of democracy differ tremendously in their aggregation rules. What substantive differences do those alternatives imply? (3) *Application*: finally, in terms of real-world developments and resulting research questions, how are the conceptualization and nature of a democracy measure related to its applicability? In line with the contributions to this thematic issue, we argue that the relevance of this question goes far beyond a two-measure comparison in the field of democratization. New measures allow us to address new questions, to revise the answers to old questions, to delve deeper into the realm of causal mechanisms, and, last but not least, more valid application in general. In sum, it seems reasonable to revisit and compare measures of democracy from different perspectives despite all the positive developments in the field.

This editorial serves both as an introduction to and as a summary of the thematic issue. First, we provide a short summary of several development phases in the history of democracy measurement. We show that these developments can be traced back to very similar motives and origins. We then outline why and to which regard there still is the necessity for improvement in the field of democracy measurement. Finally, we discuss how this thematic issue addresses some of the existing problems by presenting the main findings of all nine contributions in the broader context of future avenues for democracy measurement.

2. A Brief History of Democracy Measurement: Improvements, Recurring Topics, and Persistent Shortcomings

In a way, democracy measurement is a prime example of empirical research in political science. The following brief history of the field, which spans more than six decades, demonstrates that scholars adapted democracy measurement in response to major political events or changes in the nature and distribution of political regimes. Whenever existing measures fell out of touch with real-world developments or new substantively important questions arose, democracy measurement was quickly responding. At the same time, democracy measurement remained in close contact with developments in the discipline itself, particularly to democratic theory and political methodology. Three questions drive the development of democracy measurement: (1) what is an appropriate *definition of democracy*, (2) what does an appropriate *measurement* of that definition look like, and (3) are the measures *applicable to real-world phenomena*?

Clearly, there are significant problems regarding research on democracy measurement. As is often the case in the social sciences, it is hardly possible to reach consensus in terms of conceptualization or measurement. 20 years later it is still worth quoting Vanhanen's (1997, p. 31) assessment of democracy measurement in full: "It has been much more difficult to find suitable measures of democracy and to measure the variation in the level of democracy than to formulate a definition of democracy. In fact, nearly all researchers who have attempted to measure democracy have used different indicators. The situation is confusing". Vanhanen describes a very important problem that has yet to be solved. There have been important advancements but we also observe recurring topics and problems in all historical phases of democracy measurement. We count four major phases of development that will be presented in the following with a special focus on the applicability of measures, conceptualization of democracy, and adequacy of measurement.[1]

Phase 1: Modernization Theory and Democracy. Efforts to measure democracy stretch back to the 1950s and 1960s. Following the horrors of World War II and the preceding crisis of democracy in many countries, scholars pondered on the relationship between democracy and modernization. In the course of the debate, Lipset's (1959) "Some Social Requisites of Democracy" advanced as a theoretical and empirical model for numerous future studies. Lipset's seminal piece was not just the first to explicate the link between economic development and democracy (Wucherpfennig & Deutsch, 2009, p. 1), it also translated Dahl's (1956) procedural conception of democracy into a term fit for empirical investigation. In Lipset's words, democracy "is a social mechanism for the resolution of societal decision making among conflicting interest groups" that permits the participation of the largest possible share of the population (Lipset, 1959, p. 71). The questions intriguing Lipset and many others (e.g., Adelman & Morris, 1971; Coleman, 1960; Cutright, 1963; Cutright & Wiley, 1969; R. W. Jackman, 1973; Johnson, 1976; Neubauer, 1967; Smith, 1969) was how to achieve *stable* democracy and how to recognize it.

An occasionally vicious debate over parsimonious and not so parsimonious conceptions of democracy en-

[1] Obviously, it is impossible to provide a complete summary of all different approaches to democracy measurement published since the 1950s. However, the development and nature of all measures follow certain time-specific patterns, which allows us to provide a comprehensive overview nevertheless—albeit on a more abstract level and focusing on the most relevant measures and scholars.

sued, a debate which has continued until today (e.g., Coppedge et al., 2011; Held, 2010; Przeworski, 1999; Schmitter & Karl, 1991). According to Lipset, in the 1950s, stable European democracies had provided uninterrupted rule since World War I and had not met any major domestic anti-democratic movement in the previous 25 years (Lipset, 1959, p. 73). He followed these criteria to separate "stable democracies" from "unstable democracies and dictatorships" (Lipset, 1959, p. 74). Scholars objected to Lipset's admittedly crude dichotomy for various reasons. On the one hand, they criticized his binary classification, which disregarded gradual differences between democracies (e.g., Cutright, 1963). On the other hand, scholars added new properties to the concept of democracy, including aspects of the constitutional state, party competition, as well as the participation (Neubauer, 1967) and representation of citizens (Cutright & Wiley, 1969; Lauth, Pickel, & Welzel, 2000, p. 11). Thus, strictly procedural conceptions of political democracy (Dahl, 1971; Downs, 1957; Schumpeter, 1950) clashed with more universal, social conceptions of democracy which went well beyond the properties of political competition (Bollen, 1991).

Early attempts to measure democracy featured problems of theory, methodology, and applicability. Concepts often did not adequately distinguish between the properties of democracy and its consequences. Moreover, conceptual attributes and indicators of democracy were occasionally only loosely connected. Finally, the resultant indices rarely covered more than a handful of countries or years and were often selected on the basis of data availability rather than substantive criteria. Those deficits in measuring democracy resulted in a barrage of contradictory findings on fundamentally important questions such as the connection between the levels of political democracy and economic inequality (Bollen, 1980).

Phase 2: Differentiation and Sophistication. During the late 1970s and the 1980s, applications of democracy measurements spread to new research areas, e.g., political economy and international relations. More importantly, those years constitute a first blooming of truly comparative measurements of democracy. Many of the most influential measurements of democracy emerged during that period. Freedom House's report on Freedom in the World began its annual circulation in 1978. Although never intended to meet the standards of scientific research (Gastil, 1991, p. 21), its civil liberties and political rights scales found their way into countless scientific applications. In 1975, the first version of the Polity data was used to study patterns of political authority (Gurr Jaggers, & Moore, 1991, p. 73). Interestingly, measuring democracy was not their primary intent, as Polity's famous 21-point scale aggregates patterns of authority related to either autocracy or democracy (Gurr, et al., 1991, p. 79). Vanhanen's (1971) index of democratization, in contrast, directly measured competition and participation in elections, recognizing them as necessary features of democracy (Vanhanen, 2000, p. 256). Finally, Bollen (1980) presented his Political Democracy Index and his extensive use of structural equation models (SEMs) changed the methodological standard in measuring democracy.

SEMs and related approaches use path diagrams to communicate the structure of terms and even entire theories. Path diagrams might be seen as a minor byproduct of a highly specific approach to empirical analysis, but that does not diminish the fact that they made democracy measurement more rigorous. More importantly, Bollen's work reinforced the move towards graded scales in measuring democracy. His analyses built on the firm convictions that (1) democracy is a matter of degree and (2) that dichotomous or even trichotomous measurements of democracy (e.g., Gasiorowski, 1990) introduce substantial measurement errors into the analysis (Bollen, 1991, p. 14; S. Jackman, 2008). Bollen was clearly not the first to advocate degrees of democracy. However, his rigorous assessments of measurement transparency, reliability, and validity justified the use of graded scales in the shadow of a fierce controversy over the proper order of classification and quantification in the social sciences (Sartori, 1970). In short, during this phase best practices emerged in the field. Those prepared the ground for some very important features of modern democracy measurement, e.g., concept trees, theory-consistent aggregation rules (Munck & Verkuilen, 2002), and the conscious choice of data types and sources (Bollen, 1991).

Phase 3: The Age of Hybridization. When the Cold War ended in the early 1990s and Huntington's (1991) "third wave" of democratization surged, the measurement of democracy faced new challenges. On one hand, the number of political systems that were at least minimally democratic had grown substantially. This international spread of democracy also underlined the need for more precise measurements (Lauth et al., 2000, p. 8). On the other hand, democracy indices were criticized for their alleged Western bias. The debate pitted cultural universalism against relativism (Sowell, 1994) and forced existing indices to justify why their conception of democracy should apply across time and space. As more cross-national survey data became available, new opportunities for measuring democracy arose. Hitherto, democracy indices had either privileged factual, easily observable properties of democracy, or had relied on expert knowledge. Now, it became possible to exploit citizen perceptions of democracy for cross-national empirical research and even policy advice.

Severe theoretical and methodological critiques of existing measures of democracy were one of the first types of response to those three developments. Nothing exemplifies them better than the ACLP dataset (Alvarez, Cheibub, Limongi, & Przeworski, 1996; Cheibub, Gandhi, & Vreeland, 2010; Przeworski, Alvarez, Cheibub, & Limongi, 2000). Its binary distinction between democracy and dictatorship excelled with its theoretical clarity and methodological rigor, proving that the debate between discrete and graded measurements of democracy

was far from over. In fact, exchanges between the opponents and proponents of classificatory measurement schemes continued well into the new millennium (Bogaards, 2012; Cheibub, et al., 2010; Collier & Adcock, 1999; Elkins, 2000; Mainwaring, Brinks, & Pérez-Liñán, 2001). Moreover, in the course of the 1990s and early 2000s, Bollen and Paxton repeatedly discussed pitfalls in measuring democracy. They provided evidence for low validity and method factors in important, subjective measures of democracy (Bollen, 1993; Bollen & Paxton, 1998, 2000). Others scrutinized the dimensionality and precision of the Polity data (Gleditsch & Ward, 1997), elaborated on how differences between measurements of democracy resulted in divergent empirical findings (Casper & Tufis, 2003), or developed frameworks for the systematic comparison of measures of democracy (Munck & Verkuilen, 2002). Each of these responses highlight the continuing search for the necessary and theoretically proper degree of precision in measuring democracy.

Citizen evaluations entered the field from two different directions. First, scholars used cross-national surveys to study and compare citizen evaluations of democracy. Those contributions walked the line between measuring democracy and studying political culture. For instance, Welzel, Inglehart and Kligemann (2003) and later on Inglehart and Welzel (2005) used the World Values Survey (WVS) to link macro-level modernization to individual-level aspirations for democracy. Similar connections were later made by Ferrín and Kriesi (2016) who used European Social Survey (ESS) data and thereby demonstrated the ongoing relevance of this research. However, as survey research must rely on standardized questionnaires, those data are not the most granular. Enter the democratic audit, which "constitutes the simple but ambitious project of assessing the state of democracy in a single country" (Beetham, 1994, p. 26). Whereas other measures aim to compare levels of democracy over space and time, the democratic audit proposes a framework for evaluating the living experience of citizens with a democracy. It is every bit as much an analytical as a political tool to empower citizens and it has been used in over 25 countries (Landman, 2012). In retrospect, the audit's success foreshadowed a fundamental shift from measuring the level of democracy to measuring its quality.

Phase 4: Quality, Varieties, and Rollback of Democracy. By the early 2000s, scholars moved away from measuring the *level* of democracy and had begun to take more interest in its *quality* (see Altman & Pérez-Liñán, 2002; Diamond & Morlino, 2005; Pharr & Putnam, 2000). Two observations caused this shift. On one hand, the third wave of democratization left many political regimes stuck in the murky middle ground between democracy and autocracy where multiparty political competition coexists with severe deficits in democratic government. On the other hand, extant measurements rewarded many if not all established democracies with the highest marks, glossing over what struck many as decisive differences. Yet, what is a "good" democracy? Answering that question required newer, broader, mid to wide-range concepts of democracy (e.g., Beetham, 2004; Held, 2010; Merkel, 2004) as well as new data (e.g., Bühlmann, Merkel, Müller, & Weßels, 2012; Coppedge et al., 2011; Lauth, 2015). However, the empirical domain of the quality of democracy remains contested (e.g., Lijphart, 1999; cf. Munck, 2016) as do the attributes of democracy which impact its quality (e.g., Diamond & Morlino, 2005; cf. Lauth, 2011) and the way they matter (e.g., Bochsler & Kriesi, 2013; cf. Giebler & Merkel, 2016). Mirroring the earlier development of the field, the quality of democracy has inspired much conceptual innovation, but no consensus at all (see Fishman, 2016).

While the "value-laden and hence controversial" (Diamond & Morlino, 2005, p. iv) quality of democracy has spawned a rich debate, other scholars have refocused methodological aspects of democracy measurement (see Giebler, 2012) and also types of democracy. Echoing Bollen and Paxton, a series of publications demonstrated the utility of latent variable models for reducing measurement error (Treier & Jackman, 2008), probing dimensionality (Armstrong, 2009), and enhancing validity in democracy measurement (Pemstein, Meserve, & Melton, 2010). Those contributions differ tremendously in their methodological specifics and intent, but they fundamentally agree that extant measures of democracy "capture similar, but often distinct, aspects of what makes states more or less democratic" (Pemstein et al., 2010, p. 427). In other words, extant measures constitute variations on a single theme. The work on varieties of democracy debates that point, arguing that the numerous configurations of institutions and practices can be reduced to a few distinct patterns of democracy. Lijphart's (1999) distinction between majoritarian and consensus democracy constitutes what is probably the best-known example of that debate. Yet, the Varieties of Democracy project has gone much further, outlining and measuring electoral, liberal, majoritarian, participatory, deliberative, and egalitarian conceptions of democracy (Coppedge et al., 2011).

The moment Huntington coined his maritime vocabulary, he warned of a "two-step-forward, one-step-backward" (Huntington, 1991, p. 25) dynamic. All past waves of democratization had been followed by a relapse to authoritarianism and it remained to be seen how persistent the results of the third wave would be. Soon Puddington (2008) titled "Freedom in Retreat: Is the Tide Turning?" and Diamond (2008) announced "The Democratic Rollback", such that other scholars promptly began to ask "Are Dictatorships Returning?" (Merkel, 2010). The answer depends very much on the democracy index used. For Freedom House, which declared an outright crisis of democracy after twelve consecutive years of decline (Abramowitz, 2018, p. 1), the case is blatantly clear. Comparing data provided by Freedom House, Polity IV, the Economist Intelligence Unit, and the Bertelsmann Transformation Index, Levitsky and Way (2015), however,

found scant evidence for a global retreat of democracy between 2000 and 2013. In contrast, based on the latest Varieties of Democracy data, Mechkova, Lührmann and Lindberg (2017, p. 162) produced evidence for a moderate rollback "with democracies becoming less liberal and autocracies less competitive and more repressive". The debate on the rollback of democracy rages on and the field of democracy measurement has effectively come full circle: Once again, scholars are asking what stable democracy looks like.

3. Lessons Learnt? Locating This Thematic Issue in the Debate

Again, the three big questions of the field boil down to matters of conceptualization, measurement, and application. What is democracy? What is a good measure of democracy? Finally, do our measurements of democracy reflect real-world political developments? For decades, scholars have pitted different conceptions of democracy against each other (e.g., Alvarez, et al., 1996; Bollen & Jackman, 1989; Collier & Adcock, 1999) and they have quarreled over the translation of theoretical terms into observational concepts (e.g., Munck & Verkuilen, 2002; Pickel, Stark, & Breustedt, 2015). Disagreement continues and, ultimately, measures of democracy have to prove their value by registering real-world developments. As more sophisticated answers to each of those questions have become available, more practical wisdom is required in order to exploit the full potential of contemporary measures of democracy. Hence, it is time to revisit and compare measures of democracy to show that measurement choice actually matters.

In that regard, prior publications on democracy measurement leave something to be desired. For instance, many of the democracy indices introduced and discussed in the chapters of Inkeles' (1991) ground-breaking edited volume have been abandoned. Although the book still offers an instructive read, it provides limited orientation on *contemporary* democracy measurement. Later contributions such as the now classic review by Munck and Verkuilen (2002) and Munck's (2009) book-long treatment of the topic provide invaluable, comparative assessments of select democracy indices. However, they do not demonstrate in great detail how or why those differences affect empirical research nor do they provide much methodological advice on democracy measurement beyond concept building. Later review articles such as Pickel et al. (2015) often emphasize measuring the quality of democracy as does a recent thematic issue edited by Geißel, Kneuer and Lauth (2016). Moreover, the latter explicitly refrains from supplementing the methodological debate that accompanies measuring democracy since the inception of the field (Geißel et al., 2016, p. 572). However, "theoretical and methodological concerns must go hand-in-hand" (Blalock, 1982, p. 9), and it is the stated intent of this thematic issue to provide practical wisdom on both.

Regarding the *conceptualization of democracy*, Dahl's (1971) concept of polyarchy has held the field together in the past. Virtually every measure of democracy pays respect to its twin dimensions of contestation and participation (Coppedge, Alvarez, & Maldonado, 2008). However, there is a vast number of conceptualization approaches that go beyond Dahl's minimalistic and institution-centered approach (Shapiro, 2003). Moreover, if societies and challenges to democracy change, democracy itself may do so as well, requiring modifications to existing conceptualizations of democracy. Several contributions to this thematic issue tackle that challenge head-on. Fleuß, Helbig and Schaal (2018) show the need for careful theoretical reflection and conceptualization to integrate the demanding concept of democratic deliberation, for which systematic measures exist mainly on the micro and meso level, into democratic performance measures at the macro-level. They highlight that there may not be a one-size-fits-all solution to measures of democratic deliberation and propose a modular approach that builds on different parameters to capture democratic deliberation on the macro level. Their contribution can be considered as a roadmap for future researchers aiming at measuring democratic deliberation at the system level.

Echoing developments in the fourth phase of democracy measurement highlighted above, the contribution by Fuchs and Roller (2018) argues that the quality of democracy is based both on objective (institutional as well as procedural characteristics) and subjective criteria (public opinion). Hence, they translate different normative models of democracy to the level of public opinion data and measure their acceptance in different countries all over the world. While this does not contest the various models in terms of their conceptualization it does indeed challenge any institutional or formal approach to democracy measurement. In a similar way, Mayne and Geißel (2018) turn their attention to the crucial role of citizens in democratic quality assessments. Aiming to identify what constitutes a "good" citizen they conceptualize and discuss potential measures of three citizen dispositions that make up the citizen component of democratic quality "breathing life" into democratic institutions. Moreover, Mayne and Geißel (2018) raise awareness of the fact that different institutional models of democracy consider different types of citizen (i.e., different dispositions) as being "good" (or bad) for democratic quality.

Landman (2018) turns to another important point of debate, asking to what degree democracy and human rights overlap. Recent contributions claim inextricable theoretical and empirical connections between democracy and human rights (Hill, 2016; Hill & Jones, 2014). Landman, in contrast, demonstrates that those connections are variable and depend systematically on the conception of democracy employed. In a related effort, Lührmann, Tannenberg and Lindberg (2018) revisit the debate on degrees of democracy and types of political regimes. Their contribution underlines the importance

of a conceptually sound distinction between closed and electoral autocracies on one hand and electoral and liberal democracies on the other hand. Based on the latest V-Dem data, the article also serves as an interesting proof of concept: Classification and quantification can go hand in hand if the underlying data receive careful attention.

In fact, *democracy measurement* requires as much careful attention as conceptualization. Although the field has accumulated much methodological wisdom over the years, several lacunae remain. The contribution of Lauth and Schlenkrich (2018) presents an approach to tackle the long-standing question of whether there is a trade-off between certain democratic principles—first and foremost whether both freedom and equality can be maximized in a democracy. In doing so, it allows for a more valid operationalization of various democratic models without demanding a normative decision on which model constitutes the best model of democracy. Elff and Ziaja (2018) carry on the work of Bollen and others. The authors use confirmatory factor analysis to gauge potentially biasing method factors in four high profile measures of democracy. Their results serve as a sobering reminder not to take measures of democracy at face value. As the authors show, apparent differences between or trends within countries may be more telling about the measurement instrument itself than about real-world developments.

Skaaning's (2018) discussion of different types of data sources makes a similar point. Premised on the assumption that observational features of political regimes have as many drawbacks as in-house coding and expert or population surveys, Skaaning reflects on the numerous trade-offs involved in measuring democracy reliably and validly. The article carefully considers each type of data and formulates numerous best-practices for the production and application of democracy data. Even though scholars have developed several measures that are able to detect differences even in established democracies, Fuchs and Roller (2018) show that the variation might still be underestimated if public opinion data on democratic quality is not taken into account. Hence, they link the debate on data types to the debate on more hybrid manifestations of democracy.

In terms of *application*, finally, the thematic issue showcases efforts to explain divergent empirical findings by theoretical and methodological differences between extant measures of democracy and it provides guidance on best practices. The article by Escher and Walter-Rogg (2018) constitutes a mixture of replication and genuine research. It sheds light on the question of whether democracy is good or bad for climate protection. In contrast to earlier approaches, the article makes use of the multi-level and multi-branch tree approach to democracy measurement. Distinguishing between different features and sub-dimensions of democracy, the authors show that only certain features of democracy have a positive impact on climate protection and that the underlying mechanisms are impossible to identify if scholars focus only on highly aggregated democracy scores. Linking this to the trade-off approach developed by Lauth and Schlenkrich (2018), there is clear evidence that scholars should also make use of democracy measures below their highest levels of aggregation.

Moreover, data from different sources can be better compared or combined at lower levels of aggregation. The potential gain is twofold. First, there is an increased awareness of ambiguities implied by differences in concept building and operationalization between measures of democracy. The grand tour of regime classifications provided by Lührmann et al. (2018), for instance, shows in a scrupulously precise way how strongly even minor differences between measures of democracy affect descriptive inference at higher levels of aggregation. Second, the combination of different data sources or types promises to overcome the limitations inherent to each of them. This holds true at all stages of the research process as shown by the contributions of Elff and Ziaja (2018) and Skaaning (2018). Finally, the contributions by Fleuß et al. (2018) as well as Landman (2018) highlight how the application of certain more maximalist definitions of democracy could change the comparison of different democratic regimes significantly.

Without doubt, this thematic issue will not be the final contribution to the vast body of democracy measurement literature and having read this introduction, many will agree that such a convergence is unlikely. Democratic regimes are confronted with new and different developments and challenges, as are the researchers who try to measure the state of such democracies. This is not a problem but rather distinguishes scientific approaches from normative teleology and doomsday rhetoric.

As our look into the history of democracy measurement has shown, these days we are blessed with much more advanced and nuanced measures—in theoretical as well as in methodological terms. These allow us to address both new and old questions which are of relevance for many different research areas. Choice indeed matters! Ongoing debates in democracy measurement, which certainly will be influenced by the contributions to this thematic issue, underline that complacency is not a virtue in academia. Fortunately, this thematic issue reveals that further improvement is not only necessary but also possible.

Acknowledgments

We would like to thank Daniel Bochsler from the Democracy Barometer project as well as Daniel Kübler from the National Center of Competence in Research (NCCR) Democracy at the University of Zürich for providing the funds to conduct an author workshop in Zürich in May 2017. We also would like to thank Lea Heyne and Yvonne Rosteck for their help in organizing the event. Our thanks go to the German Institute for Global Area Studies (GIGA) for providing additional funds. Moreover, we are indebted to Katarina Pollner who proofread an earlier ver-

sion of this editorial. Finally, we would like to thank all the authors who contributed to this thematic issue for all their efforts and excellent articles.

Conflict of Interests

The authors declare no conflict of interests.

References

Abramowitz, M. J. (2018). Democracy in crisis. *Freedom in the world 2018*. Washington, DC, and New York, NY: Freedom House.

Adelman, I., & Morris, C. T. (1971). *A conceptualization and analysis of political participation in underdeveloped countries. Part 2 of the final report to the Agency of International Development*. Washington, DC: Agency of International Development.

Altman, D., & Pérez-Liñán, A. (2002). Assessing the quality of democracy: Freedom, competitiveness and participation in eighteen Latin American Countries. *Democratization*, *9*(2), 85–100.

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

Armstrong, D. A. (2009). Measuring the democracy–repression nexus. *Electoral Studies*, *28*(3), 403–412.

Beetham, D. (1994). Key principles and indices for a democratic audit. In D. Beetham (Ed.), *Defining and measuring democracy* (pp. 25–43). London: Sage Publications.

Beetham, D. (2004). Towards a universal framework for democracy assessment. *Democratization*, *11*(2), 1–17.

Blalock, H. M. (1982). *Conceptualization and measurement in the social sciences*. Beverly Hills, CA: Sage Publications.

Bochsler, D., & Kriesi, H. (2013). Varieties of democracy. In H. Kriesi, S. Lavenex, F. Esser, J. Matthes, M. Bühlmann, & D. Bochsler (Eds.), *Democracy in the age of globalization and mediatization* (pp. 96–102). Basingstoke: Palgrave Macmillan.

Bogaards, M. (2012). Where to draw the line? From degree to dichotomy in measures of democracy. *Democratization*, *19*(4), 690–712.

Bollen, K. A. (1980). Issues in the comparative measurement of political democracy. *American Sociological Review*, *45*(3), 370–390.

Bollen, K. A. (1991). Political democracy. Conceptual and measurement traps. In A. Inkeles (Ed.), *On measuring democracy. Its consequences and concomitants* (pp. 3–20). New Brunswick and London: Transaction Publishers.

Bollen, K. A. (1993). Liberal democracy: Validity and method factors in cross-national measures. *American Journal of Political Science*, *37*(4), 1207–1230.

Bollen, K. A., & Jackman, R. W. (1989). Democracy, stability, and dichotomies. American Sociological Review, 54(4), 612-621.

Bollen, K. A., & Paxton, P. (1998). Detection and determinants of bias in subjective measures. *American Sociological Review*, *63*(3), 465–478.

Bollen, K. A., & Paxton, P. (2000). Subjective measures of liberal democracy. *Comparative Political Studies*, *33*(1), 58–86.

Bühlmann, M., Merkel, W., Müller, L., & Weßels, B. (2012). The democracy barometer: A new instrument to measure the quality of democracy and its potential for comparative research. *European Political Science*, *11*(4), 519–536.

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.

Cheibub, J. A., Gandhi, J., & Vreeland, J. R. (2010). Democracy and dictatorship revisited. *Public Choice*, *143*(1), 67–101.

Coleman, J. S. (1960). Conclusion: The political systems of the developing areas. In G. A. Almond & J. S. Coleman (Eds.), *The politics of the developing area* (pp. 532–576). Princeton, NJ: Princeton University Press.

Collier, D., & Adcock, R. (1999). Democracy and dichotomies: A pragmatic approach to choices about concepts. *Annual Review of Political Science*, *2*(1), 537–565.

Coppedge, M., Alvarez, A., & Maldonado, C. (2008). Two persistent dimensions of democracy: Contestation and inclusiveness. *The Journal of Politics*, *70*(3), 632–647.

Coppedge, M., Gerring, J., Altman, D., Bernhard, M., Fish, S., Hicken, A., . . . Teorell, J. (2011). Conceptualizing and measuring democracy: A new approach. *Perspectives on Politics*, *9*(2), 247–267.

Cutright, P. (1963). National political development: Measurement and analysis. *American Sociological Review*, *28*(2), 253–264.

Cutright, P., & Wiley, J. A. (1969). Modernization and political representation: 1927–1966. *Studies in Comparative International Development*, *5*(2), 23–44.

Dahl, R. A. (1956). *A preface to democratic theory*. Chicago, IL, and London: University of Chicago Press.

Dahl, R. A. (1971). *Polyarchy: Participation and opposition*. New Haven, CT: Yale University Press.

Diamond, L. (2008). The democratic rollback. The resurgence of the predatory state. *Foreign Affairs*, *87*(2), 36–48.

Diamond, L., & Morlino, L. (2005). *Assessing the quality of democracy*. Baltimore, MD: The Johns Hopkins University Press.

Downs, A. (1957). *An economic theory of democracy*. New York, NY: Harper & Row Publishers.

Elff, M., & Ziaja, S. (2018). Method factors in democracy indicators. *Politics and Governance*, *6*(1), 92–104.

Elkins, Z. (2000). Gradations of democracy? Empirical tests of alternative conceptualizations. *American Journal of Political Science*, *44*(2), 293–300.

Escher, R., & Walter-Rogg, M. (2018). Does the concep-

tualization and measurement of democracy quality matter in comparative climate policy research? *Politics and Governance*, *6*(1), 117–144.

Ferrín, M., & Kriesi, H. (Eds.). (2016). *How Europeans view and evaluate democracy*. Oxford: Oxford University Press.

Fishman, R. M. (2016). Rethinking dimensions of democracy for empirical analysis: Authenticity, quality, depth, and consolidation. *Annual Review of Political Science*, *19*(1), 289–309.

Fleuß, D., Helbig, K., & Schaal, G. S. (2018). Four parameters for measuring democratic deliberation: Theoretical and methodological challenges and how to respond. *Politics and Governance*, *6*(1), 11–21.

Fuchs, D., & Roller, E. (2018). Conceptualizing and measuring the quality of democracy: The citizens' perspective. *Politics and Governance*, *6*(1), 22–32.

Gasiorowski, M. J. (1990). The political regimes project. *Studies in Comparative International Development*, *25*(1), 109–125.

Gastil, R. D. (1991). The comparative survey of freedom: Experiences and suggestions. In A. Inkeles (Ed.), *On measuring democracy. Its consequences and concomitants* (pp. 21–46). New Brunswick and London: Transaction Publishers.

Geißel, B., Kneuer, M., & Lauth, H.-J. (2016). Measuring the quality of democracy: Introduction. *International Political Science Review*, *37*(5), 571–579.

Giebler, H. (2012). Bringing methodology (back) in: Some remarks on contemporary democracy measurements. *European Political Science*, *11*(4), 509–518.

Giebler, H., & Merkel, W. (2016). Freedom and equality in democracies: Is there a trade-off? *International Political Science Review*, *37*(5), 594–605.

Gleditsch, K. S., & Ward, M. D. (1997). Double take: A reexamination of democracy and autocracy in modern polities. *The Journal of Conflict Resolution*, *41*(3), 361–383.

Gurr, T. R., Jaggers, K., & Moore, R. T. (1991). The transformation of the Western state: The growth of democracy, autocracy, and state power since 1800. In A. Inkeles (Ed.), *On measuring democracy. Its consequences and concomitants* (pp. 69–104). New Brunswick and London: Transaction Publishers.

Held, D. (2010). *Models of democracy*. Cambridge: Polity Press.

Hill, D. W. (2016). Democracy and the concept of personal integrity rights. *The Journal of Politics*, *78*(3), 822–835.

Hill, D. W., & Jones, Z. M. (2014). An empirical evaluation of explanations for state repression. *American Political Science Review*, *108*(3), 661–687.

Huntington, S. P. (1991). Democracy's third wave. *Journal of Democracy*, *2*(2), 12–34.

Inglehart, R., & Welzel, C. (2005). *Modernization, cultural change, and democracy. The human development sequence*. Cambridge: Cambridge University Press.

Inkeles, A. (Ed.). (1991). *On measuring democracy. Its consequences and concomitants*. New Brunswick and London: Transaction Publishers.

Jackman, R. W. (1973). On the relation of economic development to democratic performance. *American Journal of Political Science*, *17*(3), 611–621.

Jackman, S. (2008). Measurement. In J. M. Box-Steffensmeier, H. E. Brady, & D. Collier (Eds.), *The Oxford handbook of political methodology* (pp. 119–151). Oxford: Oxford University Press.

Johnson, K. F. (1976). Scholarly images of Latin American political democracy in 1975. *Latin American Research Review*, *11*(2), 129–140.

Landman, T. (2012). Assessing the quality of democracy: The international IDEA framework. *European Political Science*, *11*(4), 456–468.

Landman, T. (2018). Democracy and human rights: Concepts, measures, and relationships. *Politics and Governance*, *6*(1), 48–59.

Lauth, H.-J. (2011). Quality criteria for democracy. Why responsiveness is not the key. In G. Erdmann & M. Kneuer (Eds.), *Regression of democracy? Zeitschrift für Vergleichende Politikwissenschaft Comparative Governance and Politics* (pp. 59–80). Wiesbaden: VS Verlag für Sozialwissenschaften.

Lauth, H.-J. (2015). *The matrix of democracy. A three-dimensional approach to measuring the quality of democracy and regime transformations*. Würzburg: Universität Würzburg.

Lauth, H.-J., Pickel, G., & Welzel, C. (2000). Grundfragen, Probleme und Perspektiven der Demokratiemessung [Fundamental questions, problems and perspectives on democracy measurement]. In H.-J. Lauth, G. Pickel, & C. Welzel (Eds.), *Demokratiemessung. Konzepte und Befunde im internationalen Vergleich* [*Democracy measurement. Concepts and results in comparative perspective*] (pp. 7–26). Wiesbaden: VS Verlag für Sozialwissenschaften.

Lauth, H.-J., & Schlenkrich, O. (2018). Making trade-offs visible: Theoretical and methodological considerations about the relationship between dimensions and institutions of democracy and empirical findings. *Politics and Governance*, *6*(1), 78–91.

Levitsky, S., & Way, L. (2015). The myth of democratic regression. *Journal of Democracy*, *26*(1), 45–58.

Lijphart, A. (1999). *Patterns of democracy. Government forms and performance in thirty-six countries*. New Haven, CT, and London: Yale University Press.

Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *The American Political Science Review*, *53*(1), 69–105.

Lührmann, A., Tannenberg, M., & Lindberg, S. I. (2018). Regimes of the world (RoW): Opening new avenues for the comparative study of political regimes. *Politics and Governance*, *6*(1), 60–77.

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

Mayne, Q., & Geißel, B. (2018). Don't good democracies need "good" citizens? Citizen dispositions and the study of democratic quality. *Politics and Governance*, *6*(1), 33–47.

Mechkova, V., Lührmann, A., & Lindberg, S. I. (2017). How much democratic backsliding? *Journal of Democracy*, *28*(4), 162–169.

Merkel, W. (2004). Embedded and defective democracies. *Democratization*, *11*(5), 33–58.

Merkel, W. (2010). Are dictatorships returning? Revisiting the 'democratic rollback' hypothesis. *Contemporary Politics*, *16*(1), 17–31.

Munck, G. L. (2009). *Measuring democracy. A bridge between scholarship and politics*. Baltimore, MD: Johns Hopkins University Press.

Munck, G. L. (2016). What is democracy? A reconceptualization of the quality of democracy. *Democratization*, *23*(1), 1–26.

Munck, G. L., & Verkuilen, J. (2002). Conceptualizing and measuring democracy: Evaluating alternative indices. *Comparative Political Studies*, *35*(1), 5–34. Neubauer, D. E. (1967). Some conditions of democracy. *The American Political Science Review*, *61*(4), 1002–1009.

Pemstein, D., Meserve, S. A., & Melton, J. (2010). Democratic compromise: A latent variable analysis of ten measures of regime type. *Political Analysis*, *18*(4), 426–449.

Pharr, S. J., & Putnam, R. D. (Eds.). (2000). *Disaffected democracies: What's troubling the trialateral countries?* Princeton, NJ: Princeton University Press.

Pickel, S., Stark, T., & Breustedt, W. (2015). Assessing the quality of quality measures of democracy: A theoretical framework and its empirical application. *European Political Science*, *14*(4), 496–520.

Przeworski, A. (1999). Minimalist conception of democracy. A defense. In I. Shapiro & C. Hacker-Cordón (Eds.), *Democracy's value* (pp. 23–55). Cambridge: Cambridge University Press.

Przeworski, A., Alvarez, M., Cheibub, J. A., & Limongi, F. (2000). *Democracy and development: Material well-being in the world, 1950–1990*. Cambridge and New York, NY: Cambridge University Press.

Puddington, A. (2008). Is the tide turning? *Journal of Democracy*, *19*(2), 61–73.

Sartori, G. (1970). Concept misformation in comparative politics. *The American Political Science Review*, *64*(4), 1033–1053.

Schmitter, P. C., & Karl, T. L. (1991). What democracy is…and is not. *Journal of Democracy*, *2*(3), 75–88.

Schumpeter, J. A. (1950). *Capitalism, socialism & democracy*. New York, NY: Harper & Brothers.

Shapiro, I. (2003). *The state of democratic theory*. Princeton, NJ: Princeton University Press.

Skaaning, S.-E. (2018). Different types of data and the validity of democracy measures. *Politics and Governance*, *6*(1), 105–116.

Smith, A. K. (1969). Socio-economic development and political democracy: A causal analysis. *Midwest Journal of Political Science*, *13*(1), 95–125.

Sowell, T. (1994). *Race and culture. A world view*. New York, NY: Basic Books.

Treier, S., & Jackman, S. (2008). Democracy as a latent variable. *American Journal of Political Science*, *52*(1), 201–217.

Vanhanen, T. (1971). *Dependence of power on resources. A comparative study of 114 states in the 1960s* (Vol. 1). Jyväskylä: University of Jyväskylä.

Vanhanen, T. (1997). *Prospects of democracy. A study of 172 countries*. London: Routledge.

Vanhanen, T. (2000). A new dataset for measuring democracy, 1810–1998. *Journal of Peace Research*, *37*(2), 251–265.

Welzel, C., Inglehart, R., & Kligemann, H.-D. (2003). The theory of human development: A cross-cultural analysis. *European Journal of Political Research*, *42*(3), 341–379.

Wucherpfennig, J., & Deutsch, F. (2009). Modernization and democracy: Theories and evidence revisited. *Living Reviews in Democracy*, *1*, 1–9.

About the Authors

Heiko Giebler is a Senior Researcher at WZB Berlin Social Science Center, Germany. He is also the head of two large-scale projects dealing with the psychological underpinnings of populism and the consequences of populism for democracy. He works on political behavior, populism, democracy measurement, and applied quantitative methodology. Heiko Giebler has published articles in several peer-reviewed journals, among others the British Journal of Political Science, German Politics, and Electoral Studies as well as several edited volumes and special issues.

Saskia P. Ruth a Research Fellow at the German Institute of Global and Area Studies in Hamburg, Germany. Her research focuses on the consequences of populism and clientelism for the quality of democratic governance with a special focus on the Latin American region. She has published articles in The Journal of Politics, Latin American Politics and Society, Political Studies, as well as Swiss Political Science Review.

Dag Tanneberg is a Research Fellow at the Chair of Comparative Politics, University of Potsdam, Germany. His research interests span authoritarian rule, contentious politics, state repression, and applied methods in political science. Dag Tanneberg has published in several peer-reviewed journals, among them Politische Vierteljahresschrift, Zeitschrift für Vergleichende Politikwissenschaft, and Contemporary Politics.

COGITATIO

Politics and Governance (ISSN: 2183–2463)
2018, Volume 6, Issue 1, Pages 11–21
DOI: 10.17645/pag.v6i1.1199

Article

Four Parameters for Measuring Democratic Deliberation: Theoretical and Methodological Challenges and How to Respond

Dannica Fleuß *, Karoline Helbig and Gary S. Schaal

Department of Political Science, Helmut-Schmidt-University, 22043 Hamburg, Germany; E-Mails: dannica.fleuss@hsu-hh.de (D.F.), karoline.helbig@hsu-hh.de (K.H.), gschaal@hsu-hh.de (G.S.S.)

* Corresponding author

Submitted: 30 September 2017 | Accepted: 8 November 2018 | Published: 19 March 2018

Abstract
Although measuring democratic deliberation is necessary for a valid measurement of the performance of democracies, it poses serious theoretical and methodological challenges. The most serious problem in the context of research on democratic performance is the need for a theoretical and methodological approach for "upscaling" the measurement of deliberation from the micro and meso level to the macro level. The systemic approach offers a useful framework for this purpose. Building on this framework, this article offers a modular approach consisting of four parameters for conceptualization, measurement, and aggregation which can be adjusted to make the measurement of democratic deliberation compatible with the various general measurement approaches adopted by different scholars.

Keywords
deliberation; democracy; democratic performance; measurement of democracy; systemic approach

Issue
This article is part of the issue "Why Choice Matters: Revisiting and Comparing Measures of Democracy", edited by Heiko Giebler (WZB Berlin Social Science Center, Germany), Saskia P. Ruth (German Institute of Global and Area Studies, Germany), and Dag Tanneberg (University of Potsdam, Germany).

© 2018 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

For a long time, liberal democracies' legitimacy mainly rested on voting and representation. In the course of the so-called "participatory revolution" (Kaase, 1984), the means of participation have increased dramatically in Western democracies. The term "democratic innovation" refers to the new, multi-faceted forms of participation which go beyond voting. Most of them are built around the idea and practice of deliberation in one way or another (Geißel & Newton, 2012). Today, the theory of deliberative democracy is considered to be the most important normative theory of democracy (Dryzek, 2015; Elstub, 2015). Accordingly, authors such as Dryzek (2010, pp. 21–42) argue that deliberative legitimacy has become the most important paradigm of legitimacy in contemporary political *theory* as well as democratic *practice*.

Due to the vastly increased theoretical importance and empirical impact of democratic theories of deliberation, measures of democracy need to include deliberation to achieve valid and empirically meaningful results. The Discourse Quality Index (DQI) already offers a sophisticated and widely acclaimed measuring instrument for democratic deliberation at the *micro/meso level*. However, an evaluation of the deliberative performance of democratic political systems at large requires measuring *deliberation at the macro level*. Niemeyer (2014) questioned whether scaling up deliberation was possible. A theoretically and methodologically grounded framework for this purpose is still required (Niemeyer, Curato, & Bächtiger, 2015, p. 4).

Addressing this gap in research, we develop an approach for measuring the deliberative performance[1] of political systems and systematically outline four "parameters for measuring macro deliberation" (PMMD). Thereby, we propose a modular approach that can be adopted by other scholars with different normative and theoretical presuppositions. The parameters are based on the assumption that the so-called "systemic approach", as originally developed in a seminal book by Mansbridge et al. (2012), is the only framework for conceptualizing democratic deliberation[2] at the macro level to have been suggested so far which represents a suitable basis for measuring deliberative performance at this level. In this approach, deliberation is conceptualized as an "emergent property" (Niemeyer et al., 2015) which cannot be reduced to a mere aggregation of other qualities of the political system (see O'Conner & Wong, 2015). Rather than isolated deliberative fora, "the interdependence of sites within a larger system" as well as the interactions of deliberations in different institutions (and *loci* in general) represent the focal point of this understanding (Bohman, 2012, p. 73; Mansbridge et al., 2012, pp. 1f.).

In order to have a common point of departure, we take the original systemic approach as a starting point to develop a framework for "scaling up" the measurement of democratic deliberation from the micro to the macro level (see Niemeyer, 2014).[3] This does not imply any commitment to Mansbridge et al.'s (2012) specific *normative* premises (and especially not to their concept of deliberation), but only to the basic framework of the systemic approach. Following an explanation of the need for a new approach to measure deliberation at the macro level in section 2, we outline the four "parameters" that have to be considered in the process of conceptualization and operationalization: *the theory of democracy*, *the concept of deliberation*, *the selection of loci*, *and the aggregation rule* (section 3). In the concluding section of this article, we identify some challenges future research will need to deal with regarding the measurement of the deliberative performance of democratic political systems.

2. Why "Parameters for Measuring Macro Deliberation"?

Elstub, Ercan and Mendonça (2016) identify four generations of the deliberative democracy school of thought. They started out with an explicitly normative theory on the rational, impartial justification of norms in the first generation (Cohen, 1989, 1997; Habermas, 1996), continued to adapt the definition of deliberation to the increasing plurality and complexity of contemporary democracies in the second generation (Chambers, 2003, p. 322; Elstub et al., 2016, pp. 141f.), and began the empirical evaluation of deliberation under "laboratory conditions" in the third generation (Fishkin, 1995). In the following decades, "real world deliberation" became an object of scholars' interest. With the DQI, the "gold standard" for the evaluation of institutions or the public sphere's deliberative performance was developed (Mansbridge, 2010; Steiner, Bächtiger, Spörndli, & Steenbergen, 2004) and scholars such as Fung (2006, p. 66) discussed the "range of institutional possibilities for public participation".

The *fourth generation* is characterized by the so-called "systemic turn", which attempts to combine "the insights gained from three preceding generations, namely the strong normative premises, institutional feasibility, and empirical results" (Elstub et al., 2016, p. 143). The major innovation of the systemic approach is the acknowledgement of the "importance of looking at the system as a whole, as well as its different parts" rather than the previous focus on "isolated instances of deliberation" (Erman, 2016, p. 263). Thereby, "the deliberative system reconnects deliberative democratic theory to its initial macro ambitions: to enhance and understand democracy at the large scale" (Boswell & Corbett, 2017, p. 3). This implies that the systemic approach attempts to conceptualize democratic deliberation(s) as taking place all over a society or a political system and seeks to systematically account for the interactive relationships between various deliberative practices (Elstub et al., 2016, p. 140).

In spite of the considerations of *fourth generation* scholars, even two of the most sophisticated contemporary measures of democracy are unable to provide an appropriate measure of democratic deliberation at the macro level: while the *Democracy Barometer* (Merkel et al., 2016) does not take democratic deliberation into account at all, *Varieties of Democracy* (Coppedge et al., 2016) explicitly offers a "deliberative component index". However, the latter demonstrates one of the major problems of addressing democratic deliberation in a democracy index: in their attempt to transfer criteria that were applicable at a micro/meso level to a larger scale, Coppedge et al. (2016) do not take account of the difference between the criteria for deliberation at the micro/meso level and at the macro level. Although they claimed to measure "deliberation at all levels" (Coppedge et al., 2016, p. 6), the items mostly address formalized deliberation by political elites. More-

[1] The concept used by John S. Dryzek (2009); Stevenson and Dryzek (2012) and many subsequent researchers (e.g. Felicetti, Niemeyer, & Curato, 2016; Niemeyer et al., 2015) is the concept of "deliberative capacity". Although this concept undoubtedly "provides diagnostic criteria for assessing the system" (Felicetti et al., 2016, p. 429), we rely on the concept of "(deliberative) performance", which is much more common in the context of measuring the performance of democratic political systems.

[2] We explicitly address the measurement of "democratic deliberation", not "deliberative democracy". While the latter denotes a political system that meets the demanding normative standards of deliberative theories, the latter refers to a form of communication, originally derived from these theories, put into practice in (liberal) democracies (Mansbridge, 2007).

[3] More recent developments by Hendriks (2016a, 2016b), Ercan, Hendriks and Boswell (2017) and Erman (2016) will be integrated in section 3 of this article.

over, the aggregation rule is simply a factor index (that is, an additive index weighted by the factor loadings of the items) and does not reflect the complex interdependencies between those levels, especially not those between the different loci of deliberation. This shows that a valid measurement of democratic deliberation at the systems level "requires more than scaling up micropolitical concepts" (Niemeyer et al., 2015, p. 5). The different *kinds* of potential loci and deliberative practices, as well as the interactive relations, need to be taken into consideration and appropriately reproduced in the aggregation rule.

Thus, the systemic approach itself provides opportunities for research on democratic deliberation while addressing the "'scaling-up' problem" (Chambers, 2012; Elstub et al., 2016, p. 140; Erman, 2016, p. 263): the systemic approach conceptualizes the "deliberative quality" of a democracy as an "emergent property" of the system as a whole. This means that it is "irreducible" to the properties of parts of the system (the quality of deliberations in individual loci). Consequently, an aggregation rule that merely adds up the deliberative performances *within* those loci without taking account of the interactions of the "individual" deliberations is insufficient. By providing a framework for the complex interactive relationships between various "deliberative activities", the systemic approach enables us "to identify which standards to employ when assessing the deliberative performance of a system as a whole" (Elstub et al., 2016, p. 140; see also Boswell, Hendriks, & Ercan, 2016; Dryzek, 2009).

There are two exemplary approaches for the measurement of deliberative performance on the basis of the systemic approach that shall serve as a starting point for our elaborations: Ercan et al. (2017, p. 197) "offer an interpretive response to…the empirical questions posed by the systemic turn". But even though they are able to identify the contribution of qualitative case studies to the study of deliberative systems, they are aware of the fact that "interpretive studies are typically limited to discrete or small-n case studies" (Ercan et al., 2017, p. 206). John S. Dryzek (2009) attempts to evaluate the "deliberative capacity" of deliberative systems by referring to criteria of authenticity[4] of the respective processes, their inclusiveness and their consequentiality. In doing so, he takes the core theoretical ideas of scholars of the systemic approach for granted and translates them into operationalizations. Thus, his approach seems to be limited to scholars who (to a large extent) accept the systemic approach's premises, which means that it is hardly compatible with the measuring approaches of numerous other scholars who base their work on other theoretical foundations.

To avoid such problems, we will use the following parts of this article to propose a *modular* approach that is compatible with different indices of democratic performance or quality. The four PMMDs and the decisions that are to be made in these steps offer a large degree of flexibility since they can be adjusted according to the theoretical and empirical focus of the researcher. Furthermore, the approach makes it possible to specifically address these challenges of scaling up the measurement of micro deliberation in conceptualization, measurement, and aggregation (Munck & Verkuilen, 2002). We are aware of the fact that an empirical specification of the systemic approach faces fundamental problems.[5] Nevertheless, we use the *conceptual framework* of the systemic approach (Boswell & Corbett, 2017; Dryzek, 2009, 2016a, 2016b; Mansbridge et al., 2012) as a point of departure for our suggestions, as it seems to be the only framework available so far that does justice to the importance of the interactive relationships between different forms of deliberation for the deliberative performance of the political system as a whole.

3. "Parameters for Measuring Democratic Deliberation" at the Systemic Level

3.1. Core Elements of the Systemic Approach

The systemic approach considers deliberation to be an "emergent property" (Niemeyer et al., 2015): the deliberative quality is a property of the system as a whole and cannot (1) be located in a specific part of the system (locus) or (2) be reduced to a mere aggregation of other qualities of the political system. In the different *loci*, deliberations of various degrees of formality take place. One of Mansbridge's (1999) original concerns was to include "everyday political talk" in the analysis of the deliberative performance of political systems (see also Conover & Searing, 2005, pp. 269f.). But this does not mean that the significance and deliberative character of formal institutions (such as parliaments or courts) should be underestimated or "downplayed" (Gaus, 2016, p. 511). To evaluate the deliberative performance of a political system, one has to consider formal *and* nonformal, institutionalized *and* non-institutionalized practices as well as "the interdependence of sites within a larger system" and their interactions (Mansbridge et al., 2012, p. 1; see also Bohman, 2012, p. 73).

The framework proposed in this article relies on three major claims of the systemic approach:

1. Deliberations in *different loci* (formal institutions, informal political talk, everyday conversations, etc.) of the political system are relevant;
2. In these loci, we will find *more or less formalized deliberative practices*: assumedly, in a federal court the standards for "good deliberative performance" and adequate ways of "taking and giving of reasons" will be much higher than in less for-

[4] "Authenticity can be understood in light of the tests just introduced (deliberation must induce reflection in noncoercive fashion, connect particular claims to more general principles, and exhibit reciprocity)" (Dryzek, 2009, p. 1382). Schouten, Leroy and Glasbergen (2012) operationalize this concept by using the criteria of the DQI (Steiner et al., 2004).

[5] Boswell and Corbett (2017, p. 4) identify the major challenges of empirical applications of the systemic approach that result from a lack of empirical specification, such as the measurement of "the deliberative effects of non-deliberative acts".

mal contexts such as the public sphere or "everyday talk" which might be considered relevant;
3. There are *interactive relationships* between these deliberations that need to be considered in the attempt to determine the overall deliberative performance of the political system: the deliberative quality of the whole democratic political system is not the same as the mere accumulation of the deliberative performances in the individual loci. Instead, the interactions between deliberations in these loci need to be taken into account, as well as the performance within the loci.

In the following, we aim to propose a measurement approach compatible with different theories of democracy and concepts of democratic deliberation by building on these core elements of the systemic approach. In order to offer a useful guideline for future research, we name the relevant conceptual decisions (section 3.2.) that need to be made in the process of conceptualization, operationalization, and aggregation of "democratic deliberation".

3.2. Parameters for the Conceptualization and Operationalization of Democratic Deliberation (PMMDs)

In this section, we suggest four parameters that need to be considered when measuring democratic deliberation at the macro level. By adjusting them, this measure can be made compatible with different conceptual frameworks and measurement approaches. To show the relevance of these conceptual decisions, we present examples to show why and to what extent the adjustment of the respective parameter makes a difference to the measurement of democratic deliberation and the results of the measurement process.

3.2.1. First Parameter: Theory of Democracy

The aim of this article is to develop suggestions for the measurement of democratic deliberations that are compatible with different measures of democracy and democratic performance. Even though the most frequently cited measurement approaches seem to measure the very same object (democracy or democratic performance), they refer to different democratic theories. This applies in particular to the way democratic institutions and their interactions are described as well as the way legitimacy is thought to be generated. Accordingly, the first conceptual question to be asked is which theory of democracy is at the heart of the measurement approach adopted. This is actually not only a preliminary question as it has rather important implications for the following steps.

Obviously, there is such an extensive range of theories of democracy that any attempt to summarize them here would be pointless.[6] Instead of offering a comprehensive account of all (possible) choices that might need to be considered regarding the theory of democracy adopted, we want to illustrate the importance of the careful adjustment of parameter one with the help of one major division between contemporary schools of democratic thought: liberal and deliberative democratic theory. What are the implications of choosing one or the other, and to what extent does an affiliation with either side impact the decisions made in the process of measuring democratic deliberation? The fundamental difference between liberal and deliberative theories' understanding of democracy is the logic of democratic legitimacy that is presupposed: while liberal theories follow the premise that an adequate representation of prepolitical or endogenous preferences is the core criterion, deliberative theories assume that preferences are exogenous and legitimacy is generated by an inclusive debate that ideally results in consensus (or at least compromise). Thus, the role ascribed to democratic deliberation differs in both models. In the case of a liberal theory, its function is to validate endogenous preferences and thereby improve the epistemic quality of decisions reached in debates in representative institutions or to justify elite decisions vis-à-vis the public. In a deliberative theory, the democratic criterion is that the addressees of law also need to be the authors of law. This is the result of an inclusive deliberative process, which ideally leads to a consensual agreement between all stakeholders.

Depending on the theoretical framework adopted, there can be path dependencies for the measurement of democratic deliberation. First, the concepts of deliberation implied by the theory are likely to differ: a liberal model might value "(good) deliberation" mostly for its epistemic benefits and define it accordingly, whereas a deliberative understanding is much more likely to presuppose a procedural concept of deliberation (see *parameter two*, section 3.2.2.). The specific definition of "good deliberation" has important implications for the evaluation of the deliberative performance of democratic political systems: the operationalization, measurement, and even the loci considered to be relevant for the total score in this dimension (see *parameter three*, section 3.2.3.) will depend on the adjustment of this first parameter.

3.2.2. Second Parameter: Concept of Deliberation

The second parameter—a decision for a clearly defined concept of deliberation—is of particular importance, as there is an intense theoretical debate on what deliberation "really" is. This theoretical debate has serious implications for empirical research on democratic deliberation, as:

> [this] lack of agreement about what constitutes deliberation makes it extremely difficult for empirical

[6] *Varieties of Democracy*, for example, considers seven theoretical approaches of democracy and derives the respective "principles" from the thinking underlying them (Coppedge et al., 2016, pp. 4–6).

researchers to address the claims of normative theory. How can one safely assert that deliberation has occurred when there are no necessary and sufficient conditions routinely applied to this concept? (Mutz, 2008, p. 526)

Obviously, we cannot offer a comprehensive discussion of all concepts of deliberation here, but we want to point out two particularly important decisions that have to be made in this step. Before that, there are two preliminary questions scholars should ask: do their general measurement approach and the theoretical understanding presupposed by this approach (see *parameter one*) imply a definite and non-ambiguous definition of deliberation? And if so, should this concept of deliberation be used in the measurement of democratic deliberation as well? If both questions are answered negatively, the scholar needs to decide upon a concept of deliberation. In this step, two questions are of particular importance: (1) *is the criterion for "high deliberative performance" an epistemic or a procedural criterion?*[7] (2) *Do I want to adopt a wide or a narrow concept of deliberation?* Both questions will be elaborated on in this section.

Question (1). While a procedural concept of deliberation would evaluate the deliberative performance of a system by the characteristics of the process(es) of deliberation (inclusiveness, fairness, etc.), an epistemic concept would evaluate the performance based on the output of this very process and the conformity of this outcome to an external criterion of "rightness" (Estlund, 2008). From a strictly theoretical perspective epistemic and procedural deliberation seem to be incompatible, therefore an explicit decision for one of them would be necessary. Nevertheless, this theoretical issue is not as pressing in a more pragmatic (empirical) approach. Deliberation in actual political systems can obviously have different functions simultaneously; it can promote the epistemic value of the decisions and the inclusion of all people affected by it, and it can, of course, be valued for both.[8] In evaluating deliberative performance, one nevertheless needs to be aware of the fact that the scores for achieving the epistemic and the procedural goal can vary independently: expert deliberation can be highly beneficial for generating a "qualified" decision by being exclusive at the same time. In this article, we do not want to argue in favor of one concept or another, but simply want to raise awareness of the fact that different "kinds" of deliberation (depending on theoretical assumptions and chosen loci) might need to be evaluated by different standards.

Question (2). There is a broad range of concepts of deliberation applying more or less rigid standards. In the original normative approaches, a narrow concept of deliberation was used: "deliberation" meant the taking and giving of reasons in the strictest sense, i.e., the exchange of rational (non-emotional), neutral, impartial arguments (Cohen, 1989; Habermas, 1996). In the confrontation with diversity theories, "we see a definite expansion of the sorts of things that could be considered arguments and reasons" (Chambers, 2003, p. 322). Partly as a result of "deep theorizing about reason", and partly as a "result of confrontations with real-world practices", a stretching of the original concept took place by taking into account that there are actually different "styles" and "cultures" of communication and reason-giving (Chambers, 2003).

Coming from the framework of the systemic approach, it seems reasonable to lower the standards for what counts as deliberation (as taking and giving of "good" reasons) in specific contexts. For example, "average citizens have few opportunities to deliberate rigorously in formal institutional settings. Most of their political discussions are therefore quite unstructured" (Conover & Searing, 2005, pp. 269f.).

If we regard different institutional settings, we also need to consider that these different deliberations should be "evaluated…by different standards" (Christiano, 2012, p. 28): an evaluation of the deliberative performance of a federal court and everyday political talk using the same standards would hardly make sense (cf. Christiano, 2012).

Although it is "a core axiom of the deliberative systemic approach: that non-deliberative practices can have positive systemic deliberative consequences, and as such should be treated as part of the system" (Dryzek, 2016a, p. 211), we object to any attempt to stretch the concept of deliberation too far. The inclusion of *non-deliberative practices* in the measurement of the overall deliberative performance of a political system has been criticized by various scholars (prominently: Owen & Smith, 2015, who suggest a *reductio ad absurdum* of this claim; see Dryzek, 2016b, p. 12).[9] Additionally, an extensive lowering of standards would miss the point of setting a *normative standard* for the evaluation of the performance and quality of democracies: "A too realistic ideal is merely an apology for the status quo" (Neblo, 2007, p. 536; see also Elstub et al., 2016, p. 146). This does not mean that it is not possible to assume "a continuum of deliberative standards for assessing the parts of the system", but only that they need "to be kept normatively robust and stringent" (Elstub et al., 2016, p. 146). Thus, on the one hand, a feasible and realistic approach (Bohman, 1998) that is compatible with the systemic approach cannot presuppose only "rational, reasonable, etc." exchanges of logically valid arguments and "good" reasons in the strictest sense. On the other hand, if the concept is stretched too

[7] In the theoretical debate, the controversy between "proceduralists" and "epistemic democrats" is one of the major cleavages (Estlund, 2008; Peter, 2007, 2013; for a discussion of this controversy see Fleuß, 2017, Chapter 3).

[8] In the systemic approach, deliberation can fulfil three functions (epistemic, ethical, democratic) in different degrees. These functions partly correspond to different but not incompatible theoretical understandings of democratic deliberation (Mansbridge et al., 2012).

[9] The validity of this *reductio* would have to be discussed in detail. Nevertheless, it highlights a problematic issue about the inclusion of *non-deliberative* practices in the evaluation of the *deliberative* performance of democracies.

far (see Steiner, 2008), there is no normative standard to compare deliberative performance with at all.

Therefore, we will stick to the claim that "deliberation" implies at least that the exchange of arguments and reasons *of some kind* occurs and we suggest that bargaining or story-telling should not be considered as "deliberative practices" (cf. Bächtiger & Wyss, 2013). Although this means that we exclude certain "communicative styles" from the concept of deliberation, there still remains a range of concepts of deliberation with a variety of scopes that might refer to a different range of phenomena, which (depending on the researcher's conceptual decision) would have to be measured.

3.2.3. Third Parameter: Loci of Deliberation

In this section, we present a systematized list of loci that is suitable for the comparative measurement of democratic deliberation. Following Conover and Searing (2005, p. 270), we assume that deliberations relevant to assess the deliberative performance of the whole system can take place[10] in three arenas of decreasing degree of formality:

(1) *Highly formal deliberations* "occur within institutions such as national courts, parliaments, and civil science departments" (Conover & Searing, 2005). These deliberations are probably most compliant with high standards of rationality;

(2) *Semi-formal deliberations* are "conversations between constituents and government officials, and conversations in political parties, interest groups, and the media" (Conover & Searing, 2005). Here, a lowering of the "rationality-standards" is probably necessary for the evaluation of the deliberative performance in these spheres;

(3) *Informal deliberations* are the "less deliberative everyday discussions among political activists, attentive publics and general publics; a form of political talk that is essential to the system's democratic character" (Conover & Searing, 2005). We expect informal deliberations to be least compliant with demanding normative standards of rationality and impartiality.

Depending on theoretical and conceptual decisions, different potential loci of deliberation will be selected and prioritized for the evaluation of the overall deliberative performance of a political system. In the selection of these loci, the parameters one and two and the choices made in these steps are relevant as well: a liberal theory, for example, will suggest a different relative weight of parliamentary deliberation than a deliberative theory and might not consider some deliberations named in category (3) to be relevant for the deliberative performance of the political system at all. Also, the selection of a wide or narrow concept of deliberation will have an impact on the selection of the loci: depending on how far one is willing to "stretch the concept", a different range of phenomena will be included in the measurement conducted on this basis.

From categories (1)–(3), we can derive a systematized list of potential loci of deliberation, which offers a much more useful framework for an empirical analysis than the enumerations given by Mansbridge et al. (2012, pp. 2, 7, 10; see also Conover & Searing, 2005; Ercan et al., 2017). This procedure also matches the *loci* to spheres in which different *kinds of deliberative* practices (which have to be evaluated by different standards) take place.[11] As we are about to demonstrate, the loci of each of the corresponding categories (1)–(3) require a different measurement approach, depending on the nature of the deliberative practices in question, and—pragmatically speaking—the accessibility of data. This is why, after the explication of each category, we will point out what has to be considered if a measurement of deliberations occurring in the respective loci is envisaged. Depending on what kind of deliberation is to be measured and what kind of data is available for the respective deliberation, different methodological approaches need to be considered.

The logical assignment of the different loci to the categories proposed by scholars has so far been somewhat vague and rarely more precise than: outlining "a spectrum of venues for deliberation, including representative assemblies, public assemblies, the public sphere, and everyday talk, and 'moving along this range entails moving along a similar range, from formal to informal'" (Elstub et al., 2016, p. 145). Not all potential loci of deliberation can be assigned to just one category: different kinds of deliberation can appear in one locus, though usually there is a tendency for certain kinds of deliberation to occur in a certain locus.

With regard to measurement approaches and operationalizations, the loci in the different categories need to be treated quite differently. This is partly determined by the availability of data on deliberations—while parliamentary deliberations are generally recorded, deliberation in less formalized loci such as marketplaces usually happens spontaneously, without audience or record. Furthermore, different kinds of deliberation occur within different frameworks in terms of timeframe, the presence of the participants, and strictness of rules. On one hand, there are the contributions of MPs to parliamentary debates which usually follow a general pattern, have a certain timeframe (according to the respective protocol), and are restricted to defined topics and a

[10] In the following, we refer to "*potential* loci of deliberation" for two reasons: First, comparative research faces the problem that depending on the institutional reification of the "deliberative system" in question, certain loci might not exist. Second, the existence of a specific *locus* does not necessarily mean that (relevant) deliberation is actually taking place in this space.

[11] Thereby we attempt to provide a framework that is more systematic and at the same time (due to its modular character) more flexible than previous research using the deliberative systemic approach.

defined type of language. On the other hand, there are online debates whose participants are generally free in their expressions concerning structure and language, as well as in their use of pictures, videos or other sign systems such as hashtags, likes and emoji. Online debates do not have limitations in terms of time or space; anyone can log in from anywhere in the world anytime. To analyze these different modes of deliberation scholars of deliberative performance need to use different methods. We will give examples for each category of deliberation in order to illustrate what measurement approaches can be used.

Loci of "highly formal deliberations" (1). As cited above, Conover and Searing (2005, p. 270) assign loci such as national courts, parliaments, and civil science departments to the level of highly formal deliberation (in their terminology: "structured deliberation"). However, we propose to include in this category only loci that are constitutionally or otherwise legally installed, that follow certain (procedural) rules while deliberating and that have the power to make collectively binding decisions.[12] Thus, we differ from Conover and Searing by excluding any locus of deliberation that does not meet those criteria (such as the civil science departments they suggest). In addition, "highly formal deliberation" is not necessarily restricted to deliberation taking place in constitutional or representational bodies: certain "democratic innovations" (such as mini publics) can be subsumed under this category if they are empowered to make collectively binding decisions (Fung, 2006). To measure the deliberative performance in loci of highly formal deliberation, scholars can use minutes and reports of the deliberations, as well as written statements or legislative proposals prepared in advance of the deliberations.

Loci of "semi-formal deliberations" (2). Conover and Searing (2005, p. 270) define semi-formal deliberations as "conversations between constituents and government officials, and conversations in political parties, interest groups, and the media". This correlates with what Habermas calls "the public sphere", which is situated around the political center and which functions as the transition sphere of political ideas and arguments to that center. Habermas (1996) also includes journals, interest groups, clubs, professional associations, academies and universities, as well as grass root initiatives. We would like to complement this list with NGO-related spaces and meetings, trade unions, and other lobby groups. Thus, this category remains quite vague and cannot be described by more specific criteria than: (1) it is *the zone where members of the political elite and members of the public sphere deliberate, or where such encounters are prepared*, and (2) there is a *certain degree of institutionalization*. Again, we differ from Conover and Searing who assign party deliberations to this category. Measuring the deliberative performance in these loci can be attempted with the help of minutes (if existent) or by interviewing insiders and experts.

Loci of "informal deliberations" (3).[13] In contrast to the other two categories, informal deliberation is not at all institutionalized (in the sense of being regulated by formalized rules), i.e. the "less deliberative everyday discussion among political activists, attentive publics and general publics" (Conover & Searing, 2005, p. 270). Loci of this kind of deliberation can be "ad hoc forums, or online spaces within which ordinary citizens, members of social movements, and civil society actors can engage in discussion and debate" (Smith, 2016, p. 154), offline and online comments in response to news items, as well as marketplaces and their culturally specific equivalents. Sources of data for measuring deliberative performance can again be interviews with insiders and experts. Furthermore, online deliberation within social media platforms and in comment feeds are especially helpful in that they enable scholars to explore informal deliberation in great detail with the help of computational text mining devices.

3.2.4. Fourth Parameter: Aggregation Rule

As previously stated (section 2), the deliberative quality of the entire democratic political system does not equal the mere accumulation of the deliberative performances in the individual loci. Rather, the interactions between deliberations in these loci also need to be taken into consideration. So far, the interactive relationships between deliberations in different loci have been addressed in case studies (Boswell et al., 2016; Ercan et al., 2017) and in various approaches comparing deliberative systems (Boswell & Corbett, 2017). However, a comprehensive approach to taking the interactive relationships between different loci systematically into account, instead of merely scaling up micro level measurement of deliberative performance, is still missing. The fourth parameter addresses questions and choices that should be considered when developing such an aggregation rule.

In line with the systemic approach, we regard two kinds of interaction to be crucial for the evaluation of the deliberative performance of political systems at the macro level: the *transmissions* between deliberative procedures taking place in the more or less formalized spheres as well as their (potential) *complementarity*. Thus, these two should be reflected in the aggregation rule. Generally, aggregation rules consist of three

[12] Depending on their power and authority, some constitutional courts fall into this category with regard to some judicial matters. Since constitutional courts are not conventionally regarded as part of the political system, we exclude them from our further discussion.

[13] One important innovation of the systemic approach was the inclusion of non-deliberative practices in the evaluation of the overall deliberative performance. In the exposition of parameter (2), we explained why *explicitly* non-deliberative practices are to be excluded in the measurement. In addition to the normative and conceptual considerations regarding this matter outlined above, there are pragmatic reasons for excluding certain communicative styles from the analysis (see Boswell & Corbett, 2017, p. 15).

kinds of element: variables, weights, and operations. In our framework, the variables describe the deliberative performances measured for the different loci (see Parameter 3). In the following, we will show that the weights can be based on the degree of transmission and that the operations depend on the relationships between the loci as well as on their complementarity.

3.3. "Transmissions" and Weighting

The aggregation rule needs to take into account that the results of deliberative processes reached in different loci and "spheres" (1–3) "must be proliferated across and among sites so that they can be challenged and 'laundered' through the system" (Boswell et al., 2016, p. 264). There have already been some attempts at capturing this "interplay" of deliberations, which is crucial for the deliberative performance of the whole system (Boswell et al., 2016). However, a systematic way that is compatible with different indices of democracy is a serious challenge.

Mansbridge et al. (2012, p. 23) do not offer an explicit definition of the interactions between loci, but rather speak (sometimes in a metaphorical way) of "coupling" (see also Hendriks, 2016a, p. 44). While the concept of coupling focuses on the relationships between the loci of the deliberative system (tight coupling of loci vs. loose coupling of loci) (Hendriks, 2016a), the concept of transmissions refers to the transfer of reasons given and results achieved in deliberations among the various loci. For feasibility reasons, we suggest that scholars use the concept of transmissions for the measurement of deliberative performance: the identification of reasons and results of reason-giving processes that might or might not be transferred to another locus seems to be much easier than the measurement of the degree of "coupling" of the loci of the respective deliberative practices.

There are three ways of tracking the transmissions of topics between different loci and assigning them a score in order to compare different democracies in terms of transmissions. Firstly, scholars could track certain topics as they evolve throughout the system (as done by Ercan et al., 2017, pp. 201-203), counting the number of loci they pass through as well as recording whether they have been present in all three categories of deliberation. Secondly, scholars could track certain individuals who potentially transmit ideas from one locus to another (cf. Mendonça, 2016) in terms of how many different loci in which categories they frequent and how many topics they pass from one to another.

However, we strongly recommend a third approach: observing certain loci with regard to where the transmitted elements (that is, deliberated ideas, reasons, resolutions, etc.) come from and where they are transmitted to (as done by Boswell et al., 2016, pp. 270-273; Hendriks, 2016a). Translated into the quantitative measurement of deliberative quality that would be: how many elements are transmitted? And how many other loci are involved in these transmissions? That approach is the most feasible for three reasons. It (1) is easily integrated into the measurement of the loci, since these loci are being assessed anyway. Thus, scholars could use elaborate methods like participant observation, but they could also take the materials they already use for assessing the deliberative quality and browse them with the help of computers for citations, expert opinions, and references to news articles, activist groups and such. Since that would only deliver a fairly accurate approximation of transmission, it should be complemented with the tracking of certain topics (cf. the first approach) in order to at least gauge the accuracy of the approximation. Furthermore, this approach would (2) be far more systematic since all loci in the study would be included and could be assessed by the same methods and it would (3) provide an approach for the weighting within an aggregation rule. The scope of transmission of each locus could be used as a weight, either in an inclusive sense for the whole system, or for each respective locus. The theoretical implication in terms of the deliberative systemic approach would be: the larger the scope of transmission, the better the deliberative quality. Furthermore, the importance of the deliberation in one locus could be assessed with that method as well: the more it is referred to (and refers to itself), the more important it becomes for the whole system—thus providing another option for a systematic aggregation rule for the measurement of macro deliberation. Consequently, for reasons of feasibility and compatibility with different democracy indices that approach should be the most suitable for most studies of deliberative performance. However, the choice of method always depends on the aim of the study and the instrument it is to be integrated within.

3.4. "Complementarity" and Operations

One fundamental assumption of the systemic approach is that "[t]hough there may be little or no perfect democratic deliberation in any site, the collective work done across the system may still produce a suitably deliberative democratic whole" (Boswell et al., 2016, p. 263). Accordingly, an aggregation rule taking this line of thinking seriously needs to take account of the complementarity of deliberation in different loci, and therefore the substitutability of the deliberation in one locus. There are several assumptions to be made about what the defining criteria for the degree to which a locus is substitutable are. Firstly, it might depend on the level of formalization of the locus. Secondly, it might depend on the (legally ascribed) political importance of that locus, which correlates with—but is not identical to—the degree of formalization. Thirdly, it might depend on its importance for the deliberative system, which can be assessed by the method recommended above—the more transmission links and transmitted elements, the higher the importance. Fourthly, it might depend on the structure of the locus. On that line of thought, Boswell and Corbett

(2017) present an approach to compare deliberative systems via "family resemblances":

> At its centre are recurring "traits" that come and go, to varying degrees, across units within the same broad family. Such traits might include institutional variants, but they tend to entail a decentred, interpretive account of these institutions—one that sees them not as given, but as constructed and continually reproduced through social interaction.

This approach can be transferred to the level of loci. "Traits" can integrate some of the criteria mentioned above (such as the level of formalization). Loci, with similar structural traits (thus belonging to one "family"), could be deemed complementary, and the more members of the respective family, the more substitutable the single locus. However, some loci might not be substitutable at all, in spite of family resemblances to other loci. Examples are deliberations in parliaments and courts. That "unsubstitutability" has to be marked as a trait as well.

Although some democracy indices' aggregation rules seem very elaborate,[14] there are two mathematical operations which all aggregation rules are based on: addition and multiplication. Those rules imply different theoretical assumptions concerning the relationship of the attributes that are to be aggregated: "If one's theory indicates that both attributes are necessary features, one could multiply both scores, and if one's theory indicates that both attributes are sufficient features, one could take the score of the highest attribute" (Munck & Verkuilen, 2002, p. 24), or, alternatively, cumulate all scores. Consequently, complementary deliberations in loci of one family could be added up to one score, while deliberations in non-complementary loci should be multiplied. In the first case, there are two options. Either, the total scores of deliberative performance in each locus are cumulated, or the scores for the chosen criteria for deliberative performance (Parameter 2) are cumulated across loci, prior to using the chosen aggregation rule for deliberative performance—thus, the deliberation within one family of loci would be treated as one truly complementary unit. In the second case, low deliberation scores in "unsubstitutable" loci would vastly lower the total score, and a zero would reduce the overall score to nil.

The choices to be made concerning the aggregation rules depend on the selection of the democratic theory on one hand, and on the understanding of deliberation on the other. For example, an index based on liberal democratic theory will probably place greater weight on highly formal deliberation (by individual weights as well as the use of multiplications) than an index based on deliberative democratic theory. Thus, the fourth parameter, again, builds upon the choices made concerning the previous parameters.

4. Conclusion: Challenges of Measuring Democratic Deliberation

In this article, we argued for the need to include the measurement of democratic deliberation into the evaluation of the democratic performance of political systems. Accordingly, we developed guidelines for a theoretically grounded measurement of deliberative performance at the macro level. Since we intended to make our suggestions compatible with different available approaches to measuring democratic performance, we proposed a modular approach. The core elements of this approach are the four PMMD that can be adjusted in various ways to fit in with the measurement approach adopted. The specific indicators which should be used to conduct the measurement have to be decided upon in accordance with the specific adjustments of these PMMDs. In the suggestions we provided concerning these parameters, we tried to do justice to the specific requirements of the measurement of deliberation at the macro level which are to a large extent based on the systemic approach (Beste, 2016; Dryzek, 2015, 2016a, 2016b; Mansbridge et al., 2012).

We are aware of the fact that this specific theoretical framework not only has its own theoretical pitfalls (Hendriks, 2016b; Owen & Smith, 2015) but that it also carries intricate methodological challenges, especially in terms of feasibility. The most complicated challenge is probably how to adequately reflect the interactive relationships of deliberations in different loci—their transmission and their complementarity—in the aggregation rule. Here, future research should further address not only the question of how transmissions or complementarity can be adequately theorized, but also how they can be measured in practice at the macro level in comparative large-n studies.

We firmly believe that it is impossible to develop a one-size-fits-all solution for this issue. Rather, the solution adopted for the integration of the measurement of deliberative performance will to a large extent be dependent on the theoretical, conceptual, and methodological "parameters" previously chosen by the researcher.

Conflict of Interests

The authors declare no conflict of interests.

References

Bächtiger, A., & Wyss, D. (2013). Empirische Deliberationsforschung—eine systematische Übersicht [Empirical deliberation—A systematic review]. *Zeitschrift für Vergleichende Politikwissenschaft*, *7*(2), 155–181.

Beste, S. (2016). Legislative frame representation: Towards an empirical account of the deliberative systems approach. *Journal of Legislative Studies*, *22*(3), 295–328.

[14] For example, V-Dem uses factor indices that are based on point estimates drawn from a Bayesian factor analysis model (cf. Coppedge et al., 2016, p. 11).

Bohman, J. (1998). Survey article: The coming of age of deliberative democracy. *Journal of Political Philosophy*, *6*(4),400–425.

Bohman, J. (2012). Representation in the deliberative system. In J. Parkinson & J. Mansbridge (Eds.), *Deliberative systems: Deliberative democracy at the large scale* (pp. 72–94). Cambridge: Cambridge University Press.

Boswell, J., & Corbett, J. (2017). Why and how to compare deliberative systems. *European Journal of Political Research*, *56*(4), 801–819.

Boswell, J., Hendriks, C. M., & Ercan, S. A. (2016). Message received? *Critical Policy Studies*, *10*(3), 263–283.

Chambers, S. (2003). Deliberative democratic theory. *Annual Review of Political Science*, *6*(1), 307–326.

Chambers, S. (2012). Deliberation and mass democracy. In J. Parkinson & J. Mansbridge (Eds.), *Deliberative systems: Deliberative democracy at the large scale* (pp. 52–71). Cambridge: Cambridge University Press.

Cohen, J. (1989). Deliberation and democratic legitimacy. In A. Hamlin & P. Petit (Eds.), *The good polity. Normative analysis of the state* (pp. 17–34). Oxford: Polity Press.

Cohen, J. (1997). Procedure and substance in deliberative democracy. In J. Bohmann & W. Regh (Eds.), *Deliberative democracy. Essays on reason and politics* (pp. 87–106). Oxford: Blackwell.

Christiano, T. (2012). Rational deliberation among experts and citizens. In J. Parkinson & J. Mansbridge (Eds.), *Deliberative systems: Deliberative democracy at the large scale* (pp. 27–51). Cambridge: Cambridge University Press.

Conover, P. J., & Searing, D. D. (2005). Studying 'everyday political talk' in the deliberative system. *Acta Politica*, *40*(3), 269–283.

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S. E., Teorell, J., Altman, D., . . . Zimmerman, B. (2016). *V-Dem codebook v6*. University of Gothenburg: Varieties of Democracy (V-Dem) Project.

Dryzek, J. S. (2009). Democratization as deliberative capacity building. *Comparative Political Studies*, *42*(11), 1379–1402.

Dryzek, J. S. (2010). *Foundations and frontiers of deliberative governance*. Oxford: Oxford University Press.

Dryzek, J. S. (2015). Deliberative engagement: The forum in the system. *Journal of Environmental Studies and Sciences*, *5*(4), 750–754.

Dryzek, J. S. (2016a). Symposium commentary: Reflections on the theory of deliberative systems. *Critical Policy Studies*, *10*(2), 209–215.

Dryzek, J. S. (2016b). The forum, the system, and the polity. *Political Theory*, *44*(1), 1–27.

Elstub, S. (2015). A genealogy of deliberative democracy. *Democratic Theory*, *2*(1), 100–117.

Elstub, S., Ercan, S., & Mendonça, R. F. (2016). Editorial introduction: The fourth generation of deliberative democracy. *Critical Policy Studies*, *10*(2), 139–151.

Ercan, S. A., Hendriks, C. M., & Boswell, J. (2017). Studying public deliberation after the systemic turn: The crucial role for interpretive research. *Policy and Politics*, *45*(2), 195–212.

Erman, E. (2016). Representation, equality, and inclusion in deliberative systems: Desiderata for a good account. *Critical Review of International Social and Political Philosophy*, *19*(3), 263–282.

Estlund, D. M. (2008). *Democratic authority. A philosophical framework*. Princeton, NJ: Princeton University Press.

Felicetti, A., Niemeyer, S., & Curato, N. (2016). Improving deliberative participation: Connecting mini-publics to deliberative systems. *European Political Science Review*, *8*(3), 427–448.

Fishkin, J. (1995). Bringing deliberation to democracy: The British experiment. *The Good Society*, *5*(3), 45–49.

Fleuß, D. (2017). *Prozeduren, Rechte, Demokratie. Das legitimatorische Potential von Verfahren für politische Systeme* [*The normative legitimacy of democracies. On the limits of proceduralism*] (Doctoral dissertation). University of Heidelberg, Heidelberg, Germany. Retrieved from http://archiv.ub.uni-heidelberg.de/volltextserver/23203/1/Fleuss-Prozeduren%20Rechte%20Demokratie-2017.pdf

Fung, A. (2006). Varieties of participation in complex governance. *Public Administration Review*, *66*, 66–75.

Gaus, D. (2016). Discourse theory's sociological claim: Reconstructing the epistemic meaning of democracy as a deliberative system. *Philosophy & Social Criticism*, *42*(6), 503–525.

Geißel, B., & Newton, K. (Eds.). (2012). *Evaluating democratic innovations*. New York, NY: Routledge.

Habermas, J. (1996). *Between facts and norms: Contributions to a discourse theory of law and democracy*. Cambridge, MA: MIT Press.

Hendriks, C. M. (2016a). Coupling citizens and elites in deliberative systems: The role of institutional design. *European Journal of Political Research*, *55*(1), 43–60.

Hendriks, C. M. (2016b). Integrated deliberation: Reconciling civil society's dual role in deliberative democracy. *Political Studies*, *54*(3), 486–508.

Kaase, M. (1984). The challenge of the 'participatory revolution' in pluralist democracies. *International Political Science Review*, *5*(3), 299–318.

Mansbridge, J. (1999). Everyday talk in the deliberative system. In S. Macedo (Ed.), *Deliberative politics. Essays on democracy and disagreement* (pp. 211–239). Oxford: Oxford University Press.

Mansbridge, J. (2007). Deliberative democracy or democratic deliberation? In S. Rosenberg (Ed.), *Deliberation, participation and democracy* (pp. 251–271). London: Palgrave Macmillan.

Mansbridge, J. (2010). Deliberative polling as the gold standard. *The Good Society*, *19*(1), 55–62.

Mansbridge, J., Bohman, J., Chambers, S., Christiano, T., Fung, A., Parkinson, J., . . . Warren, M. E. (2012). A systemic approach to deliberative democracy. In J.

Parkinson & J. Mansbridge (Eds.), *Deliberative systems* (pp. 1–26). Cambridge, MA: Cambridge University Press.

Mendonça, R. F. (2016). Mitigating systemic dangers: The role of connectivity inducers in a deliberative system. *Critical Policy Studies*, *10*(2), 171–190.

Merkel, W., Bochsler, D., Bousbah, K., Bühlmann, M., Giebler, H., Hänni, M., . . . Wessels, B. (2016). *Democracy barometer. Methodology* (Version 5). Aarau: Zentrum für Demokratie.

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

Mutz, D. C. (2008). Is deliberative democracy a falsifiable theory? *Annual Review of Political Science*, *11*(1), 521–538.

Neblo, M. A. (2007). Family disputes: Diversity in defining and measuring deliberation. *Swiss Political Science Review*, *13*(4), 527–557.

Niemeyer, S. (2014). Scaling up deliberation to mass publics: Harnessing mini-publics in a deliberative system. In K. Grönlund, A. Bächtiger, & M. Setälä (Eds.), *Deliberative mini-publics. Involving citizens in the democratic process* (pp. 177–202). Colchester: ECPR Press.

Niemeyer, S., Curato, N., & Bächtiger, A. (2015). *Assessing the deliberative capacity of democratic polities and the factors that contribute to it*. Paper presented at the conference “Democracy: A Citizen Perspective”, Turku, Finland. Retrieved from http://www.abo.fi/fakultet/media/33801/niemeyerbachtigercurato_assessingthecapacityofdeliberativesystems.pdf

O’Connor, T., & Wong, H. Y. (2015). Emergent properties. *Stanford Encyclopedia of Philosophy*. Retrieved from https://plato.stanford.edu/entries/properties-emergent/#OntEme

Owen, D., & Smith, G. (2015). Survey article: Deliberation, democracy, and the systemic turn. *Journal of Political Philosophy*, *23*(2), 213–234.

Peter, F. (2007). Democratic legitimacy and proceduralist social epistemology. *Politics, Philosophy & Economics*, *6*(3), 329–353.

Peter, F. (2013). The procedural epistemic value of deliberation. *Synthese*, *190*(7), 1253–1266.

Schouten, G., Leroy, P., & Glasbergen, P. (2012). On the deliberative capacity of private multi-stakeholder governance: The roundtables on responsible soy and sustainable palm oil. *Ecological Economics*, *83*, 42–50.

Smith, W. (2016). The boundaries of a deliberative system: The case of disruptive protest. *Critical Policy Studies*, *10*(2), 152–170.

Steiner, J. (2008). Concept stretching: The case of deliberation. *European Political Science*, *7*(2), 186–190.

Steiner, J., Bächtiger, A., Spörndli, M., & Steenbergen, M. R. (2004). *Deliberative politics in action. Analyzing parliamentary discourse*. Cambridge: Cambridge University Press.

Stevenson, H., & Dryzek, J. S. (2012). The discursive democratization of global climate governance. *Environmental Politics*, *21*(2), 189–210.

About the Authors

Dannica Fleuß is a postdoctoral researcher at Helmut-Schmidt-University, Hamburg. She completed her PhD on proceduralist theories of democracy and is currently working on applications of deliberative theories as well as the corresponding measurement approaches. Her research interests include political philosophy, measuring democracy, modern theories of democracy, and political-cultural studies.

Karoline Helbig is a doctoral researcher for Political Theory at the Helmut-Schmidt-University, Hamburg. She studied sociology and mathematics and currently is working on her PhD project regarding the influence of digital algorithms such as search engines and news feeds on democratic societies, with a particular focus on their deliberative procedures. Her further research interests include measuring democracy, methods of social research, democratic theory, digitalization, and the sociology of knowledge.

Gary S. Schaal is full professor of Political Theory at the Helmut-Schmidt-University, Hamburg. He worked on a wide variety of topics, including contemporary normative political theory, emotions in the political sphere and digital democracy.

Politics and Governance (ISSN: 2183–2463)
2018, Volume 6, Issue 1, Pages 22–32
DOI: 10.17645/pag.v6i1.1188

Article

Conceptualizing and Measuring the Quality of Democracy: The Citizens' Perspective

Dieter Fuchs [1,*] and Edeltraud Roller [2]

[1] Department of Social Science, University of Stuttgart, 70174 Stuttgart, Germany;
E-Mail: dieter.fuchs@sowi.uni-stuttgart.de
[2] Department of Political Science, Johannes Gutenberg University Mainz, 55099 Mainz, Germany;
E-Mail: roller@politik.uni-mainz.de

* Corresponding author

Submitted: 29 September 2017 | Accepted: 7 December 2017 | Published: 19 March 2018

Abstract
In recent years, several measurements of the quality of democracy have been developed (e.g. Democracy Barometer, Varieties of Democracy Project). These objective measurements focus on institutional and procedural characteristics of democracy. This article starts from the premise that in order to fully understand the quality of democracy such objective measurements have to be complemented by subjective measurements based on the perspective of citizens. The aim of the article is to conceptualize and measure the subjective quality of democracy. First, a conceptualization of the subjective quality of democracy is developed consisting of citizens' support for three normative models of democracy (electoral, liberal, and direct democracy). Second, based on the World Values Survey 2005–2007, an instrument measuring these different dimensions of the subjective quality of democracy is suggested. Third, distributions for different models of democracy are presented for some European and non-European liberal democracies. They reveal significant differences regarding the subjective quality of democracies. Fourth, the subjective quality of democracy of these countries is compared with the objective quality of democracy based on three indices (electoral democracy, liberal democracy and direct popular vote) developed by the Varieties of Democracy (V-Dem) Project. Finally, further research questions are discussed.

Keywords
democracy; measuring democracy; models of democracy; political culture; quality of democracy; social science concepts; subjective quality of democracy; varieties of democracy

Issue
This article is part of the issue "Why Choice Matters: Revisiting and Comparing Measures of Democracy", edited by Heiko Giebler (WZB Berlin Social Science Center, Germany), Saskia P. Ruth (German Institute of Global and Area Studies, Germany), and Dag Tanneberg (University of Potsdam, Germany).

© 2018 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

In recent years, several measurements of the *quality* of democracy have been developed. Focussing on political regimes classified as democracies they examine differences in the quality of these democracies. These measurements include the Democracy Barometer (Bühlmann, Merkel, Müller, & Weßels, 2012) and the Varieties of Democracy (V-Dem) Project (Coppedge, Lindberg, Skaaning, & Teorell, 2016). As these measurements focus on institutional and procedural characteristics of democracy, they can be called *objective measurements*.

This article deals with *subjective measurements* of the quality of democracy which are based on the perspective of citizens. It starts from the premise that in order to fully understand the quality of a democracy, objective measurements have to be complemented by subjective measurements at the level of the citizens. In

some cases, the above-mentioned objective measurements use such subjective indicators as well, particularly if objective indicators for theoretical constructs are missing. For example, the Democracy Barometer uses citizens' confidence in the legal system as an indicator of the quality of the legal system (Merkel et al., 2016, p. 17). However, this pragmatic strategy is questionable because objective structures and subjective evaluations of citizens constitute entirely different dimensions and can vary independently. It is possible that objective measurements of the democratic structure and processes are of high quality whereas subjective measurements of these objects are of low quality and vice versa.

The aim of this article is to propose a conceptualization and measurement of the subjective quality of democracy. Two assumptions are central. First, the subjective quality of democracy consists of citizens' support for normative conceptions of democracy. Accordingly, the more citizens support normative conceptions of democracy, the higher the subjective quality of a democracy. We want to point out that the subjective quality of democracy does not consist of citizens' evaluations of the democracy in their own country which is a quite common conception (see section 2) but instead it consists of their basic conceptions of democracy. Our conceptualization is compatible with a situation where citizens prefer democracy in general while at the same time critically evaluate the democracy in their own country. Klingemann (2014), who introduced the term "critical citizens" for such individuals, could demonstrate empirically that this type of citizen is widespread in old and new democracies and that these citizens are inclined to demand democratic reforms. Hence, the existence of critical citizens does not indicate low subjective quality of a democracy but instead is a sign of a living democracy and of high subjective quality of democracy. Second, in conceptualizing citizens' support for normative conceptions of democracy we draw on the established notion of different models of democracy and relate the subjective quality of democracy to these models. In particular, we distinguish between three well-established normative models, i.e. electoral, liberal and direct democracy, which build a hierarchy from less to more demanding models. This hierarchy is taken into account when conceptualizing the subjective quality of democracy.

The analysis proceeds in four steps. First, the state of the art on subjective quality of democracy is presented by discussing recently published studies with comparable goals but differing with respect to the central aspects of our analysis. Second, after presenting arguments for why objective measurements of the quality of democracy have to be complemented by subjective measurements at the level of citizens, a conceptualization of the subjective quality of democracy is developed including definitions of different dimensions of subjective quality of democracy. Third, based on the sixth wave of the World Values Survey (WVS) 2005–2007 an instrument measuring different dimensions of the subjective quality of democracy is suggested. Fourth, distributions for these different dimensions (normative conceptions of democracy) are presented for some European and non-European liberal democracies and compared to objective measurements of the quality of democracy based on three indices (electoral democracy, liberal democracy and direct popular vote) developed by the V-Dem Project (Coppedge et al., 2017). In the concluding chapter, further research question are discussed.

2. State of the Art

Recently, three studies have been published which are of relevance for our study on the subjective quality of democracy. Two studies claim that objective measurements of the quality of democracy have to be complemented by subjective measurements (Mayne & Geissel, 2016; Pickel, Breustedt, & Smolka, 2016) while a third study develops refined measures of citizens' attitudes towards democracy and constructs different models of democracy on this basis (Ferrín & Kriesi, 2016)

Mayne and Geissel (2016, p. 634) state the goal of their analysis concisely in the title of their article: "Putting the Demos back into the Concept of Democratic Quality". They argue that the concept of democratic quality encompasses two dimensions, namely an institutional opportunity-structure component and a citizen component. While several conceptualizations of the quality of democracy are available for the institutional dimension, conceptualizing the citizen component "which refers to the ways in which citizens can and do breathe life into existing institutions opportunities" is a neglected research topic (Mayne & Geissel, 2016, pp. 635–637). To grasp the citizen component they develop an analytic framework and introduce two theoretical dimensions. On one hand, by explicitly referring to the V-Dem Project they identify three key models of democracy: minimal-elitism, liberal-pluralism as well as participatory and deliberative democracy (Mayne & Geissel, 2016, pp. 636–639). On the other hand, they distinguish three key citizens' dispositions, namely democratic commitments, political capacities and political participation (Mayne & Geissel, 2016, p. 634). For each combination of key models of democracy and key citizens' dispositions, they suggest specific attitudes or specific modes of behaviour capturing the democratic quality of the citizen component (Mayne & Geissel, 2016, p. 641). For the minimal-elitism model, for example, they list acceptance of elected elites as sole decision makers and commitment to comply with law of the land as forms of democratic commitment.

While the proposed attitudes and modes of behaviour involve measurable constructs, the authors do not list concrete indicators from available surveys and data collections. Furthermore, a definition of the democratic quality of the citizen component is missing as well as information on how this variety of attitudes and modes of behaviour is related to the democratic quality of the citizen component.

Pickel et al. (2016, p. 646) initially praise the new indices for measuring the quality of democracy such as the Democracy Barometer as being innovative achievements. However, they also criticize these indices for relying mainly on macro indicators and neglecting the micro level of citizens which might involve a biased perspective. Pickel et al. (2016, p. 645) do not start their analysis with the premise that the inclusion of the citizen perspective is a necessity when measuring the quality of democracy; instead they ask the precedent question "why include the citizens' perspective?" According to them, it first has to be demonstrated that the inclusion of the citizens' perspective improves the measurement of the quality of government. To answer this question, they compare a measurement of the democratic quality at the macro level with a measurement at the micro level of citizens. For the macro level, they rely on Democracy Barometer data and for the micro level they use data on views and evaluations of democracy collected by the European Social Survey 2012. The measurement instrument of the European Social Survey 2012 asks for several democratic principles (e.g. protection of rights of minority groups, equal treatment by the courts) how important they are for democracy in general (views) and to what extent these principles apply in their country (evaluations). Pickel et al. (2016, p. 648) relate the individual level indicators of the European Social Survey to the macro level concepts of the Democracy Barometer. Their empirical analysis of 20 European democracies reveal similarities as well as considerable differences between the macro level data on the one hand and the views and evaluations of democracy on the other. Pickel et al. (2016, p. 653) conclude that citizens views and evaluations "provide a meaningful complementary perspective to 'objective' measures of the quality of democracy".

By demonstrating empirically that objective and subjective evaluations of the quality of democracy differ, the study of Pickel et al. (2016) provides evidence for including the subjective perspective when measuring quality of democracy. Their study does not include a definition of the subjective quality of democracy, but this was not their leading question. In addition, they do not distinguish between normative models of democracy.

The already-mentioned data of the European Social Survey 2012 were analysed in a book edited by Ferrín and Kriesi (2016). Based on the above-described item battery, the authors construct three models of democracy (liberal, social justice and direct democracy) on the level of views and on the level of evaluations. Although they construct different normative models of democracy, the authors do not address the topic of the subjective quality of democracy. The main goal of their book is to develop new concepts and measurements of political support and political legitimacy (Ferrín & Kriesi, 2016, pp. 12–13).

The following analysis distinguishes from these three contributions in several respects. First, a conceptualization of the subjective quality of democracy is developed covering definitions of the different dimensions of the subjective quality of democracy. Second, the subjective quality of democracy consists of citizens' support for normative models of democracy. It does not refer to citizens' evaluations of the democracy in their country. Third, a measurement of the subjective quality of democracy is proposed on the basis of the WVS 2005–2007 and empirical findings are presented for European as well as non-European democracies. Fourth, these results are compared to objective measurements of the quality of democracy developed by the V-Dem Project (Coppedge et al., 2017).

3. Conceptualizing the Subjective Quality of Democracy

Before conceptualizing the subjective quality of democracy, we present arguments for why it is reasonable to complement objective measurements of the quality of democracy with subjective ones and propose a general strategy for how to do this.

The starting point is the paradigm of political culture and its fundamental idea of there being a separation between an institutional structure and a political culture (Almond & Verba, 1963). Political culture is based upon the (aggregated) political attitudes of the citizens and thus has a subjective basis while the institutional structure refers to an objective level. The innovative notion of Almond and Verba is the introduction of the citizen perspective into political science and providing arguments for its significance. Its relevance is expressed in the following postulate of the political culture paradigm: a political regime is more stable, the stronger the congruency between the institutional structure and the political culture (Almond & Verba, 1963, p. 21). In subsequent research on democratic regimes, the concept of political culture has been defined more precisely. Nowadays it is a widely accepted notion among political scientists that the stability, as well as the functioning of a democracy, depends mainly on citizens' support of democracy (e.g. Diamond, 1999; Easton, 1975; Linz & Stepan, 1996; Lipset, 1981).

We start from the assumption that this basic postulate of the political culture paradigm can be expanded to the quality of democracy, i.e. support of democracy is not only of relevance for the stability and the functioning of democracy but also for the quality of democracy. Based on this premise we develop our general strategy of conceptualizing the subjective quality of democracy. In doing so we refer to the concept of support of democracy. It is very common to distinguish between at least three levels of support of democracy: commitment to democratic values and principles, support of the democratic regime of one's own country, and support of political authorities (e.g. Dalton, 2004; Fuchs, 2007; Norris, 2011). These three levels create a hierarchy; the highest level consists of the commitment to democratic values—including the value of democracy—and democratic principles. It is the most important level because

it determines the support of democracy at the lower levels, specific democratic attitudes, as well as democratic modes of behaviour. We assume that the commitment to democratic values and principles is of relevance for the subjective quality of democracy in at least in three ways. First, the unambiguous and doubtless support of democracy implies that citizens do not see any alternative to this form of government. This can be conceived as a criterion for the subjective quality of democracy. Second, citizens' basic democratic values and principles are used for evaluating the democracy of one's own country. This confrontation may result in citizens demanding reforms to improve their democracy which is regarded as a feature of a vibrant democracy and thus as a criterion for the subjective quality of democracy. Third, the general commitment to democratic values and principles influences more specific democratic attitudes and modes of behaviour. For example, it could motivate citizens to participate actively and cooperatively to both articulate their interests, as well as to engage in the political decision-making processes (Putnam, 1993). These are characteristics of a living democracy and of a high quality of democracy.

To summarize: when interested in the quality of a democracy, the perspective of citizens has to be taken into account because citizens are the ultimate sovereign of democracy. Their attitudes and behaviour depend decisively on their commitment to democratic values and principles. This is the starting point for conceptualizing and measuring the subjective quality of a democracy. The subjective quality of democracy does not refer to citizens' evaluations of the democracy of their country. The proposed subjective measurement is not intended to replace objective measurements of the quality of democracy but to complement them.

In conceptualizing the subjective quality of democracy we use the "framework for the analysis of data" developed by Munck and Verkuilen (2002, updated in Munck, 2009, p. 15) and add some ideas from the analysis of social science concepts by Goertz (2006). According to Munck (2009, p. 15), the challenge of conceptualization consists of two tasks: the "identification of attributes" used to define the concept, and the "vertical organization of attributes by level of abstraction". This vertical organization includes three levels: the highest level is the concept, the next level comprises the attributes of the concept, and the lowest level includes the components of attributes. These components of attributes are also called "leaves" and serve as a point of reference for the measurement (Munck, 2009, p. 21). As a result a so-called conceptual tree can be developed (Figure 1); it is comparable to the "three-level concepts" by Goertz (2006, pp. 50ff.). Munck (2009) illustrates his three-level model of democracy with objective criteria by referring mainly to Dahl (1989).

For our purpose, this model has to be specified for the subjective quality of democracy. Hence, the concept level consists of the subjective quality (SQ) of democracy (Figure 3). To identify the attributes we draw on a conclusive notion of the V-Dem Project. Coppedge et al. (2016) argue that the question regarding the quality of democracy depends largely on the normative standards used for evaluation. Accordingly, they distinguish several models of democracy and suggest a number of components and indicators for each model with which the quality of democracy can be assessed. The idea of the V-Dem Project is to "offer a fairly comprehensive accounting of the concept of democracy as it is employed today" (Coppedge et al., 2011, p. 253). Consequently, they use a broad list of models including electoral, liberal, participatory, deliberative, and egalitarian democracy (Coppedge et al., 2016). In the following, we pursue a more simplified conceptual approach and reduce the list to three established models of democracy: electoral, liberal and direct democracy. These models are common both in macro-level research on objective measurements of the quality of democracy and in micro-level research on the support of democracy. For example, such models have been suggested by Altman (2013), Diamond (1999) as well as Ferrín and Kriesi (2016). Using these three models implies that hybrid models of democracy such as delegative democracy (Morlino, 2009; O'Donnell, 1994) are excluded from the analysis. The advantage of focusing on these three established models, among other things,

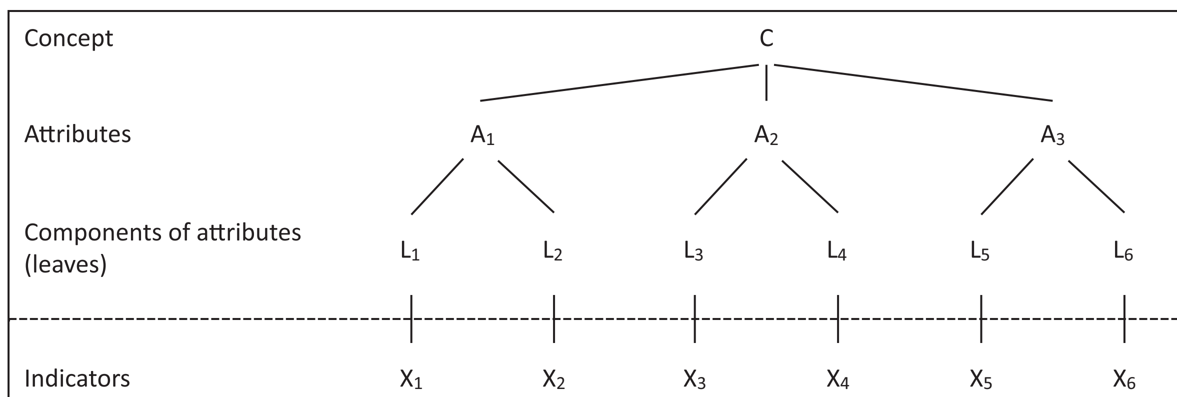


Figure 1. The logical structure of concepts (based on Munck, 2009).

Figure 2. Models of democracy (institutions, hierarchical order).

	Electoral democracy	Liberal democracy	Direct democracy
Competitive elections	1	1	1
Liberal rights	0	1	1
Direct participation	0	0	1

Notes: 1 = assignment; 0 = non-assignment.

is parsimony and the possibility of establishing a systematic relationship between subjective and objective measurements of the quality of democracy.

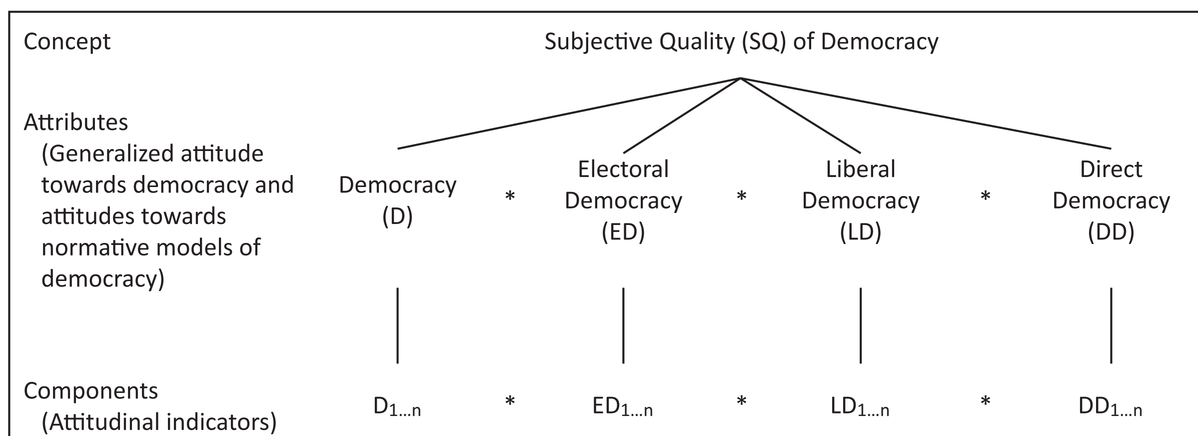
In defining these three models we rely on concise formulations of the V-Dem Project (Coppedge et al., 2016, pp. 582–583): electoral democracy "embodies the core value of making rulers responsive to citizens through competition for the approval of a broad electorate during periodic elections". Liberal democracy "embodies the intrinsic value of protecting individual and minority rights against a potential 'tyranny of the majority' and state repression more generally". Participatory democracy "embodies the values of direct rule and active participation by citizens in all political processes". These three models can be ordered along a normative hierarchy, i.e. the more demanding models include the less demanding ones (Coppedge et al., 2016; Diamond, 1999). Electoral democracy is above all defined by the institution of competitive elections; liberal democracy additionally includes liberal rights, and direct democracy covers forms of direct participation as well (Figure 2).

Attitudes towards these three models of democracy constitute the attributes of the subjective quality of democracy on the second level (Figure 3). In addition to these attitudes towards the three normative models of democracy, a basic and generalized attitude towards democracy is postulated. Only if democracy in general is clearly supported in the first instance does the question of support for different normative models of democracy arise. As a result, the subjective quality of democracy is defined by four attitudes that create the following normative hierarchy: democracy (D), electoral democracy (ED), liberal democracy (LD) and direct democracy (DD).

We started from the assumption that the first and minimal criterion of the subjective quality of democracy is support of democracy as a form of government in general. Consequently, the subjective quality (SQ) of democracy (D) could be defined as follows: *independent of the specific model of democracy institutionalized in a country, the minimal subjective quality of the country's democracy becomes higher, the more citizens unambiguously and doubtlessly support democracy as a form of government*.

To judge a country's subjective quality of democracy a theoretical criterion is needed that determines the minimum percentage of citizens who support democracy. As a rule of thumb the majority rule can be used; i.e. at least 50% of the citizens have to support democracy.

On a structural level, *electoral democracy* is defined by the institutionalization of free, competitive, fair and frequent elections; this institution unanimously makes up the indispensable core of democracy (Dahl, 1989; Diamond, 1999). On the subjective level, a corresponding citizens' attitude towards electoral democracy is support of important characteristics of the institution of elections. However, this criterion has to be supplemented by another because, as we know from autocracy research, there are many autocracies which hold seemingly democratic elections (Levitsky & Way, 2010; Schedler, 2006). Hence, even at this objective level, the institution of elections does not suffice to separate democracy from autocracy. A similar situation exists on the subjective level



Notes: * = Multiplication.

Figure 3. Conceptual structure of the subjective quality of democracy.

when citizens support important characteristics of elections without showing a clear preference for democracy. Consequently, support for elections has to be linked to support of democracy.

Such a link between two attributes is defined by Goertz (2006, pp. 50ff.) as an "AND"-relationship and by Munck (2009, p. 50) as an "interactive, noncompensatory"-relationship. As a logical expression of this "AND"-relationship Goertz (2006, p. 61) uses the multiplication sign "*". If the absence of an attribute in an "AND"-relationship is coded as 0, then a multiplication of both attributes equals also 0, thus the concept does not exist. Following Goertz, a multiplication sign between democracy (D) and electoral democracy (ED) is included in Figure 3.

On the basis of the previous argument, a second definition of the subjective quality of democracy (SQ) can be given which refers to electoral democracy (ED): *if an electoral democracy is fully institutionalized in a country, the quality of this electoral democracy becomes higher, as more citizens support democracy as a form of government and support the most important characteristics of electoral democracy*.

The first part of this definition points to a further important characteristic of our conceptualization of the subjective quality of democracy. It says that assessing the subjective quality of electoral democracy presupposes the institutionalisation of an electoral democracy in a country. The same logic applies to the following models of liberal democracy and direct democracy. This kind of reference to the institutional dimension of democracy is constitutive for the conceptualization of the subjective quality of democracy. In that way, the paradoxical coexistence of there being a high subjective quality of democracy within a non-democratic regime is excluded on the conceptual level.

Liberal democracy is a more demanding concept; it presupposes electoral democracy and complements it by the institutionalization of values which originate from the tradition of liberal thought. According to many authors (e.g. Dahl, 1989; Diamond, 1999; Merkel, 2004), a liberal democracy is the only type of democracy which sufficiently corresponds to the meaning of democracy. In order to be meaningful, elections need to be complemented by the guarantee of political rights and civil liberties. The third dimension of the subjective quality of democracy (SQ) referring to liberal democracy (LD) can be defined as follows: *if a liberal democracy is fully institutionalized in a country, then the quality of this liberal democracy is higher, the more citizens are in favour of electoral democracy while simultaneously supporting the most important characteristics of liberal democracy*. The "AND"-relationship between the subjective quality of electoral democracy and liberal democracy is marked by a multiplication sign in Figure 3.

Direct democracy is characterized by direct participation of citizens in political decisions. Historically, a pure form of direct democracy only existed within the city-state of the ancient Athens, where all important issues were decided by the people themselves, and the people literally governed themselves. In contemporary nation-states, organized as representative democracy, the model of direct democracy consists of the supplementation of representative democracy by forms of direct citizen participation such as referendums (Altman, 2011).

In defining the subjective quality of direct democracy a conceptual problem arises. To what extent do forms of direct participation have to be institutionalized so that we are able to refer to them as a new type of democracy which can be called direct democracy? Creating a threshold is difficult. However, the notion of supplementing representative democracy with forms of direct citizen participation involves the idea of a continuum with a purely representative democracy at one end and a representative democracy which incorporates some degree of direct participation at the other end. Usually, Switzerland with by far the most forms of direct participation is placed at the end of this continuum.

The ambiguity of the concept of direct democracy has also consequences for the definition of the subjective quality of this model. In contrast to electoral democracy and liberal democracy, it cannot start from the premise that direct democracy is fully institutionalized. This is why the definition of the subjective quality (SQ) of direct democracy (DD) has to take into account different degrees of direct participation: *provided that a liberal democracy has institutionalized some forms of direct participation and therefore can be understood as a direct democracy, then the subjective quality of the direct democracy is higher, the stronger citizen support for these forms of direct participation is*. Again, the "AND"-relationship between the subjective quality of liberal democracy and direct democracy is marked by a multiplication sign in Figure 3.

The construction of an "AND"-relationship between the subjective quality of the three models of democracy is the technical implementation of the assumption that the three models form a normative hierarchy, i.e. the more demanding models include the less demanding ones (cf. Figure 2). At the objective level, the V-Dem Project constructs the indices for the different models in a similar way (Coppedge et al., 2016, 2017).

The next and lowest level of the conceptual tree is made up of components of attributes. In the conceptual structure in Figure 3, these refer to the attitudinal indicators of the attributes and are addressed in the next section.

4. Measuring the Subjective Quality of Democracy

For the measurement of the subjective quality of democracy, survey data are required asking for support of democracy in general and support of the electoral, liberal, and direct model of democracy. Currently, only two comparative representative surveys including such indi-

cators are available: the European Social Survey 2012 covering only European countries and the WVS including European as well as non-European countries. In order to be able to include non-European countries as well, we draw on the sixth wave of the WVS (2005–2007). The indicators and the construction of the corresponding indices "Democracy (D)", "Electoral Democracy (ED)", "Liberal Democracy (LD)", and "Direct Democracy (DD)" are described in Table 1.

The different models of democracy represent normative models, therefore each indicator assigned to each construct constitutes a necessary feature of the respective model. In order to construct the four indices, each indicator measuring support of democracy or democratic principles is dichotomized. Respondents who support democracy in general or electoral democracy or liberal democracy or direct democracy are coded 1; those who do not support democracy or electoral democracy or liberal democracy or direct democracy are coded 0. Through the multiplication of two measurements one gets a 0 if any one of the measurements is 0; in this case, the respective democratic attitude does not exist.

For constructing the index "Democracy (D)" an indicator is available which directly asks whether a democratic political system is good or bad. Yet, as respondents may associate different things with democracy, a correct understanding of democracy which at least roughly corresponds with the theoretical definition cannot be assumed. However, the understanding can be checked using two indicators which unequivocally measure a rejection of autocracy. One indicator asks whether one considers a "strong leader who does not have to bother with parliament and elections" to be good or bad; the other asks the same for "having the army rule". Initially the three indicators, measured on a four-point scale, are dichotomized separating between good and bad. Respondents who support democracy in general are coded 1, those who do not support democracy are coded 0. In order to correct for a reasonable understanding of democracy, the multiplication rule is applied. If a respondent assesses at least one of the two autocracy items as good, 0 is assigned; the multiplication with support of democracy results in 0, which means that the respondent does not support democracy unambiguously and doubtlessly.

Table 1. Measuring the subjective quality of democracy on the basis of the WVS 2005–2007.

Quality dimension	Original Items	Measurement
Democracy (D)	• Having a democratic political system • Having a strong leader who does not have to bother with parliament and elections? • Having the army rule (1 = very good, 2 = fairly good, 3 = fairly bad, 4 = very bad)	• All three variables are dichotomized (1–2 = 1; 3–4 = 0). • Construction of index "rejection of autocracy": 1 = if "strong leader" equal 0 or "army rule" equal 0; 0 = all other logical combinations • Construction of index "Democracy (D)" = democratic system * index "rejection of autocracy" (1 = "democratic system" equal 1 * "rejection of autocracy" equal 1; 0 = all other multiplicative terms)
Electoral Democracy (ED)	• People choose their leaders in free elections • Women have the same rights as men (1 = not an essential characteristic democracy…10 = an essential of characteristic of democracy)	• Both variables are dichotomized (1–7 = 0; 8–10 = 1). • Construction of index "elections": 1 = if "free elections" equal 1 and "same rights" equal 1; 0 = all other logical combinations • Construction of index "Electoral Democracy (ED)" = index "elections" * index "Democracy (D)" (1 = "elections" equal 1 * index "Democracy (D)" equal 1; 0 = all other multiplicative terms)
Liberal Democracy (LD)	• Civil rights protect people's liberty against oppression. (1 = not an essential characteristic of democracy…10 = an essential characteristic of democracy)	• Variable "civil rights" is dichotomized (1–7 = 0; 8–10 = 1). • Construction of index "Liberal Democracy (LD)" = variable "civil rights" * index "Electoral Democracy (ED)" (1 = "civil rights" equal 1 * index "Electoral Democracy (ED)" equal 1; 0 = all other multiplicative terms)
Direct Democracy (DD)	• People can change the laws in referendums. (1 = not an essential characteristic of democracy…10 = an essential characteristic of democracy)	• Variable "referendums" is dichotomized (1–7 = 0; 8–10 = 1). • Construction of index "Direct Democracy (DD)" = variable "referendums" * index "Liberal Democracy (LD)" (1 = "referendums" equal 1 * index "Liberal Democracy (LD)" equal 1; 0 = all other multiplicative terms)

Notes: * = Multiplication.

For measuring the support of electoral, liberal, and direct democracy an item battery covering several principles of democracy (e.g. people choose their leaders in free elections) is used. The respondents are asked for each whether it is an essential characteristic of democracy or not. The scale ranges from 1 (= not an essential characteristic of democracy) to 10 (= an essential characteristic of democracy). In order to measure clear support of a democratic principle and to avoid a soft middle category, the scale is dichotomized as follows: The values 8–10 are coded 1 and the remaining values are coded 0. Of course, other modes of dichotomizing would produce different results.

Support of electoral democracy is measured by two indicators: "people choose their leaders in free elections" and "women have the same rights as men". They refer to two basic principles of electoral democracy, namely free elections and political equality. The second indicator "women have the same rights as men" is ambiguous because the protection of women's rights can also be regarded as a criterion for a liberal democracy. But since, according to Dahl (1989), elections are not sufficiently defined without the criterion of equality, this indicator is used as a proxy for the equality criterion. Only if both of these principles, i.e. free elections and political equality, are considered to be essential characteristics of democracy, can reasonable support of electoral democracy be assumed. In order to construct the index "Electoral democracy (ED)" the multiplication rule with "Democracy (D)" is applied.

The indicator measuring liberal democracy represents an appropriate operationalization of a basic principle of negative liberties: "civil rights protect people's liberty against oppression". Finally, the indicator for direct democracy focuses on a central form of direct participation of the citizens: "People can change the laws in referendums". Referendums are the dominant form in which direct democracy can be realized in modern democracies. For constructing the respective indices, "Liberal Democracy (LD)" and "Direct Democracy (DD)", the values are multiplied by the preceding models.

As the item battery of the WVS asks whether each element is an essential characteristic of democracy it might be criticized for measuring primarily a cognitive and not the conceptually demanded evaluative dimension. In the first instance, we would argue that the criterion "essential" includes more of an evaluative component than the criterion "important" used by the comparable battery of the European Social Survey 2012. At any rate, it is plausible to assume that the specific characteristics of democracy gathered by the WVS clearly has an evaluative meaning, if democracy is unambiguously preferred as a form of government (see multiplication with "Democracy (D)" in Table 1).

5. Empirical Results

To demonstrate the applicability of the proposed measure of the subjective quality of democracy and to compare it with an objective measure a sample of five European and five non-European liberal democracies was selected (France, Germany, Norway, Sweden, Switzerland vs. Brazil, India, South Korea, Taiwan, USA). Liberal democracies were identified on the basis of the liberal democracy index of V-Dem (Coppedge et al., 2017). The scale of this index ranging from 0 to 1; we define that values greater than 0.5 indicate the existence of a liberal democracy. This definition is validated by the democracy index of Freedom House (2017); all selected countries are classified as liberal democracies by this index, too.

Table 2 includes the distribution of the four indices measuring the subjective quality of democracy (columns "subjective") as well as data for the objective quality of democracy for each of the three normative models (columns "objective"). The objective data was provided by V-Dem (Coppedge et al., 2017). The values for the three indices—electoral democracy index, liberal democracy index, direct popular vote index—vary from 0 to 1.

Table 2. Subjective and objective quality of democracy in selected liberal democracies.

	Democracy (D)	Electoral Democracy (ED)		Liberal Democracy (LD)		Direct Democracy (DD)	
Country	**Subjective**	**Subjective**	**Objective [1]**	**Subjective**	**Objective [2]**	**Subjective**	**Objective [3]**
Norway	81.7	74.2	0.92	57.5	0.89	44.0	0.3
Switzerland	79.8	70.8	0.91	64.8	0.87	62.2	0.68
Sweden	79.7	79.8	0.91	75.0	0.88	57.8	0.15
Germany	78.4	67.7	0.91	63.0	0.88	54.7	0.01
USA	59.1	49.9	0.87	44.3	0.81	34.4	0
France	57.6	43.3	0.91	37.1	0.83	28.9	0
South Korea	39.3	27.3	0.84	21.8	0.76	17.0	0.03
Taiwan	37.0	30.6	0.77	27.6	0.68	22.8	0.21
Brazil	27.3	20.7	0.89	15.5	0.78	13.7	0.2
India	27.3	19.4	0.73	16.9	0.58	12.1	0

Notes: [1] Electoral democracy index (V-Dem, 2005); [2] Liberal democracy index (V-Dem, 2005), [3] Direct popular vote index (V-Dem, 2005). Database: WVS 2005–2007.

The higher the value, the more the democracy of each country conforms to the normative model and the higher the objective quality of that democracy. We start with the description of the subjective data and then move on to the objective data.

As the subjective democracy indices are constructed in such a way that the more demanding models presuppose the less demanding ones, the level of subjective quality of democracy successively decreases from "Democracy (D)" to "Direct Democracy (DD)". In Table 2, the countries are ordered according to the level of support for "Democracy (D)".

The most striking feature of the distribution of the subjective quality of democracy is the significant differences between the countries. This is already true for the basic support of democracy as a form of government, i.e. "Democracy (D)". The minimal subjective quality of democracy is the highest in Norway (81.7%) and the lowest in India and Brazil (27.3%). Notably, the values of all established Western countries, i.e. the five European countries and the USA, are above the 50-percent-level, while the values of the remaining non-European countries are below.

The pattern is similar for "Electoral Democracy (ED)", where Norway shows the highest (74.2%) and India (19.4%) the lowest degree of subjective quality. Even though all countries are classified as liberal democracies, the data for the subjective quality of "Liberal Democracy (LD)" reveal considerable differences. Only four European countries (Norway, Switzerland, Sweden, and Germany) exceed the 50-percent-level and with the exception of the USA, all non-European countries reveal support levels below 30%.

Country differences are also large for the subjective quality of "Direct Democracy (DD)". The lowest figures can be found in India (12.1%) and the highest level of subjective quality of direct democracy is—as expected—in Switzerland (62.2%). Interestingly, the values for Sweden and Germany are also above the 50-percent-level. Overall, the established Western democracies show higher values than the other non-Western countries and the differences within the first group of countries are much higher than in the second group. While in the established Western democracies values range from 28.9% (France) to 62.2% (Switzerland), support for direct democracy in the other non-Western countries varies only from 12.1% (India) to 22.8% (Taiwan). Obviously, it is difficult to interpret the pattern for the subjective quality of direct democracy without objective data on the degree of institutionalization of direct democracy.

As an initial conclusion, we can state that the data reveal significant differences regarding the subjective quality of all three models of democracy. The lowest levels exist for non-Western democracies. Applying less strict criteria in dichotomizing the indicators would indeed result in higher percentages for the subjective quality of democracy but the substantial differences between the countries as well as the ranking of countries would remain.

What about the objective quality of democracy in the selected European and non-European countries? For electoral and liberal democracy, the differences between the countries at the objective level are considerably lower compared to the subjective level. For electoral democracy, only the figures of Taiwan and India are below 0.80. The liberal democracy index reveals two thresholds, one which separates Western from non-Western countries (above and below 0.80, respectively) and the other within the non-Western countries where India clearly differs from the rest (below 0.68).

In order to measure the objective quality of direct democracy we do not draw on the very broadly defined participatory democracy index of V-Dem; instead, the direct popular vote index is used (Coppedge et al., 2017, pp. 52–53, 62–63). It corresponds to the subjective indicator asking for approval of referendums. The most striking result for this objective measurement is that only Switzerland has a notable value (0.68). This reflects the fact that in other countries direct democracy or referendums are not or are hardly institutionalized. However, it is noteworthy that despite the low degree of institutionalization of direct democracy, the subjective quality score is between 44 and 57.8% in several countries (Norway, Sweden, and Germany). This can be interpreted as citizens' demand for the introduction of referendums into the institutional setting of the liberal democracy of their country.

The most important result of the comparison between objective and subjective measures of the quality of democracy is the following: differences between countries at the objective level are considerably lower than at the subjective level. This implies different things for the various models of democracy. First, with respect to electoral and liberal democracy, this can mean, among other things, that even a successful institutionalization of an electoral or liberal democracy in a country does not, at the same time, ensure a corresponding level of citizens' commitment to democratic values and principles. Such a configuration is already known from democratization research, where democratic consolidation encompasses not only the existence of democratic institutions but also support of democracy by the main political actors (Diamond, 1999; Linz & Stepan, 1996). A commitment to democratic values and principles depends on different historical traditions and different cultural contexts (Fuchs & Roller, 2016), it cannot be directly generated by implementing democratic institutions. Second, with respect to direct democracy, a reverse pattern exists between the subjective and objective quality of democracy. The discrepancy between both dimensions can mean that citizens demand that the institutional setting of liberal democracy in their country is complemented with instruments of direct democracy.

6. Conclusions

The aim of the article is to contribute to the current discussion on assessing the quality of democracy in several

ways. First, by arguing that a measurement of the subjective quality of democracy is reasonable. We start from the assumption that the basic postulate of the concept of political culture, according to which support of democracy is of relevance for the stability and functioning of a democracy, can be expanded to the subjective quality of democracy as well. Second, by conceptualizing the subjective quality of democracy as support for different models of democracy (distinguishing between three established normative models, i.e. electoral, liberal and direct democracy) and not as citizens' evaluation of democracy in their country. Accordingly, the more citizens support normative conceptions of democracy, the higher the subjective quality of a democracy. Third, by developing an instrument measuring the suggested dimensions of the quality of democracy for European and non-European democracies. Fourth, by demonstrating the applicability of this instrument and providing initial empirical results. Fifth, by comparing these results with objective measurements provided by V-Dem. The comparison of the subjective with the objective quality of democracy for the electoral, liberal and direct models of democracy shows that at the objective level the differences between the countries are significantly lower than those at the subjective level. Hence, the subjective perspective of citizens is not fully determined by the objective institutions and processes. For the subjective perspective, historical traditions and cultural contexts play a crucial role.

The proposed measurement for assessing the subjective quality of democracy is preliminary, it is not claimed to be the final solution. For future research, at least three desiderata remain. First, the validity of our measurement instrument should be examined using multiple indicators for each model of democracy. This could be done, at least for European countries, on the basis of the European Social Survey 2012 which includes several indicators for the different models of democracy. Second, supplementing the subjective quality of democracy by including support for normative conceptions of democratic community, i.e. the relationship between citizens such as tolerance toward other groups and trust in unknown others. Empirically, it could be demonstrated that the largest differences between Western and non-Western countries exist for these normative conceptions of democratic community (Fuchs & Roller, 2016). Third, a measurement instrument could be developed to combine the objective and subjective dimensions of quality of democracy.

Conflict of Interests

The authors declare no conflict of interests.

References

Almond, G. A., & Verba, S. (1963). *The civic culture*. Princeton, NJ: Princeton University Press.

Altman, D. (2011). *Direct democracy worldwide*. Cambridge: Cambridge University Press.

Altman, D. (2013). Bringing direct democracy back in: Toward a three-dimensional measure of democracy. *Democratization*, *20*(4), 615–641.

Bühlmann, M., Merkel, W., Müller, L., & Weßels, B. (2012). The democracy barometer: A new instrument to measure the quality of democracy and its potential for comparative research. *European Political Science*, *11*(4), 519–536.

Coppedge, M., Gerring, J., Altman, D., Bernhard, M., Fish, S., Hicken, A., . . . Teorell, J. (2011). Conceptualizing and measuring democracy: A new approach. *Perspectives on Politics*, *9*(2), 247–267.

Coppedge, M., Gerring, J., Lindberg, S., Skaaning, S.-E., Altman, D., Bernhard, M., . . . Staton, J. (2017). *V-Dem Codebook v7.1*. Retrieved from https://www.v-dem.net/media/filer_public/84/a8/84a880ae-e0ca-4ad3-aff8-556abfdaff70/v-dem_codebook_v71.pdf

Coppedge, M., Lindberg, S., Skaaning, S.-E., & Teorell, J. (2016). Measuring high level democratic principles using the V-Dem Data. *International Political Science Review*, *37*(5), 580–593.

Dahl, R. A. (1989). *Democracy and its critics*. New Haven, CT: Yale University Press.

Dalton, R. J. (2004). *Democratic challenges, democratic choices*. Oxford: Oxford University Press.

Diamond, L. (1999). *Developing democracy*. Baltimore, MD: Johns Hopkins University Press.

Easton, D. (1975). A re-assessment of the concept of political support. *British Journal of Political Science*, *5*(4), 435–457.

Ferrín, M., & Kriesi, H. (Eds.). (2016). *How Europeans view and evaluate democracy*. Oxford: Oxford University Press.

Freedom House. (2017). Freedom in the world. *Freedom House*. Retrieved from http://www.freedomhouse.org

Fuchs, D. (2007). The political culture paradigm. In R. J. Dalton & H.-D. Klingemann (Eds.), *The Oxford handbook of political behaviour* (pp. 161–184). Oxford: Oxford University Press.

Fuchs, D., & Roller, E. (2016). Demokratiekonzeptionen der Bürger und demokratische Gemeinschaftsorientierungen [Citizens' conceptions of democracy and democratic community]. In S. Schubert & A. Weiss (Eds.), *"Demokratie" jenseits des Westens* [*"Democracy" beyond the West*] (pp. 296–317). Baden-Baden: Nomos.

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

Klingemann, H.-D. (2014). Dissatisfied democrats. In R. J. Dalton & C. Welzel (Eds.), *The civic culture transformed* (pp. 116–157). Cambridge: Cambridge University Press.

Levitsky, S., & Way, L. A. (2010). *Competitive authoritarianism*. Cambridge: Cambridge University Press.

Linz, J. J., & Stepan, A. (1996). *Problems of democratic transition and consolidation*. Baltimore, MD: Johns Hopkins University.

Lipset, S. M. (1981). *Political man*. Baltimore, MD: The Johns Hopkins University Press.
Mayne, Q., & Geissel, B. (2016). Putting the demos back into the concept of democratic quality. *International Political Science Review*, *37*(5), 634–644.
Merkel, W. (2004). Embedded and defective democracies. *Democratization*, *11*(5), 33–58.
Merkel, W., Bochsler, D., Bousbah, K., Bühlmann, M., Giebler, H., Hänni, M., . . . Wessels, B. (2016). *Democracy barometer. Codebook. Version 5*. Aarau: Zentrum für Demokratie.
Morlino, L. (2009). Are there hybrid regimes? Or are they just an optical illusion? *European Political Science Review*, *1*(2), 273–296.
Munck, G. L. (2009). *Measuring democracy*. Baltimore, MD: The Johns Hopkins University Press.
Munck, G. L., & Verkuilen, J. (2002). Conceptualizing and measuring democracy: Evaluating alternative indices. *Comparative Political Studies*, *35*(1), 5–34.
Norris, P. (2011). *Democratic deficit*. Cambridge: Cambridge University Press.
O'Donnell, G. A. (1994). Delegative democracy. *Journal of Democracy*, *5*(1), 55–69.
Pickel, S., Breustedt, W., & Smolka, T. (2016). Measuring the quality of democracy: Why include the citizens' perspective? *International Political Science Review*, *37*(5), 645–655.
Putnam, R. (1993). *Making democracy work*. Princeton, NJ: Princeton University Press.
Schedler, A. (Eds.). (2006). *Electoral authoritarianism*. Boulder, CO: Lynne Rienner.

About the Authors

Dieter Fuchs is Professor Emeritus of Political Science at the University of Stuttgart. His research mainly focusses on democratic theory, comparative analysis of political cultures, and European integration.

Edeltraud Roller is Professor of Political Science at the Johannes Gutenberg University Mainz. Her research mainly focusses on welfare state cultures, the performance of democracies, and political support within old and new democracies.

COGITATIO

Politics and Governance (ISSN: 2183–2463)
2018, Volume 6, Issue 1, Pages 33–47
DOI: 10.17645/pag.v6i1.1216

Article

Don't Good Democracies Need "Good" Citizens? Citizen Dispositions and the Study of Democratic Quality

Quinton Mayne [1,*] and Brigitte Geißel [2]

[1] Kennedy School of Government, Harvard University, Cambridge, MA 02138, USA; E-Mail: quinton_mayne@harvard.edu
[2] Institute of Political Science, Goethe University, 60629 Frankfurt am Main, Germany; E-Mail: geissel@soz.uni-frankfurt.de

* Corresponding author

Submitted: 5 October 2017 | Accepted: 29 January 2018 | Published: 19 March 2018

Abstract
This article advances the argument that quality of democracy depends not only on the performance of democratic institutions but also on the dispositions of citizens. We make three contributions to the study of democratic quality. First, we develop a fine-grained, structured conceptualization of the three core dispositions (democratic commitments, political capacities, and political participation) that make up the citizen component of democratic quality. Second, we provide a more precise account of the notion of inter-component congruence or "fit" between the institutional and citizen components of democratic quality, distinguishing between static and dynamic forms of congruence. Third, drawing on cross-national data, we show the importance of taking levels of inter-dispositional consistency into account when measuring democratic quality.

Keywords
citizens; democracy; democratic commitments; political capacity; political participation; quality of democracy

Issue
This article is part of the issue "Why Choice Matters: Revisiting and Comparing Measures of Democracy", edited by Heiko Giebler (WZB Berlin Social Science Center, Germany), Saskia P. Ruth (German Institute of Global and Area Studies, Germany), and Dag Tanneberg (University of Potsdam, Germany).

© 2018 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

> The health and stability of a modern democracy depends, not only on the justice of its "basic structure" but also on the qualities and attitudes of its citizens.
> (Kymlicka & Norman, 1994, p. 352)

1. Introduction

Large-scale, cross-national indices of democratic quality have traditionally paid little systematic attention to citizens as a constitutive component of democratic quality. In earlier work (Mayne & Geissel, 2016), we challenged this orthodoxy by highlighting the importance of citizens as central to the conceptualization of democratic quality. Specifically, we argued that democratic quality consists of two necessary, but independently insufficient, components. First, an "institutional component", which dominates research on democratic quality, refers to the institutional and structural opportunities that allow for democratic rule. Second, a "citizen component" relates to the ways in which citizens can and do breathe life into existing institutional opportunities for democratic rule.[1] We identified three broad categories of citizen dispositions as constitutive of the citizen component: namely, democratic commitments, political capacity, and political participation.

Providing a structured account of how citizens lie at the core of the concept of democratic quality is important for both scholarly and practical reasons. By includ-

[1] This should not be confused with quality-of-democracy research that takes citizens into account using data on mass public assessments of political actors and institutions (see, e.g., Logan & Mattes, 2012; Pickel, Breustedt, & Smolka, 2016). These kinds of public opinion data are unrelated to what we refer to here as the citizen component of democratic quality; instead, they provide (non-expert evaluative) information on the institutional component of democratic quality.

ing measures of electoral turnout, many existing studies in this area partially or implicitly acknowledge that high-quality institutions are not enough; a high-quality democracy also requires citizens to use these institutions. Participation in elections is, however, only one part of the citizen component of democratic quality. To advance theoretical and empirical research in this area, it is necessary to develop a systematic understanding of the place of citizens within the concept of democratic quality. That is the goal of this article. To this end, we expand the empirical underpinnings of the study of democratic quality by bringing it into conversation with research from political behavior and political psychology. We also deepen the engagement of quality-of-democracy research with political theory and political philosophy. Existing work on democratic quality is often anchored in partial readings of normative accounts of democracy, focused on extracting what these accounts have to say about institutions. As we show here, these same accounts have a great deal to say about the kinds of citizen dispositions that are constitutive of a high-quality democracy.

Taking citizens more seriously in how we understand democratic quality also brings research in line with the realities of national and international programs aimed at supporting and deepening democracy. Publicly-funded and philanthropic work has long had both an institutional *and* a citizen component, seeking to improve the quality of institutions as well as impact the values, competences, and participatory proclivities of citizens. Traditionally this work has been targeted at new, low- and middle-income democracies, but—amidst growing fears of democratic deconsolidation (Foa & Mounk, 2017; Levitsky & Ziblatt, 2018)—there has been an upsurge in interest in democratic programming in advanced industrial societies. By fully incorporating citizens into the conceptualization of democratic quality, we hope to bridge the existing gap between practice and research, enabling work on quality of democracy to speak to efforts of leaders and organizations working in the space of democracy promotion.

In the pages that follow, we aim to provide a solid analytic foundation and conceptual framework to incorporate data on the citizen component of democratic quality in future empirical research. We do this by building on our earlier work in three ways. In the next section, we provide a more fine-grained and structured conceptualization of democratic commitments, political capacities, and political participation. In the second section of the article, we address the question of congruence or “fit” between the institutional and citizen components of democratic quality. Third, we develop the idea that the degree of consistency of democratic commitments, political capacities, and political participation with the same model of democracy is an important aspect of democratic quality. We illustrate the issues of inter-component congruence and inter-dispositional consistency using available cross-national empirical data. The article ends with a discussion of the significant limitations of existing international survey programs as sources of data for measuring the citizen component of democratic quality.

2. Citizen Dispositions

Just as the Varieties of Democracy project (Coppedge et al., 2011) has shown that different models of democracy value different kinds of institutional arrangements, it is important to recognize that there is no one-size-fits-all understanding of the core dispositions that comprise the citizen component of democratic quality. We therefore focus our attention on how three models of democracy, which have long dominated academic and policy debates, understand each disposition. This includes: minimal-elitism—epitomized by the work of Joseph Schumpeter (1950) and E. E. Schattschneider (1975); liberal-pluralism, defined and developed perhaps most famously in the work of Robert Dahl (1971, 1989); and participatory democracy, championed by scholars such as Carole Pateman (1970) and Benjamin Barber (1984).[2]

2.1. Democratic Commitments

That democratic commitments are a necessary component of democratic quality finds support in a long line of writing. As John Stuart Mill (1861/2009, p. 7) noted, “the people for whom the [democratic] form of government is intended must be willing to accept it”. Democratic commitments refer to the political beliefs, values, principles, and norms that citizens hold dear. They combine both cognitive and affective orientations, which citizens use to understand and judge the political world. A sizeable body of empirical research has emerged in recent years on how citizens understand democracy (Bratton, Mattes, & Gyimah-Boadi, 2005, Chapter 3; Canache, 2012; Carrión, 2008; Dalton, Sin, & Jou, 2007; Fuchs & Roller, 2006; Kornberg & Clarke, 1994; Miller, Hesli, & Reisinger, 1997; Silveira & Heinrich, 2017; Thomassen, 1995), but the literature on the more specific question of which democratic values and principles citizens actually endorse is still fairly limited (Carlin, 2017; Carlin & Singer, 2011; Hibbing & Theiss-Morse, 2002; Kriesi, Saris, & Moncagatta, 2016; Lalljee, Evans, Sarawgi, & Voltmer, 2013; McClosky, 1964; Schedler & Sarsfield, 2007).

The concept of democratic commitment operates at two levels: at a general level in the form of citizens’ broad preference for democracy over non-democratic forms

[2] In this article we are argue that citizens are constitutive of the concept of democratic quality; we are silent on the question of whether the concept of democracy itself includes a citizen component. We note, however, that the answer to this question has been a resounding no. Scholarship has predominantly distinguished democracies from non-democracies (or hybrid regimes) in one of two ways: based exclusively on electoral procedures (e.g., Przeworksi, Alvarez, Cheibub, & Limongi, 2000); or using a more expansive set of procedural criteria, that take account of not just the quality of electoral processes but also the protection of civil liberties and civilian control of the military, among other things (e.g., Mainwaring, Brinks, & Pérez-Liñán, 2007). The first approach distinguishes democracies based on institutions that lie at the heart of the minimal-elitist conception of democracy; the second approach is anchored more in the liberal-pluralist account of democracy. In both cases, selection criteria are essentially institutional.

of political organization; and at a more specific level in terms of citizens' support for particular principles and values. The idea that democracies require citizens' general democratic commitment finds clear support in work on democratic consolidation as well as democratic deconsolidation.[3] Building on this research, we understand democratic quality as being in part a function of how committed citizens are to democracy, even in the face of mobilization by anti-democratic forces, economic misfortune, and electoral losses.

The study of democratic quality requires going beyond this general commitment to democracy and taking into account citizens' commitments to more specific democratic values. Pragmatically, this makes it easier to identify whether citizens' general democratic commitment is in fact nominal and without meaningful content; it is also necessary because different models of democracy set store by different types of political values. A theory-driven approach requires being clear on how the model(s) of democracy underpinning one's assessment interpret core democratic principles in different ways, or even accommodate different democratic principles. To gain analytic purchase on the issue of democratic commitments, we propose that scholars focus on the principled responses that different models of democracy provide to the following two questions. First, *who* gets to decide? Second, *how* are decisions to be made?[4]

The question of *who* gets to decide is first and foremost about what citizens consider to be the proper role of elected politicians in democratic decision making. A helpful way of thinking about this issue is in terms of the checks and balances that different models of democracy expect citizens to support, which concerns the power of elected politicians relative to other "political" actors, e.g., the judiciary or subnational authorities. It also concerns checking and balancing among different classes of politician, most notably between the executive and legislature. Finally, the question of who gets to decide is crucially linked to what citizens see as their own role, acting individually or collectively, in democratic decision making.

The second question of how decisions are to be made relates to citizens' settled opinions on how core democratic principles should be instantiated in democratic processes and structures. This fundamentally concerns not just the formal rules but also the institutionalized norms of encounter and exchange between elected politicians and other political actors. Different models of democracy demand, explicitly and implicitly, different commitments from citizens when it comes to how democratic decision-making processes should take place. As a result, judgements of any one country's democratic quality will vary greatly depending on the model used to carry out the assessment. This becomes clear by looking at the democratic commitments expected of citizens by the three key models of democracy (a summary of which is available in Table 1).

The minimal-elitist account of democracy envisages citizens to be committed to forms of decision making dominated by parties, elected politicians, and the government of the day, with few checks and balances. Citizens are expected to willingly accept their own voluntary "retirement" (to borrow the words of Schumpeter, 1950, p. 295) from political life between elections. As to the question of how decisions are to be made, high-quality minimal-elitist democracy is predicated on the expectation that citizens will be tolerant of political differences and supportive of robust competition between those differences at the ballot box. However, once votes are cast, minimal-elitism expects citizens to support winner-take-all majoritarianism, which necessarily implies electoral losers (even perennial electoral losers) accepting their political marginality.

High-quality liberal-pluralist democracies are also home to citizenries that support elected politicians as the primary decision-makers. However, "good" liberal-pluralist citizens are additionally expected to be committed to the idea that politicians are checked and balanced in important ways, for example by constitutional protections and judicial oversight, or by divisions of power between the executive and legislature, i.e., decision making that involves elected *and* unelected elites. Citizens are expected to embrace their own role in democratic decision making as largely mediated: by the parties/politicians they elect, and by the interest organizations who speak on their behalf. Liberal-pluralists expect citizens to accept or even welcome that public policy will be influenced by processes of consultation and lobbying, involving politically independent intermediary organizations and associations. By extension, the "good" liberal-pluralist citizen is expected to see negotiation and compromise among elites of different political persuasions as a natural and proper part of the democratic process.[5]

The participatory model of democracy is distinct from minimal-elitism and liberal-pluralism in that it expects citizens to support unmediated forms of mass popular involvement in democratic decision making. This might include support for direct democratic mechanisms that allow citizens to vote on specific issues as well as participatory innovations (such as participatory budgeting and citizen juries) that give citizens decision-making powers. While the participatory model of democracy clearly sets great store by the idea that final decision-making powers should lie with citizens themselves in at

[3] As Juan Linz and Alfred Stepan note, "a democratic regime is consolidated when a strong majority of public opinion, even in the midst of major economic problems and deep dissatisfaction with incumbents, holds the belief that democratic procedures and institutions are the most appropriate way to govern collective life" (Linz & Stepan, 1996, p. 16; see also Diamond, 1999, p. 69; Foa & Mounk, 2017).

[4] These two questions effectively amount to two sides of the same coin of democratic decision making, and as such are likely difficult to separate empirically.

[5] Significant variations exist within the liberal-pluralist understanding of democracy, which encompasses classic forms of pluralistic decision making as well as consensus or negotiation democracy (Lijphart, 1999).

Table 1. The citizen component.

Core Dispositions	Key Elements	The "good" citizen according to:		
		Minimal-elitist model	*Liberal-pluralist model*	*Participatory model*
1. Democratic commitments	Commitment regarding: • Who gets to decide? • How decisions should be made?	Committed to decision making dominated by parties and elected politicians, with few checks and balances.	Committed to electoral democracy where politicians are checked and balanced and intermediary organizations play important role.	Committed to unmediated forms of mass popular involvement in democratic decision making and idea that politicians should actively consult citizens between elections.
2. Political capacities	Capacity to: • know • choose • influence	Capable of *selecting into* their values, preferences, and interests based on menu of options provided to them by political elites in lead up to elections.	Capable of enlightened understanding of their own interests and sufficiently tuned into politics to be able to identify and support, if need be, organizations that can defend their values and interests.	Possessing skills and knowledge that enable them to cooperate, communicate, and deliberate with fellow citizens and political elites.
3. Political participation	Participation that is: • Electoral vs. non-electoral • Mediated vs. direct • Other-regarding	Pay sufficient attention to politics during election campaign to avoid being duped and turn out to vote, if interests at stake.	No duty to participate actively in politics, but ideally occasionally undertakes mainly mediated forms of participation.	Directly and actively involved in politics on an ongoing basis, with emphasis on other-regarding and public-oriented political activities.

least certain issue or policy areas, it also demands that where elected politicians retain decision-making powers they should undertake continuous processes of consultation with citizens between elections. This is one of the chief differences between participatory democracy and minimal-elitism and liberal-pluralism when it comes to the question of "how" decisions should be made.

2.2. Political Capacity

Existing cross-national indices of democratic quality rarely include indicators capturing levels of political capacity among citizenries.[6] This contrasts with statements on the importance of political capacity made by democratic theorists of various stripes as well as the growing fears expressed by political commentators of citizens' incapacity to resist misinformation. The absence of direct measures of political capacity from existing quality-of-democracy indices also runs counter to the large body of empirical research on political capacity within the field of political behavior. Key questions that have animated this research include: are citizens able to maintain internally-consistent and ideologically-structured beliefs? How politically knowledgeable and civically literate are citizens? Do citizens interrogate their own beliefs by finding and accurately processing new or unbiased sources of political information? How capable are citizens of voting for politicians and parties that will best represent their values and interests?[7] Debates among empirical political scientists regarding how much and what kinds of political

[6] An exception is the Democracy Ranking (Campbell, 2008), which includes measures of secondary-school and university enrolment aimed at capturing the availability of "knowledge" in a society. The Economist Intelligence Unit's (EIU) Democracy Index includes data on levels of adult literacy and the share of the population that follows politics in the news.

[7] See, for example, Achen & Bartels, 2016; Alvarez & Nagler, 2000; Andersen, Tilley, & Heath, 2005; Arnold, 2012; Barabas & Jerit, 2009; Bartels, 1996; Campbell, Converse, Miller, & Stokes, 1960; Converse, 1964; Delli Carpini & Keeter, 1996; Lau, Patel, Fahmy, & Kaufman, 2014; Lavine, Johnston, & Steenbergen, 2012; Lodge & Taber, 2013; Lupia, 2016; Milner, 2002; Mutz, 2006; Nie, Junn, & Stehlik-Barry, 1996; Page & Shapiro, 1992; Rapeli, 2014; Rosema & de Vries, 2011; Zaller, 1992.

capacity are required of citizens for democracy to flourish is reflective of important conceptual disagreements about what makes a democracy high quality. All major models of democracy clearly identify political capacity as important for democracy; they differ significantly however in their understanding of what types and levels of political capacity matter for high-quality democracy.

How exactly then do different models of democracy understand the concept of political capacity? To answer this question, we propose focusing on three types of political capacity. The first is the capacity of citizens to understand or know their own values, preferences, and interests that they wish to see realized through the democratic process. The second is the capacity of citizens to identify and select elites who will defend and advance those values, preferences, and interests in the political arena. The third and final capacity is the capacity to influence political elites and the agendas they pursue. For the sake of simplicity, we refer to these three core democratic capacities as the *capacity to know, the capacity to choose, and capacity to influence*. For a summary of how these three capacities are understood by three key models of democracy, see Table 1.

Let us first turn to the capacity to know. For advocates of minimal-elitist democracy, little is expected of citizens. Schumpeter famously argued that citizens are "incapable of action other than a stampede" (1950, p. 283); such low levels of political capacity associated with stampede-like cognition and affect are seen as in no way undermining a country's quality of democracy. For minimal-elitists, citizens need only be capable of *selecting into* their values, preferences, and interests based on the menu of options provided to them by political elites during the short window of public debate that periodically occurs prior to elections. That said, as Schumpeter points out, for minimal-elitist democracy to work well, citizens must be on "an *intellectual* and moral level high enough to be proof against the offerings of the crook and the crank" (1950, p. 294, emphasis added). This suggests that the "good" citizen for minimal-elitists is able to process the content of pre-election public debate in ways that allow her to identify and resist the siren call of false information.

Liberal-pluralist and participatory models of democracy are more demanding of citizens in terms of their "capacity to know" their own values and interests. Both models share an expectation that citizens should have the capacity to arrive at what Tocqueville described as "self-interest rightly understood" or what Dahl refers to as "enlightened understanding". In *Democracy and Its Critics* (1989, pp. 111–112), Dahl writes that "to know what it wants, and what is best, the people must be enlightened". To achieve such enlightenment, Dahl argues that citizens must acquire "an understanding of means and ends, of one's interests and the expected consequences of policies for interests, not only for oneself but for all other relevant persons as well".[8] The implication of this is that citizens are expected to be capable of finding and processing information and weighing the consequences of their values and interests on those of fellow citizens.

When it comes to citizens' capacity to choose political elites who will defend and pursue their interests, minimal-elitist, liberal-pluralist, and participatory models of democracy have much in common. None of them requires citizens to be extraordinary information sleuths or indeed policy wonks; rather, they expect citizens to be capable of taking full advantage of elite-provided sources of structured information in order to choose leaders without, as Schattschneider (1975, p. 134) puts it, being duped by demagogues. The models diverge, however, along two dimensions: first, in terms of the range of elite actors that citizens are expected to select; and second, in terms of the period of time over which citizens are expected to select elites.

For minimal-elitists, the "good" citizen need only be able to tune into politics in short bursts at election time. Using information shortcuts generated by the process of political competition during the campaign period, citizens are expected to have the political wherewithal to select candidates and parties who will best serve their values and interests. For liberal-pluralists (see Galston, 1988, p. 1283) and participatory democrats, citizens are also expected to be able to make sense of available information to select the right candidates and parties at election time. In addition, they must be sufficiently tuned into politics on an ongoing basis to be able to identify and support organizations and associations that will defend their values and interests, as and when the need arises, by applying pressure on elected politicians between elections.

Finally, what do the three models have to say about citizens' capacity to influence? Minimal-elitists expect citizens to influence politics and policy making indirectly through their vote choices and certainly not between elections. Liberal-pluralist and participatory democrats expect citizens to influence elites through forms of Hirschmanian exit and voice. To influence elites via voice requires citizens to possess cognitive, expressive, and organizational capacities. This includes the ability to identify whom to target and the capacity to work with others to influence them. For participatory democrats, who argue that high-quality democracies provide wide-ranging opportunities for citizens to get involved in shaping public policy, it is particularly important that citizens possess skills and knowledge that enable them to cooperate, communicate, and deliberate with fellow citizens and political elites alike (see Barber, 1984, p. 154).

2.3. Political Participation

One of the few citizen-related indicators that routinely appears in existing cross-national quality-of-democracy

[8] For in-depth discussions of the capacities expected of the "good" liberal citizen, see Galston (1988, especially pp. 1283–1285) and Macedo (1990, especially pp. 265–273). For the capacities required of the "good" participatory citizen, see the discussion of "strong democratic talk" in Barber (1984, pp. 178–198).

indices is turnout in national elections (see Altman & Pérez-Liñán, 2002; Bühlmann et al., 2013; EIU, 2012; Levine & Molina, 2011).[9] This clearly points to a scholarly consensus that political participation is a core conceptual component of democratic quality. High-quality democracy cannot simply be understood in terms of the existence of particular kinds of democratic institutions, the most incontrovertible of which are free and fair elections; it is also defined by whether citizens actually turn out to vote in those elections. All major models of democracy set great store by electoral participation. They do differ significantly though in the importance they attach to other forms of political participation.

Over the years normative disagreements among political theorists and political philosophers have inspired and echo similar debates among scholars of political behavior. In fact, the question of what types of political participation are found in high-quality democracies goes back to one of the founding studies in the field of political behavior, *The Civic Culture* by Gabriel Almond and Sidney Verba (1963), who argued that democracies are best served by citizens who "balance" political activity and passivity. In the half-century since the publication of *The Civic Culture*, patterns of popular political participation have changed greatly. However, the question of how active citizens should be, and what forms political activity should take remains central to the study of political behavior.[10]

To capture how different models of democracy conceive of political participation, we propose that scholars of democratic quality pay particular attention to how much weight is attached to: (1) participation focused on elections versus acts of political participation that occur between elections; (2) mediated forms of political participation where citizens seek to influence politics through organized civil society versus direct forms of political action and participation; and (3) the extent to which political participation is "other-regarding" and public-oriented. See Table 1 for a summary of the discussion below.

For minimal-elitists, elections are the singular focus of citizen participation. The primary political act of the "good" citizen is therefore to turn out in periodic elections. To avoid political demagoguery, it can be assumed that minimal-elitists expect citizens to pay attention to politics during election campaign periods, consume political news, and engage in political discussions.[11] Between elections, however, citizens are expected to engage in few, if any, political acts, leaving politics to politicians and parties.

For liberal-pluralists, citizens are under no duty to participate actively in politics (Galston, 1988, p. 1284). That said, there is an expectation that they will turn out to vote, in line with their self-interest rightly understood. This implies that the "good" liberal-pluralist citizen will engage in forms of other-regarding political activities that allow her to achieve an "enlightened" understanding of her values and interests. The emphasis is placed on forms of mediated political participation, most notably engagement with organizations and associations, and by extension social movements. There is also an expectation that in a high-quality liberal-pluralist democracy, citizens will—to quote Stephen Macedo (1990, p. 274)—"take initiatives on their own [and] be prepared to combine in voluntary associations for common ends both altruistic and otherwise".

For participatory democrats, citizens are expected to be engaged in the electoral process in similar ways to the "good" liberal-pluralist citizen. However, the participatory model of democracy places a duty on citizens to be directly and actively involved in politics on an ongoing politics. As Barber (1984, p. 152) writes, "[participatory] democracy is the politics of amateurs, where every man is compelled to encounter every other man without the intermediary of expertise". Finally, as these words suggest, participatory democrats also expect citizens to undertake political activities that are expressly other-regarding and public-oriented, aimed at moving beyond "competitive interest mongering" (1984, p. 155).

3. Inter-Component Congruence

Democracies don't just need good institutions, they also need citizens who are willing and able to breathe life into those institutions. A version of this claim stands at the heart of classic studies of democratic consolidation (Diamond, 1999; Linz & Stepan, 1996) as well as more recent debates about democratic deconsolidation (see Alexander & Welzel, 2017; Foa & Mounk, 2017; Inglehart, 2016; Norris, 2017; Voeten, 2017). The basic contention of this body of work is that democracy can be considered consolidated and stable when, among other things, democratic institutions are firmly established *and* citizens are meaningfully and unwaveringly supportive of democracy.

In contrast to research on democratic consolidation, research on democratic quality has made little effort to conceptualize the relationship between institutions and citizens. The widespread inclusion of (national) electoral turnout data in existing cross-national quality-of-democracy indices points to an underlying academic con-

[9] Existing quality-of-democracy indices also routinely include other participation-related indicators. The Democracy Barometer, for example, includes data on reported rates of petitioning and demonstrations; Levine and Molina (2011) include data on the share of citizens who report having worked for a candidate or party; the EIU incorporates information on membership of political parties and political non-governmental organizations as well as participation in demonstrations.

[10] See, for example, Dalton (2008); Fung (2004); Mutz (2006); Norris (2002); Rosenstone and Hansen (1993); Stolle and Micheletti (2013); Verba, Nie and Kim (1978); Verba, Schlozman and Brady (1995).

[11] Some scholars of political participation do not consider political discussion or political news consumption a form of political participation because—to quote Verba et al. (1995, p. 40) "the target audience is not a public official". Here we adopt a more expansive understanding of political participation that allows us to accommodate the full range of political actions identified explicitly or implied by different models of democracy.

sensus that citizens are indeed conceptually constitutive of democratic quality. However, this same research has fallen short of giving any systematic conceptual consideration to how citizens matter for democratic quality beyond participation in elections. By extension, they have also failed to recognize the crucial issue that citizens matter in different ways depending on the model driving the assessment. The conceptual short shrift that researchers have given to citizens stands in marked contrast to the detailed and sophisticated discussions about how and why different kinds of institutions and structures matter for democratic quality. The goal of the previous section of this article was to address this important gap in the literature by providing a conceptual account of the citizen component of democratic quality. In this section we take a step back to address the more general conceptual question of the relationship between the citizen and institutional components of democracy.

We conceive of the relationship between institutions and citizens as it pertains to democratic quality in terms of congruence.[12] As such, we follow Mayne and Geissel (2016, p. 636) in viewing the relationship between the citizen and institutional components of democratic quality as one of mutual dependence or mutual conditionality. Our basic contention therefore is that institutions and citizens represent two sides of the same democracy coin, meaning that democratic quality is a function of the level of model-specific congruence between institutions and citizen dispositions. The more institutions and citizen dispositions are simultaneously congruent with the demands and expectations of the same model of democracy, the higher that country's quality of democracy becomes, at least when judged from the viewpoint of the model in question.

Given that both the institutional and citizen components are necessary conditions of democratic quality, it is important to be clear about a key implication of our argument. If a country's political institutions accord largely with the expectations of a particular model of democracy, but citizen dispositions do not (or vice versa), we simply cannot say that this country has a high-quality democracy. How exactly inter-component incongruence would ultimately be calculated to arrive at a country's overall democracy score is a question for future empirical research. The point we wish to make is that the value of one component must, in a non-negligible way, be contingent on the value of the other component. When considering this issue of mutual contingency, it is important to distinguish between two types of inter-component congruence: one static; the other dynamic.

3.1. Static Congruence

When one thinks about democratic quality in terms of inter-component congruence, most likely one intuitively thinks about congruence at a single point in time, i.e., static congruence. To illustrate this form of congruence, we turn now to a brief examination of the level of fit between citizen support for direct democracy, on the one hand, and the institutionalization of direct democracy, on the other. With this worked example, to be clear, we are assessing democratic quality from the perspective of the participatory model of democracy. Given the limitations of existing cross-national data sources, we must content ourselves here with this partial illustration, which offers but a small and incomplete analytic window into understanding levels of model-specific inter-component congruence.

Table 2 presents information, from a broad cross-section of countries, on the share of citizens who view referendums as an essential component of democracy, alongside a (national-level) measure of the actual institutionalization of direct-democratic mechanisms.[13] Taken together, these indicators provide at best minimally suggestive evidence for evaluating democratic quality from the perspective of the participatory model of democracy; still they are very helpful in illustrating inter-component congruence.

Surveying the data in Table 2, it is clear that citizen commitments and real-world institutions match in some countries but are totally "out of sync" in others. A majority of citizens in Switzerland and Uruguay support the idea that democracies should give people a direct say in political decision making; both countries also offer a range of direct-democratic mechanisms. Similarly, we observe higher levels of inter-component congruence (but in the opposite direction) in Finland, the Netherlands, and the United Kingdom, where less than 20 percent of citizens report strong participatory democratic commitments and where direct democratic mechanisms are weakly institutionalized. In contrast, we find evidence of inter-component incongruence in many other countries. Popular majorities in Cyprus, Argentina, and Germany, for example, prefer a participatory form of democracy, but direct democracy is weakly institutionalized in those countries. From the point of view of static inter-component congruence and using this imperfect illustration of participatory democracy as the yard stick of evaluation, we would conclude that democracy is of a higher quality in Switzerland and Uruguay than in Argentina and Cyprus.

[12] See Almond and Verba (1963); Eckstein (1998); Welzel and Inglehart (2006); Welzel and Klingemann (2011).

[13] Data on the presence of direct democracy come from the Democracy Barometer (2012). Public opinion data come from the fifth wave of the World Value Survey (fielded between 2005 and 2009). Nationally representative samples of citizens were asked whether referendums are an essential part of democracy and provided with a 1–10 response scale, with 1 indicating that referendums are not at all essential and 10 indicating that they are definitely essential. We follow other research in using response data only for the scale maximum. As Kriesi et al. (2016, p. 67) note, survey respondents "who choose a value below the scale maximum arguably allow for exceptions and do not consider the given element as required for democracy under all circumstances".

Table 2. Support for and institutionalization of direct democracy. Source: Geissel (2016).

Country	Support for direct democracy	Institutionalization of direct democracy
Argentina	53.8	1
Australia	46.7	1
Brazil	43	1
Bulgaria	39.8	2
Canada	19.5	1
Chile	37.9	0
Cyprus	55.1	0
Finland	14.3	0
France	24.6	1
Germany	51.8	0
Hungary	40.9	3
Japan	30	1
Mexico	19.1	0
Netherlands	17.3	0
Norway	31.7	0
Peru	29.8	3
Poland	42.9	1
Romania	43	1
Slovenia	39.6	3
South Africa	20.9	0
South Korea	34.3	1
Spain	45.8	0
Sweden	43	0
Switzerland	59.6	3
Turkey	45.3	1
United Kingdom	17	0
United States	25.5	0
Uruguay	50.4	3

3.2. Dynamic Congruence

Over time political institutions change, as do citizen dispositions; and in many ways these changes are profoundly connected. Existing political institutions not only bound many citizens' democratic imagination, they also play an important role in shaping how citizens participate in politics as well as the political capacities they develop. Likewise, by failing to meet citizens' expectations, political institutions can generate democratic re-imaginings, stimulate and diffuse alternative forms of political participation, and encourage citizens to develop new political capacities. The opposite is also true. Elites reform political institutions in part as a response to change over time in citizens' democratic commitments, transformations in political activism, and improvements in mass political capacities. Moreover, institutional reform and changes in citizen dispositions seldom proceed in a neat, linear fashion. This means that, when viewed over a long period of time, both types of change might slowly be moving a country away from congruence with one model of democracy, toward congruence with another model. During periods of change, however, we will necessarily observe inter-component incongruence.

From the static perspective of congruence, inter-component incongruence lowers a country's quality of democracy, which may be misleading from a long-term perspective. It is crucial to make allowances for the processes of mutual adjustment of institutions and citizen dispositions toward new equilibria. A key analytic advantage of conceiving of inter-component congruence in both static and dynamic terms is that it allows us to distinguish between two sets of democracies. On the one hand, low-quality democracies where institutions and citizens are effectively more or less permanently out of sync with each other. And on the other hand, countries where institutions and citizen dispositions are slowly moving in the same democratic direction; and where the processes of mutual adjustment underpinning these changes are in fact a powerful positive indicator of the quality of democracy.

4. Inter-Dispositional Consistency

Just as different models of democracy ideally expect the institutional and citizen components of democratic quality to be congruent, they also expect—ideally—democratic commitments, political capacities, and political participation to be consistent with each other. Inter-dispositional consistency (or *intra*-component congruence) represents an important yardstick for evaluating democratic quality, because, regardless of the model of democracy driving the assessment, a high-quality democracy depends on a particular mix and balance of commitments, capacities, and participation. For example, a high-quality participatory democracy is not just home to large numbers of citizens participating actively in politics, at and between elections, but also to large numbers of people who have the capacities to cooperate, communicate, and deliberate. Similarly, minimal-elitists might only expect citizens to participate in periodic elections, but when they do, they are also expected to be able to avoid being misled or fooled by political elites vying for their votes.

Over the years, scholars of political behavior have studied empirically how citizen dispositions relate to one another. One approach has been to examine the influence of certain kinds of democratic commitments on political participation. Recent work, by Åsa Bengtsson and Henrik Christensen (2016) and Sergiu Gherghina and Geissel (2017) finds clear associations between citizens' democratic "process preferences" and how they participate in politics. For example, citizens who support a participatory model of democracy are more likely to participate in politics (see also Bolzendahl & Coffé, 2013; Dalton, 2008). A large body of research also exists on the question of how political capacities relate to political participation. Consistently research has found a positive association between education and political participation. Compared to the legion of studies that examines the

impact of education, income and political interest (as a proxy of political capacity), very little research has been done on how the cognitive, expressive, and organizational capacities specifically identified by different models of democracy relate to participation.[14] This is mainly due to the dearth of cross-national survey data, aimed at capturing information on political capacity. All in all, we still know very little about how the varied citizen dispositions prized by minimal-elitist, liberal-pluralist, and participatory democracy "move" together.

We now turn to some worked examples. We rely here on existing cross-national survey data, from the World Values Survey, the International Social Survey Program, the European Social Survey, and the Latin American Public Opinion Project. In particular, we focus on whether democratic commitments, using data on citizen support for referendums, align with political capacity and political participation.[15] Though extremely crude, the question on referendums is helpful for the present purpose of illustrating a key aspect of citizens' democratic commitments—namely, the importance of political participation beyond regular legislative elections. We use citizens' reported interest in politics as a very rough (and imperfect) overarching proxy for citizen capacity, and create an index of non-electoral political participation using available data sources.[16]

Figure 1 plots aggregate-level support for referendums against the share of citizens who say they are interested in politics. Examining the data from the perspective of the participatory model of democracy, it becomes clear that very few countries display a high level of inter-dispositional consistency. Only three countries (Denmark, Germany, and Switzerland) are home to citizenries with the commitments and capacities expected by participatory democrats; namely, sizeable majorities (of more than 60 percent) who see referendums as a good way to decide important political issues and who are interested in politics. In many other countries, however, participatory commitments are misaligned with political capacities. Cyprus, Croatia, Spain, and Brazil are

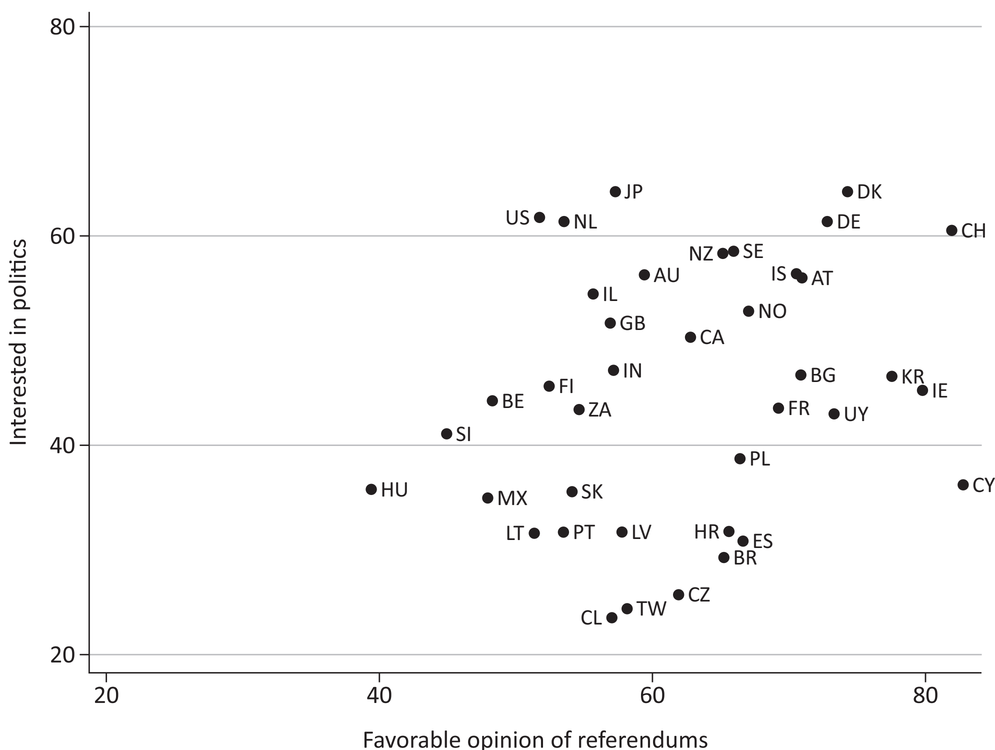


Figure 1. Interest in politics and support for direct democracy.

[14] A small body of work, mainly focused on the United States, exist on the relationship between political knowledge and political participation, see Delli Carpini and Keeter (1996); Milner (2002); Nie et al. (1996); Verba et al. (1995).

[15] The question comes from two rounds of the International Social Survey Programme (ISSP, fielded in 2004 and 2014). Survey respondents were asked if they agree that referendums are a good way to decide important political questions. We report here the results for citizens who strongly agree with this statement. We aggregate (and average, where necessary) data using available population weights to produce measures of citizen dispositions that are representative of citizenries as a whole.

[16] We use a question (from ISSP 2004 and 2014) about citizens' general interest in politics. Respondents who say they are very or somewhat interested in politics are combined. It is important to note, however, that the minimal-elitist model of democracy expects citizens to be interested in politics during elections, but not between them. The index of non-electoral political participation is based on three questions that appear in the ISSP (2004 and 2014), the World Values Survey, the European Social Survey, and Latin American Public Opinion Project. These relate to signing a petition, taking part in a demonstration, and contacting a public official. The index captures the share of citizens who report having undertaken at least one of these activities in the past 12 months.

cases in point: large numbers of citizens favor a participatory approach to political decision making, but far fewer are interested in politics. From the perspective of minimal-elitism, only Hungary comes close to displaying ideal levels of inter-dispositional consistency. More than 60 percent of Hungarians have no strong desire for referendums and an even greater share of Hungarians say they are not interested in politics. In a number of other countries, such as Chile or the Czech Republic, we find citizenries with the kinds of political capacity (reflected in low levels of political interest) that minimal-elitists argue make for a higher-quality democracy, but who also report democratic commitments that are inconsistent with a high-quality minimal-elitist democracy. Few countries are home to citizenries with commitments and capacities consistent with a liberal-pluralist account of democracy. Belgium and Slovenia arguably come closest. In many other democracies (such as France, Uruguay, and Ireland) we find citizenries with political capacities that are consistent with liberal-pluralism, but whose democratic commitments are not.

Figure 2 once again plots our indicator of democratic commitments (namely, support for referendums), but this time against a measure of non-electoral political participation. For minimal-elitists, a high-quality democracy is one where few citizens undertake political activities between elections; the opposite is true for participatory democrats. For liberal-pluralists, democracy works best when citizens exert themselves politically between elections only when the need arises. As a result, moderate levels of non-electoral political participation are arguably most consistent with high-quality democracy from the perspective of liberal-pluralists.

Hungary is the most obvious case of minimal-elitist inter-dispositional consistency. As noted earlier, Hungarians appear to be less fond of direct democracy and very few engage in political activities between elections. Iceland, New Zealand, Canada, and Norway, among others, are home to citizenries with democratic commitments and patterns of political participation consistent with participatory democracy. Applying a liberal-pluralist benchmark, Belgium and the Netherlands appear to come closest to displaying the desired types and levels of political participation and democratic commitments. Most countries, however, are home to large numbers of people with mixed patterns of citizen dispositions that are inconsistent with any one model of democracy.

5. Conclusion

In this article we have argued that democratic quality depends not only on the form and functioning of democratic institutions but also on the dispositions of citizens.[17] To date, however, cross-national indices have focused predominantly on the institutional component of democratic quality, the measures of which have become

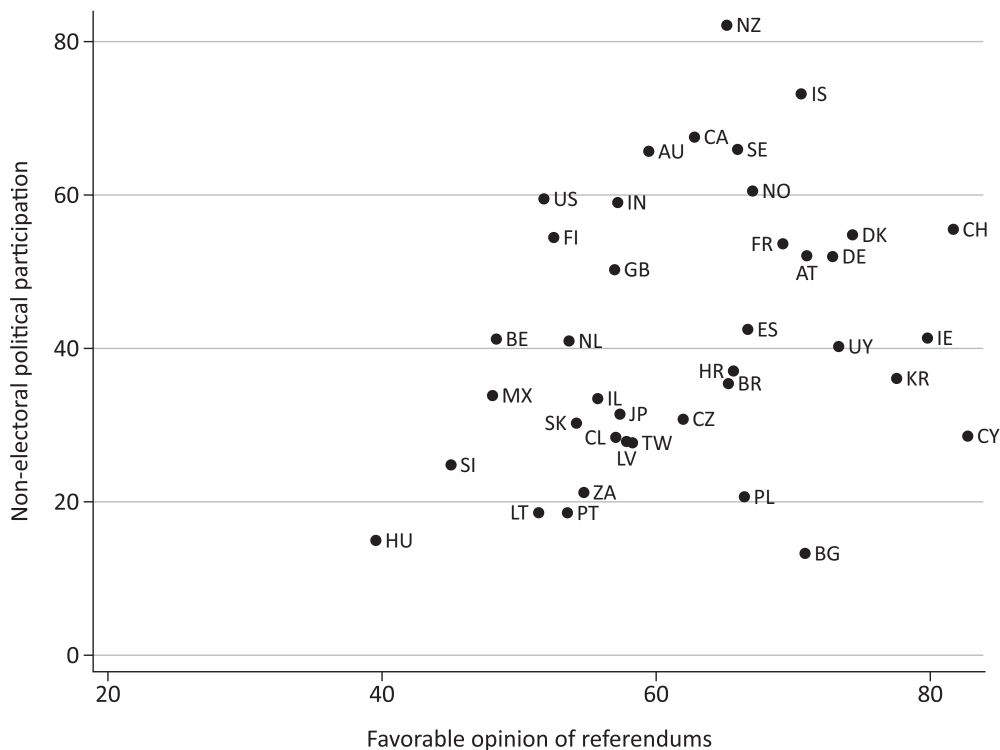


Figure 2. Non-electoral political participation and support for direct democracy.

[17] Democratic quality also depends on the dispositions of political elites, most obviously their commitment to democracy as well as their level of political competence (see, for example, Linz & Stepan, 1996; Mainwaring & Pérez-Liñán, 2013; Steiner, Bächtiger, Spörndli, & Steenbergen, 2004).

increasingly multidimensional and conceptually sophisticated. The Varieties of Democracy program (Coppedge et al., 2011) has enriched this approach by making it possible to systematically evaluate democratic institutions according to different models of democracy. The same cannot be said of the citizen component of democratic quality. Existing indices commonly incorporate information on national turnout rates, which points to an academic consensus that citizens are indeed a constitutive element of democratic quality. Few other citizen-related indicators are, however, included, and when they are it is often with little theoretical justification. The result is that citizens play conceptual second fiddle to institutions, and there is little or no recognition that different accounts of democracy demand and expect different kinds of citizen dispositions. Our aim with this article is to challenge this orthodoxy by providing a structured account of the citizen component of democratic quality, with a focus on three models of democracy—minimal-elitism, liberal-pluralism, and participatory democracy.

The first section of the article provided a fine-grained conceptualization of what we argue are the three core dispositions that make up the citizen component of democratic quality—namely, democratic commitment, political capacity, and political participation. We made the case that commitment is not just about general support for democracy but also model-specific commitments related to who gets to decide and how decisions are to be made in the political arena. We defined political capacity in terms of citizens' ability to know, choose, and influence, identifying key differences in how the three models conceive of political capacity. Finally, to capture the kinds and levels of political participation that different models of democracy expect of citizens, we argued that scholars of democratic quality should focus on the weight attached to: election-focused participation versus participation between elections; mediated versus direct forms of political action; and "other-regarding" political participation that brings together citizens with divergent political viewpoints.

The second and third sections of the article deal with two key issues that arise when taking citizens seriously in the conceptualization of democratic quality. The first is the issue of "fit" between institutions and citizens, which we refer to as inter-component congruence. Specifically, we made the case that any assessment of democratic quality must consider the extent to which both institutions and citizen dispositions are congruent with the same model of democracy. We further underscored the importance of distinguishing static congruence—where democratic quality is judged according to the level of inter-component congruence at single point in time, and dynamic congruence—where democratic quality is judged according to long-term processes of mutual adjustment between institutions and citizen disposition toward the same model of democracy.

The other significant issue we addressed was inter-dispositional consistency. Ideally, we argued, democratic commitments, capacities, and participation should all be consistent with the same model of democracy. To illustrate this point we turned to an analysis of existing cross-national survey data, which provided suggestive evidence that citizen dispositions are highly inconsistent with each other in many democracies. Cognizant of the imperfections of the available data, this nonetheless has important implications when developing composite measures of the citizen component of democratic quality.

The greatest challenge moving forward with our conceptualization of democratic quality relates to data availability. In writing this article we undertook a systematic and broad survey of existing cross-national surveys.[18] Our aim was to identify questions that could serve as indicators to operationalize the citizen component of democratic quality. In some regards, existing cross-national survey programs provide a solid foundation to build on; in other respects, however, much work remains to be done. In recent years, the measurement of democratic commitments has improved greatly. Well-established questions that gauge citizens' general support for democracy have been supplemented with new batteries of questions shedding light on citizens' commitment to specific democratic principles, capturing information on a variety of democratic principles shared by all key models of democracy. There has, however, been some effort to include one or two questions that allow researchers to distinguish commitments to principles specific to minimal-elitist, liberal-pluralist, and participatory accounts of democracy. One goal of this article has been to provide a fuller account of the commitments expected of citizens by different models of democracy.

We also found that the measurement of political participation is fairly strong, with information being frequently collected on a broad range of non-electoral forms of participation. It is difficult though to isolate "other-regarding" forms of political participation, which are important for liberal-pluralist and participatory accounts of democracy. Given the challenges of political polarization and social division that face many democratic societies today, future quality-of-democracy research would therefore benefit from being able to measure how much citizens are actually engaging in political activities that involve encountering and working with others with different political viewpoints.

Finally, and most worrying of all, we found that the measurement of political capacities is weak. Cross-national surveys often ask citizens to self-report on their general sense of political understanding or competence. Some surveys also gauge citizens' level of political knowledge, but developing cross-nationally commensurable measures of political knowledge has been challenging (Gidengil, Meneguello, Shenga, & Zechmeister, 2016). Overall though, unlike some surveys carried out in in-

[18] This included the World Values Survey, ISSP, European Social Survey, Comparative Study of Electoral Systems, European Election Study, and Latin American Public Opinion Project.

dividual countries, to date no cross-national measures have been fielded aimed at directly capturing information on citizens' cognitive, expressive, and organizational capacities. This is not to underestimate the difficulty of developing valid and reliable empirical indicators of political capacity, but the lack of data in this area poses a real problem for quality-of-democracy research. As we have argued in this article, citizens' capacity to know, choose, and influence in the political arena is central to the quality of democracy. By detailing how different models of democracy understand these three capacities in different ways, we hope that this article provides a valuable resource for developing new survey questions to fully incorporate the citizen component into future quality-of-democracy research.

Acknowledgments

We thank the anonymous reviewers and Heiko Giebler, Saskia Ruth, and Dag Tanneberg, editors of this thematic issue, for their valuable feedback. We also thank participants of a special issue workshop, held at the University of Zurich in June 2017, and members of the democratic innovations research group at Goethe University Frankfurt for their helpful comments. Finally, we thank Anna Krämling for research assistance.

Conflict of Interests

The authors declare no conflict of interests.

References

Achen, C. H., & Bartels, L. M. (2016). *Democracy for realists: Why elections do not produce responsive government*. Princeton, NJ: Princeton University Press.

Alexander, A. C., & Welzel, C. (2017). The myth of deconsolidation: Rising liberalism and the populist reaction. *Journal of Democracy (Web Exchange)*. Retrieved from https://www.journalofdemocracy.org/online-exchange-%E2%80%9Cdemocratic-deconsolidation%E2%80%9D

Almond, G. A., & Verba, S. (1963). *The civic culture: Political attitudes and democracy in five nations*. Princeton, NJ: Princeton University Press.

Altman, D., & Pérez-Liñán, A. (2002). Assessing the quality of democracy: Freedom, competitiveness and participation in eighteen Latin American Countries. *Democratization*, *9*(2), 85–100.

Alvarez, M. E., & Nagler, J. (2000). A new approach for modelling strategic voting in multiparty elections. *British Journal of Political Science*, *30*(1), 57–75.

Andersen, R., Tilley, J., & Heath, A. F. (2005). Political knowledge and enlightened preferences: Party choice through the electoral cycle. *British Journal of Political Science*, *35*(2), 285–302.

Arnold, J. R. (2012). The electoral consequences of voter ignorance. *Electoral Studies*, *31*(4), 796–815.

Barabas, J., & Jerit, J. (2009). Estimating the causal effects of media coverage on policy-specific knowledge. *American Journal of Political Science*, *53*(1), 73–89.

Barber, B. R. (1984). *Strong democracy: Participatory politics for a new age*. Berkeley, CA: University of California Press.

Bartels, L. M. (1996). Uninformed votes: Information effects in presidential elections. *American Journal of Political Science*, *40*(1), 194–230.

Bengtsson, Å., & Christensen, H. (2016). Ideals and actions: Do citizens' patterns of political participation correspond to their conceptions of democracy? *Government and Opposition*, *51*(2), 234–260.

Bolzendahl, C., & Coffé, H. (2013). Are 'good' citizens 'good' participants? Testing citizenship norms and political participation across 25 nations. *Political Studies*, *61*(S1), 45–65.

Bratton, M., Mattes, R. B., & Gyimah-Boadi, E. (2005). *Public opinion, democracy, and market reform in Africa*. Cambridge and New York, NY: Cambridge University Press.

Bühlmann, M., Merkel, W., Müller, L., Giebler, H., Wessels, B., Boschler, D., . . . Bousbah, K. (2013). *Democracy barometer. Codebook for blueprint dataset version 3*. Aarau: Zentrum für Demokratie.

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter*. Chicago, IL: University of Chicago Press.

Campbell, D. (2008). *The basic concept for the democracy ranking of the quality of democracy*. Vienna: Democracy Ranking.

Canache, D. (2012). Citizens' conceptualizations of democracy: Structural complexity, substantive content, and political significance. *Comparative Political Studies*, *45*(9), 1132–1158.

Carlin, R. E. (2017). Sorting out support for democracy: A q-method study. *Political Psychology*, 1–24. doi:10.1111/pops.12409

Carlin, R. E., & Singer, M. M. (2011). Support for polyarchy in the Americas. *Comparative Political Studies*, *44*(11), 1500–1526.

Carrión, J. F. (2008). Illiberal democracy and normative democracy: How is democracy defined in the Americas? In M. A. Seligson (Ed.), *Challenges to democracy in Latin America and the Caribbean: Evidence from the Americas Barometer 2006–2007* (pp. 21–51). Nashville, TN: Latin American Public Opinion Project, Vanderbilt University.

Converse, P. E. (1964). The nature of belief systems in mass publics. In D. E. Apter (Ed.), *Ideology and discontent* (pp. 75–169). New York, NY: Free Press.

Coppedge, M., Gerring, J., Altman, D., Bernhard, M., Fish, S., Hicken, A., . . . Teorell, J. (2011). Conceptualizing and measuring democracy: A new approach. *Perspectives on Politics*, *9*(2), 247–267.

Dahl, R. A. (1971). *Polyarchy: Participation and opposition*. New Haven, CT: Yale University Press.

Dahl, R. A. (1989). *Democracy and its critics*. New Haven, CT: Yale University Press.

Dalton, R. J. (2008). Citizenship norms and the expansion of political participation. *Political Studies*, *56*(1), 76–98.

Dalton, R. J., Sin, T., & Jou, W. (2007). Understanding democracy: Data from unlikely places. *Journal of Democracy*, *18*(4), 142–156.

Delli Carpini, M. X., & Keeter, S. (1996). *What Americans know about politics and why it matters*. New Haven, CT: Yale University Press.

Diamond, L. J. (1999). *Developing democracy: Toward consolidation*. Baltimore, MD, and London: Johns Hopkins University Press.

Eckstein, H. (1998). Congruence theory explained. In H. Eckstein (Ed.), *Can democracy take root in Post-Soviet Russia? Explorations in state–society relations* (pp. 3–34). Lanham, MD: Rowman and Littlefield.

Economist Intelligence Unit. (2012). *Democracy index*. London: Economist Intelligence Unit.

Foa, R. S., & Mounk, Y. (2017). The signs of deconsolidation. *Journal of Democracy*, *28*(1), 5–15.

Fuchs, D., & Roller, E. (2006). Learned democracy? Support of democracy in central and eastern Europe. *International Journal of Sociology*, *36*(3), 70–96.

Fung, A. (2004). *Empowered participation: Reinventing urban democracy*. Princeton, NJ: Princeton University Press.

Galston, W. A. (1988). Liberal virtues. *The American Political Science Review*, *82*(4), 1277–1290.

Geissel, B. (2016). Should participatory opportunities be a component of democratic quality? The role of citizen views in resolving a conceptual controversy. *International Political Science Review*, *37*(5), 656–665.

Gherghina, S., & Geissel, B. (2017). Linking democratic preferences and political participation: Evidence from Germany. *Political Studies*, *65*(S1), 24–42.

Gidengil, E., Meneguello, R., Shenga, C., & Zechmeister, E. (2016). Political knowledge sub-committee report. *Final Report of CSES Module 5 Political Knowledge Subcommittee*. Retrieved from http://www.cses.org/plancom/module5/CSES5_PoliticalKnowledgeSubcommittee_FinalReport.pdf

Hibbing, J. R., & Theiss-Morse, E. (2002). *Stealth democracy: Americans' beliefs about how government should work*. New York, NY: Cambridge University Press.

Inglehart, R. (2016). How much should we worry? *Journal of Democracy*, *27*(3), 18–23.

Kornberg, A., & Clarke, H. D. (1994). Beliefs about democracy and satisfaction with democratic government: The Canadian case. *Political Research Quarterly*, *47*(3), 537–563.

Kriesi, H., Saris, W., & Moncagatta, P. (2016). The structure of Europeans' views of democracy. In M. Ferrín & H. Kriesi (Eds.), *How Europeans view and evaluate democracy* (pp. 64–89). Oxford and New York, NY: Oxford University Press.

Kymlicka, W., & Norman, W. (1994). Return of the citizen: A survey of recent work on citizenship theory. *Ethics*, *104*(2), 352–381.

Lalljee, M., Evans, G., Sarawgi, S., & Voltmer, K. (2013). Respect your enemies: Orientations towards political opponents and political involvement in Britain. *International Journal of Public Opinion Research*, *25*(1), 119–131.

Lau, R. R., Patel, P., Fahmy, D. F., & Kaufman, R. R. (2014). Correct voting across thirty-three democracies: A preliminary analysis. *British Journal of Political Science*, *44*(2), 239–259.

Lavine, H., Johnston, C. D., & Steenbergen, M. R. (2012). *The ambivalent partisan: How critical loyalty promotes democracy*. New York, NY: Oxford University Press.

Levine, D. H., & Molina, J. E. (Eds.). (2011). *The quality of democracy in Latin America*. Boulder, CO: Lynne Rienner.

Levitsky, S., & Ziblatt, D. (2018). *How democracies die*. New York, NY: Crown Publishing.

Lijphart, A. (1999). *Patterns of democracy: Government forms and performance in thirty-six countries*. New Haven, CT: Yale University Press.

Linz, J. J., & Stepan, A. (1996). *Problems of democratic transition and consolidation: Southern Europe, South America, and post-communist Europe*. Baltimore, MD, and London: The Johns Hopkins University Press.

Lodge, M., & Taber, C. S. (2013). *The rationalizing voter*. New York, NY: Cambridge University Press.

Logan, C., & Mattes, R. (2012). Democratising the measurement of democratic quality: Public attitude data and the evaluation of African political regimes. *European Political Science*, *11*(4), 469–491.

Lupia, A. (2016). *Uninformed: Why people know so little about politics and what we can do about it*. New York, NY: Oxford University Press.

Macedo, S. (1990). *Liberal virtues: Citizenship, virtue, and community in liberal constitutionalism*. Oxford: Clarendon Press.

Mainwaring, S., Brinks, D., & Pérez-Liñán, A. (2007). Classifying political regimes in Latin America, 1945–2004. In G. L. Munck (Ed.), *Regimes and democracy in Latin America: Theories and methods* (pp. 123–160). New York, NY: Oxford University Press.

Mainwaring, S., & Pérez-Liñán, A. (2013). *Democracies and dictatorships in Latin America: Emergence, survival, and fall*. New York, NY: Cambridge University Press.

Mayne, Q., & Geissel B. (2016). Putting the demos back into the concept of democratic quality. *International Political Science Review*, *37*(5), 634–44.

McClosky, H. (1964). Consensus and ideology in American politics. *American Political Science Review*, *58*(2), 361–382.

Mill, J. S. (2009). *Representative government*. Munich: GRIN. (Original work published 1861)

Miller, A. H., Hesli, V. L., & Reisinger, W. M. (1997). Conceptions of democracy among mass and elite in post-Soviet societies. *British Journal of Political Science*, *27*(2), 157–190.

Milner, H. (2002). *Civic literacy: How informed citizens make democracy work*. Hanover, NH: Tufts University—University Press of New England.

Mutz, D. C. (2006). *Hearing the other side: Deliberative versus participatory democracy*. New York, NY: Cambridge University Press.

Nie, N. H., Junn, J., & Stehlik-Barry, K. (1996). *Education and democratic citizenship in America*. Chicago, IL: University of Chicago Press.

Norris, P. (2002). *Democratic phoenix: Reinventing political activism*. Cambridge and New York, NY: Cambridge University Press.

Norris, P. (2017). Is Western democracy backsliding? Diagnosing the risks. *Journal of Democracy (Web Exchange)*. Retrieved from https://www.journalofdemocracy.org/online-exchange-%E2%80%9Cdemocratic-deconsolidation%E2%80%9D

Page, B. I., & Shapiro, R. Y. (1992). *The rational public: Fifty years of trends in Americans' policy preferences*. Chicago, IL: University of Chicago Press.

Pateman, C. (1970). *Participation and democratic theory*. Cambridge: Cambridge University Press.

Pickel, S., Breustedt, W., & Smolka, T. (2016). Measuring the quality of democracy: Why include the citizens' perspective? *International Political Science Review*, *37*(5), 645–655.

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

Rapeli, L. (2014). *The conception of citizen knowledge in democratic theory*. Basingstoke and New York, NY: Palgrave Macmillan.

Rosema, M., & de Vries, C. E. (2011). Assessing the quality of European democracy: Are voters voting correctly. In S. A. H. Denters, M. Rosema, & K. Aarts (Eds.), *How democracy works: Political representation and policy congruence in modern societies: Essays in honour of Jacques Thomassen* (pp. 199–217). Amsterdam: Pallas Publications.

Rosenstone, S. J., & Hansen, J. M. (1993). *Mobilization, participation, and democracy in America*. New York, NY: Pearson Education.

Schattschneider, E. E. (1975). *The semisovereign people: A realist's view of democracy in America*. Boston, MA: Wadsworth.

Schedler, A., & Sarsfield, R. (2007). Democrats with adjectives: Linking direct and indirect measures of democratic support. *European Journal of Political Research*, *46*(5), 637–659.

Schumpeter, J. A. (1950). *Capitalism, socialism, and democracy* (3rd ed.). New York, NY, and London: Harper & Brothers.

Silveira, L., & Heinrich, H.-A. (2017). Drawing democracy: Popular conceptions of democracy in Germany. *Quality & Quantity*, *51*(4), 1645–1661.

Steiner, J., Bächtiger, A., Spörndli, M., & Steenbergen, M. R. (2004). *Deliberative politics in action: Analyzing parliamentary discourse*. New York, NY: Cambridge University Press.

Stolle, D., & Micheletti, M. (2013). *Political consumerism: Global responsibility in action*. New York, NY: Cambridge University Press.

Thomassen, J. (1995). Support for democratic values. In H.-D. Klingemann & D. Fuchs (Eds.), *Citizens and the state* (Vol. 1, pp. 383–416). New York, NY: Oxford University Press.

Verba, S., Nie, N. H., & Kim, J.-O. (1978). *Participation and political equality: A seven-nation comparison*. Cambridge and New York, NY: Cambridge University Press.

Verba, S., Schlozman, K. L., & Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Cambridge, MA: Harvard University Press.

Voeten, E. (2017). Are people really turning away from democracy? *Journal of Democracy (Web Exchange)*. Retrieved from https://www.journalofdemocracy.org/online-exchange-%E2%80%9Cdemocratic-deconsolidation%E2%80%9D

Welzel, C., & Inglehart, R. (2006). Emancipative values and democracy: Response to Hadenius and Teorell. *Studies in Comparative International Development*, *41*(3), 74–94.

Welzel, C., & Klingemann, H.-D. (2011). Democratic congruence re-established: The perspective of "substantive" democracy. In M. Rosema, B. Denters, & K. Aarts (Eds.), *How democracy works: Political representation and policy congruence in modern societies* (pp. 89–114). Amsterdam: Pallas Publications.

Zaller, J. R. (1992). *The nature and origins of mass opinion*. New York, NY: Cambridge University Press.

About the Authors

Quinton Mayne (PhD, Princeton) is Associate Professor of Public Policy at the Kennedy School of Government at Harvard University. His main research interests lie at the intersection of comparative and urban politics, with a particular interest in political behavior and social policy in advanced industrial democracies.

Brigitte Geissel (PhD, TU Berlin) is professor of political science and sociology at Goethe University Frankfurt. Her research interests include democratic innovations, new forms of governance, and political actors (new social movements, associations, civil society, parties, political elites, citizens).

COGITATIO

Politics and Governance (ISSN: 2183–2463)
2018, Volume 6, Issue 1, Pages 48–59
DOI: 10.17645/pag.v6i1.1186

Article

Democracy and Human Rights: Concepts, Measures, and Relationships

Todd Landman

School of Politics and International Relations, University of Nottingham, Nottingham, NG7 2RD, UK;
E-Mail: todd.landman@nottingham.ac.uk

Submitted: 28 September 2017 | Accepted: 20 November 2017 | Published: 19 March 2018

Abstract

The empirical literature on democracy and human rights has made great strides over the last 30 years in explaining (1) the variation in the transition to, consolidation of, and quality of democracy; (2) the proliferation and effectiveness of human rights law; and (3) the causes and consequences of human rights across many of their categories and dimensions. This work has in many ways overcome the 'essentially contested' nature of the concepts of democracy and human rights conceptually, established different measures of both empirically, and developed increasingly sophisticated statistical and other analytical techniques to provide stronger inferences for the academic and policy community. This article argues that despite these many achievements, there remain tensions between conceptualisations of democracy and human rights over the degree to which one includes the other, the temporal and spatial empirical relationships between them, and the measures that have been developed to operationalize them. These tensions, in turn, affect the kinds of analyses that are carried out, including model specification, methods of estimation, and findings. Drawing on extant theories and measures of both, the article argues that there must be greater specificity in the conceptualisation and operationalization of democracy and human rights, greater care in the development and use of measures, and greater attention to the kinds of inferences that are made possible by them.

Keywords

administrative data; big data; democracy; events data; human rights; measurement; socio-economic; standards data; statistics; survey data

Issue

This article is part of the issue "Why Choice Matters: Revisiting and Comparing Measures of Democracy", edited by Heiko Giebler (WZB Berlin Social Science Center, Germany), Saskia P. Ruth (German Institute of Global and Area Studies, Germany), and Dag Tanneberg (University of Potsdam, Germany).

© 2018 by the author; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

In 1971, Robert Dahl published *Polyarchy* in which he set out a systematic framework for measuring and understanding two fundamental dimensions of democracy: contestation and inclusiveness. The combination of these two dimensions allowed for comparative analysis of a variety of regime types around the world, while normatively his concept of 'polyarchy', which included countries with a high degree of contestation and a high degree of inclusiveness, was argued to be the most preferred system of governance. In 1988, Neil Mitchell and James McCormick published one of the first systematic comparative analyses of human rights in the journal *World Politics*, which explained the cross-national variation in the protection of civil and political rights. Both of these publications and analyses relied on (1) systematic theorisation of the concepts under inquiry, (2) methods for measuring the concepts, and (3) analysis of variation and covariation within and between the measures across country cases (see Adcock & Collier, 2001, p. 531; Landman & Carvalho, 2009, pp. 32–34). Since these seminal publications on the empirical analysis of democracy and human rights, there have been countless studies on the (1) the variation in the transition to, consolidation of, and quality and performance of democracy; (2) the proliferation and effectiveness of human rights law; and (3) the causes and consequences of human rights across many

of their different categories and dimensions (see Landman, 2005b, 2009, 2013).

This kind of work has in many ways overcome the 'essentially contested' (Gallie, 1956) nature of democracy and human rights conceptually, established different and highly varied measures of both, and developed increasingly sophisticated statistical and other analytical techniques that provide stronger inferences for the academic and policy community. This article argues that despite these many achievements, tensions remain between theories of democracy and human rights over the degree to which one includes the other, the temporal and spatial empirical relationships between and among them, and the measures that have been developed to operationalize them. These tensions, in turn, affect the kinds of analyses that are carried out, including model specification, methods of estimation, and findings. Drawing on extant theories and measures of both, the article argues that there must be greater specificity in the conceptualisation and operationalization of democracy and human rights, care in the development and use of measures, and more attention to the kinds of inferences that they make possible.

The overall motivation for this article is to provide clarity about what we mean when we talk about democracy and human rights, the degree to which they might share certain but not all attributes, and to unpack the conceptual and empirical relationships that are evident between them. Establishing conceptual clarity informs our consideration of measurement strategies and consequently any empirical relationships between democracy and human rights that might be discovered. It is important not to conflate or elide democracy and human rights. It is equally important to show how, why, and to what degree the two are inter-related, focusing on the direction, magnitude, and significance of the relationship, while at the same time remaining conscious that such relationships to date fall far below perfect correlation. This empirical gap between democracy and human rights is crucial for understanding the political challenge of progressing human rights to be closer to their legal ideal.

In order to develop these arguments, the article is structured in four sections. The first section provides a brief overview of the definition of democracy and human rights to show where and how the two concepts have a variable degree of overlap with one another. The second section shows the different strategies for measuring democracy and human rights, including (1) events-based data, (2) 'standards-based' data, (3) survey based data, (4) socio-economic and administrative data and (5) big data analytical techniques. The third section provides an overview of many of the stylized facts about the empirical relationships between measures of democracy and human rights, as well as the tendency for empirical studies to use human rights measures as measures of democracy, repression, rule of law, and good governance. The discussion shows how the associations made in theory can be tested empirically. The final section examines the remaining challenges and limitations to the current state of measurement and analysis of democracy and human rights.

2. Democracy and Human Rights

Democracy and human rights are grounded in the shared principles of accountability, individual liberty, integrity, fair and equal representation, inclusion and participation, and non-violent solutions to conflict. Modern conceptions of democracy are based on the fundamental ideas of *popular sovereignty* and *collective decision* making in which rulers through various ways are held to account by those over whom they rule (see Beetham, Carvalho, Landman, & Weir, 2008; Landman, 2013). But beyond this basic consensus, there are many varieties of democracy (see Coppedge, Lindberg, & Skaaning, 2016) or 'democracy with adjectives' (Collier & Levitsky, 1997) that have been in use by scholars, practitioners and policy makers. These definitions can be grouped broadly into three main types: (1) procedural democracy, (2) liberal democracy, and (3) social democracy, the delineation of which largely rests on the variable incorporation of different rights protections alongside the general commitment to popular sovereignty and collective decision making. Understanding these different types of democracy and the degree to which they incorporate different categories of human rights affects the ways in which measures of both can and have been used for empirical research (Dooresnspleet, 2015; Landman, 2013, 2016; Landman & Carvalho, 2009, 2017; Landman & Häusermann, 2003). Absence of consideration of these lines of overlap has led to conceptual and empirical confusion in the literature on democracy and human rights, as well as in those studies that incorporate measures of either concept in their modelling strategies (see Munck, 2009).

Procedural definitions of democracy are most closely aligned with Robert Dahl's (1971) formulation in *Polyarchy* and include the two dimensions of contestation and participation. Contestation captures the uncertain peaceful competition necessary for democratic rule; a principle which presumes the legitimacy of a significant and organised opposition, the right to challenge incumbents, protection of the twin freedoms of expression and association, the existence of free and fair elections, and a consolidated political party system. Such a procedural definition of democracy can be considered a baseline set of conditions and a minimum threshold that can be used to assess and count the number of democracies in the world (see, e.g. Banks, 1971; Landman, 2013, pp. 3–5; Przeworski, Alvarez, Cheibub, & Limongi, 2000).

Liberal definitions of democracy preserve the notions of contestation and participation found in procedural definitions, but add more explicit references to the protection of certain human rights. Definitions of liberal democracy thus contain an institutional dimension and a rights dimension (see Foweraker & Krznaric, 2000).

The institutional dimension captures the idea of popular sovereignty, and includes notions of accountability, constraint of leaders, representation of citizens, and universal participation in ways that are consistent with Dahl's 'polyarchy' model outlined above. The rights dimension is upheld by the rule of law, and includes civil, political, property, and minority rights. Such a definition is arguably richer (or 'thicker') as it includes legal constraints on the exercise of power to complement the popular elements in the derivation of and accountability for power (Coppedge, 2012, pp. 17–33).

Social definitions of democracy maintain the institutional and rights dimensions found in liberal models of democracy but expand the types of rights that ought to be protected, including social, economic and cultural rights (although some of these are included in minority rights protection seen in liberal definitions) (Beetham, 1999; Brandal, Bratberg, & Thorsen, 2013; Doorenspleet, 2005; Landman, 2005, 2013, 2016; Macpherson, 1973; Przeworski, 1985; Sørensen, 1993). This expanded form of democracy, extends 'the democratic principle from the political to the social, in effect primarily economic, realm' (Przeworski, 1985, p. 7). In the terms deployed here, the concept of social democracy thus includes the provision of social and economic welfare and the progressive realisation of economic and social rights. It could also be argued that it includes the protection of cultural rights, which are concerned with such issues as mother tongue language, ceremonial land rights, and intellectual property rights relating to cultural practices (e.g. indigenous healing practices and remedies that may be of interest to multinational companies).

In their modern manifestation, human rights have become an accepted legal and normative standard through which to judge the quality of human dignity (Landman & Carvalho, 2009). This standard has arisen through the concerted efforts of thousands of people over many years inspired by a simple set of ideas that have become codified through the mechanism of public international law and realized through the domestic legal frameworks and governmental institutions of states around the world (Landman, 2005a, 2005b; Landman & Carvalho, 2009). While the 1948 Universal Declaration of Human Rights makes reference to the right to take part in government (including through direct or indirect representatives, equal access to public services, and through periodic elections), the non-binding nature of the Universal Declaration of Human Rights along with a paucity of specific reference to democracy itself in subsequent international human rights instruments, means that human rights as such have been more legally codified through international human rights law than democracy.

According legal recognition to the moral claim of human rights through international law means that states are legally obliged to ensure that they respect, protect, and fulfil these claims (see, e.g. Koch, 2005). There is no corresponding legal obligation to respect, protect, and fulfil democracy in the same way as there is for rights, which provides a stronger foundation and core content for human rights than for democracy. As we shall see empirically, however, democracy is a form of government that appears superior to other forms of government for protecting, respecting and fulfilling human rights obligations. Respecting human rights requires the state to refrain from violating them. Protecting human rights requires the state to prevent the violation of human rights by 'third' parties, such as private companies, non-governmental organisations, paramilitary and insurgency groups, and 'uncivil' or undemocratic movements (see Payne, 2000). Fulfilling human rights requires the states to invest in and implement policies for the progressive realisation of human rights (Koch, 2005; Landman & Carvalho, 2009; Landman & Kersten, 2016).

Civil and political rights protect the 'personhood' of individuals and their ability to participate in the public activities of their countries. Economic, social and cultural rights provide individuals with access to economic resources, social opportunities for growth and the enjoyment of their distinct ways of life, as well as protection from the arbitrary loss of these rights. Solidarity rights seek to guarantee for individuals access to public goods like development and the environment, and some have begun to argue, the benefits of global economic development (Freeman, 2017; Landman, 2006; Landman & Carvalho, 2009). Taken together, there are now a large number of human rights that have been formally codified, which can be enumerated from the different treaties that have been designed to protect them.

In following Beetham (1999, p. 94) and the brief discussion of democracy and human rights, it is clear that different conceptions of democracy vary precisely around the question of the degree of overlap and interaction between the institutional and rights dimensions. Beetham (1999, p. 94) visualises this overlap as a Venn diagram with democracy in one circle and human rights in another, where different definitions and conceptualisations of democracy necessarily reflect smaller and larger degrees of overlap (see Figure 1). Thin or procedural definitions of democracy afford less space for human rights than thicker or social definitions, while it may be possible to conceive of some attributes of human rights sitting outside the conceptual space of democracy. By thinking of the association between democracy and human rights in this way, Beetham (1999) avoids the problem that democracy and human rights might be construed as mutually constitutive of one another while retaining the notion that they are 'inter-dependent and mutually reinforcing' (United Nations, 1993). Hill (2016) makes the case that respect for personal integrity is a *sine qua non* for the existence of democracy and argues that democracy and human rights are thus mutually constitutive. In the terms set out here, however, Hill's (2016) argument only focuses on physical integrity rights, which means that his conception of democracy sees a permanent overlap between the institutional dimension of democracy and this more limited set of human rights, which typi-

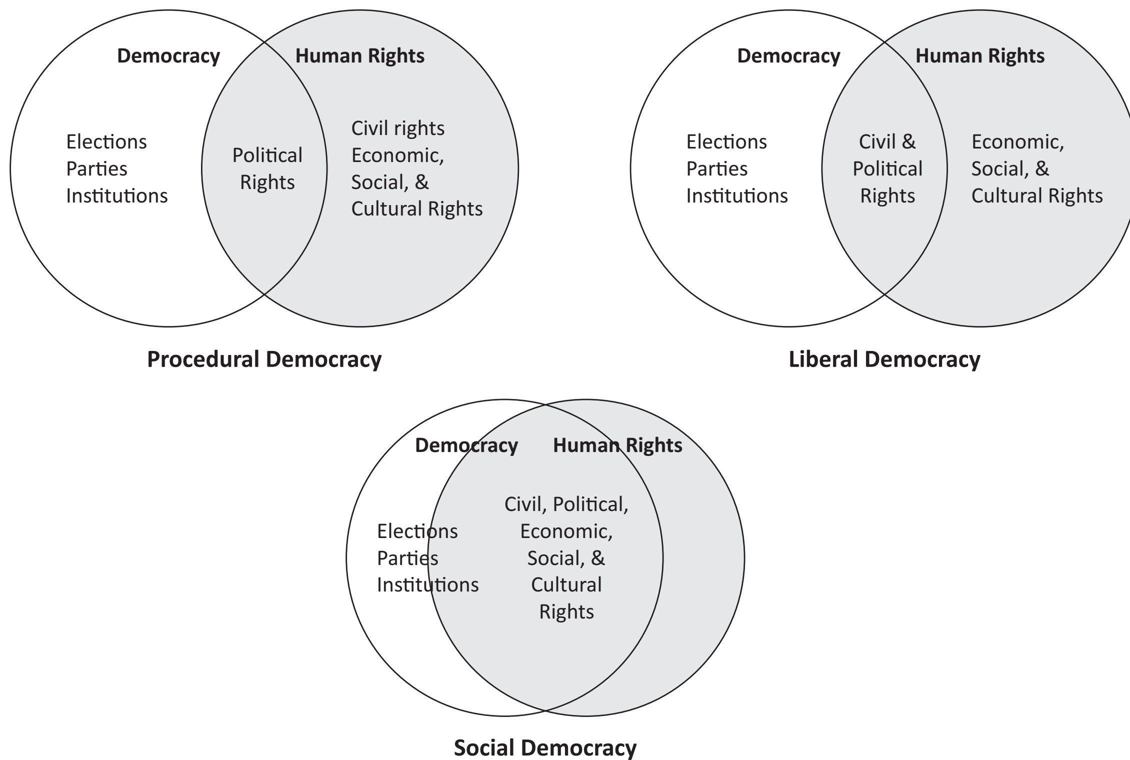


Figure 1. Definitions of democracy (adapted from Beetham, 1999, p. 94).

cally include freedom from torture, arbitrary detention, extra-judicial killing, and exile (see Poe & Tate, 1994). It is not clear from the literature on democracy or human rights that human rights beyond this more limited set are indeed necessarily part of the concept of democracy. Where Hill (2016) is correct is with respect to the endogeneity problem in the empirical analysis of the relationship between democracy and human rights as we shall see in subsequent discussions below.

The possibility of different definitions and different degrees of overlap necessarily affects the ways in which both concepts are measured and analysed (Coppedge, 2012); however, there has not been much discussion about this particular issue in the measurement literature (see Munck, 2009), since there are discussions on the measurement of democracy or human rights, but not democracy *and* human rights. Moreover, discussions of the measurement of democracy, as well as the empirical operationalisation of democracy include measures that are arguably more about human rights than democracy *per se*. For example, in her review of existing measures of democracy, Dorenspleet (2015) includes scales produced by Freedom House, where the checklists for at least one of the scales focuses almost exclusively on human rights. Helliwell (1994) combines these two Freedom House measures arithmetically and calls the combination an 'index of democracy', a move which necessarily commits him to a specific concept of democracy and inclusion of some human rights but not all. These and other tensions in the measurement of democracy and human rights are discussed in turn.

3. Measurement Strategies

The measurement of democracy and human rights has progressed significantly since the early modernization literature as seen in the seminal studies from Seymour Martin Lipset (1959) on the world and Daniel Lerner (1958) on the Middle East. Simple dichotomous coding schemes, although still adopted in some cross-national research (Przeworski et al., 2000) have given way to more complex formats that that seek to capture different dimensions of democracy (Coppedge et al., 2016) and an expanding set of human rights categories beyond civil and political rights to include economic and social rights (Fukuda-Parr, Lawson-Remer, & Randolph, 2015; Landman, 2002, 2006; Landman & Carvalho, 2009; Landman & Häuserman, 2003; Landman & Kersten, 2016). This development in measures has also included an expansion in the different types of data used to measure the two concepts, including events data, standards data, survey data, socio-economic and administrative data, and increasingly, so-called 'big' data (Landman & Kersten, 2016).

3.1. Events

Events-based data answer the important questions of what happened, when it happened, and who was involved, and then report descriptive and numerical summaries of the events. For human rights, counting such events and violations involves identifying the various acts of commission and omission that constitute or lead to human rights violations, such as extra-judicial killings,

arbitrary arrest, or torture. Such data tend to be disaggregated to the level of the violation itself, which may have related data units such as the perpetrator, the victim, and the witness (Ball, Spirer, & Spirer, 2000; Landman, 2006, pp. 82–83; Seybolt, Aronson, & Fischoff, 2013). Events data are used less frequently in research on democracy, but Lindberg (2006) used the number of elections as an indicator of the growth of democracy in Africa alongside other attributes of democracy. Other democratic events can include transitions from authoritarian rule as in the large literature on 'waves' of democracy (see Doorenspleet, 2005; Huntington, 1991; Landman, 2013; Landman & Carvalho, 2017). In his work on democratic performance, Lijphart (1994, 1999, 2012) incorporates a number of events data to judge the relative merits of consensus and majoritarian democracies, but these events are not 'democratic' per se. Rather, they are measures of government performance more generally and are hypothesised as areas of performance that should be (1) superior among democracies and (2) differentiated between consensus and majoritarian forms of democracy.

3.2. Standards

Standards-based measures of democracy and human rights are one level removed from event counting and/or violation reporting and merely apply an ordinal scale to qualitative information, where the resulting scale is derived from determining whether the reported democratic or human rights situation reaches particular threshold conditions. Standards-based scales have been the workhorse of cross-national research in comparative politics, development studies, international political economy, and international relations. One of the major challenges with standards-based scales has been the multiplicity of their use, where such scales are used as measures of democracy, human rights, the 'repressiveness of the regime', the rule of law, and 'good governance' (Foweraker & Landman, 1997; Landman, 2005a; Landman & Hauserman, 2003; Muller & Seligson, 1987). There has been a hasty and particular readiness to use such measures without careful reflection on the concepts that underpin them, the attributes that inform their coding, and the potential overlaps between democracy and human rights that arise (see also Munck, 2009).

There are prominent examples of standards-based measures of democracy. In the Cross-Polity Time-Series Data Archive, Arthur S. Banks provided standards-based scales of different institutional attributes of democracy (see Foweraker & Landman, 1997). The scales were coded for the presence of these attributes and can be totalled for a democracy score that measures the narrow form of procedural democracy. The Polity IV data set provides standards-based measures of different democratic attributes, and focusses on the constraints on the regulation, openness and competitiveness of the executive branch, alongside constraints on the executive and the regulation and competitiveness of participation. Like Banks, these different attributes can be analysed separately (see Buena de Mesquita, Downs, Smith, & Cherif, 2003) or used together to form a scale that ranges from autocracy to democracy (Marshall, Gurr, & Jaggers, 2016). In similar fashion, the 'scale of polyarchy' (Coppedge & Reinicke, 1988) indicators of freedom of expression, freedom of organization, media pluralism, and the holding of fair elections; where this approach has influenced the approach taken in the much expanded 'varieties of democracy' project.

For human rights, the most prominent standards-based examples include the Freedom House scales of civil and political liberties (Gastil, 1980; www.freedomhouse.org), the 'political terror scale' (Poe & Tate, 1994), a scale of torture (Hathaway, 2002), and a series of seventeen different rights measures collected by Cingranelli and Richards (www.humanrightsdata.com). Freedom House has a standard checklist it uses to code civil and political rights based on press reports and country sources about state practices and then derives two separate scales for each category of rights that range from 1 (full protection) to 7 (full violation). The political terror scale ranges from 1 (full protection) to 5 (full violation) for state practices that include torture, political imprisonment, unlawful killing, and disappearance. Information for these scales comes from the US State Department and Amnesty International country reports. In similar fashion, Hathaway (2002) measures torture on a 1 to 5 scale using information from the US State Department. The Cingranelli and Richards human rights data codes similar sets of rights on scales from 0 to 2, and 0 to 3, with some combined indices ranging from 0 to 8, where higher scores denote better rights protections. In addition to a series of civil and political rights, Cingranelli and Richards also provide measures for such rights as women's economic, social, and political rights, worker rights, and religious rights.

One of the key issues that emerged concerning these human rights scales has been the level of awareness in general about human rights and whether an increased awareness and expectation of accountability for human rights violations would increase the reporting of observations of human rights and thus make the world appear worse off over time than was actually the case. Christopher Fariss (2014) addresses this issue head on through the use of item-response theory (IRT) and applies it to existing human rights scales. The intuition behind the item-response theory is that discerning the location of a country on the scale involves a judgment about the degree to which a country meets the threshold condition to move it from one category in the scale to the next. Fariss (2014) finds that when the scales are readjusted for this process of discernment and raised expectation about human rights accountability the overall picture of human rights remains positive over time. Even though the world is more conscious of human rights (itself a function of successful advocacy by human rights organisations), the

underlying trends in human rights abuse over the last three decades have seen a gradual improvement.

3.3. Surveys

There are countless survey projects on democracy, in terms of electoral studies and public attitudes, support for democracy, support for democratic institutions, satisfaction with democracy and voter intention among other dimensions of democracy. Large survey projects like the World Values Survey and the 'Barometer' projects for Europe, Africa, Latin America and Asia have all used random samples and structured survey frameworks that have been used for primary and secondary analysis of citizen attitudes across wide range of concerns relevant to democracy. In addition to random sample approaches, there are 'expert judgement' surveys on democracy, such as the electoral integrity project (Norris, 2017), the Bertelsmann Transformation Index (www.bti-project.org) and the Varieties of Democracy project (see, e.g., Lührmann, Lindberg, & Tannenberg, 2017).

Household surveys have been used to provide measures for popular attitudes about rights and to uncover direct and indirect experiences of human rights violations. Some of the most notable work has been carried out by the NGO Physicians for Human Rights, which conducts surveys of 'at risk' populations (e.g. internally displaced people or women in conflict) to determine the nature and degree of human rights violations (www.physiciansforhumanrights.org). The 'minorities at risk' project certainly captures the degree to which communal groups and other national minorities suffer different forms of discrimination (www.mar.umd.edu). In addition, truth commissions, such as East Timor (www.chegareport.net) and Sierra Leone (www.sierraleonetrc.org) have carried out retrospective household mortality surveys on all deaths and illnesses during the periods under investigation. These surveys are then used alongside events data in ways that allow for better estimations of the total number of people killed or disappeared during periods of conflict, occupation, or authoritarian rule. In similar fashion, Anderson, Paskeviciute, Sandovici and Tverdova (2005) use survey data alongside standards-based measures of human rights to compare the perceptions and attitudes on human rights in Eastern Europe to reported human rights conditions (see Landman & Carvalho, 2009, pp. 91–106).

3.4. Socio-Economic and Administrative Statistics

Administrative and socio-economic statistics produced by national statistical offices or recognized international governmental organizations have been increasingly seen as useful sources of data for the indirect measure of human rights, or as indicators for rights-based approaches to different sectors, such as justice, health, education, and welfare. Government statistical agencies and intergovernmental organizations produce a variety of socio-economic statistics that can be used to approximate measures of human rights. For example, academic and policy research has used aggregate measures of development as proxy measures for the progressive realization of social and economic rights. Such aggregate measures include the Physical Quality of Life Index (PQLI), the Human Development Index (HDI) and the Social Economic Rights Fulfilment Index (SERF Index). the SERF Index (www.serfindex.org) measures on a 0 to 100 scale the extent to which states fulfil their obligations under the right to food (infant height and weight), the right to education (primary school completion, gross school enrolment, average math and science PISA score), the right to health (contraceptive prevalence, life expectancy, infant mortality), the right to housing (improved access to sanitation and water), and the right to work (poverty headcount, long-term unemployment, relative poverty) in relation to countries' maximum available resources.

The World Bank's governance indicators have collated a panoply of different indicators from which six main dimensions of governance are derived (Kaufmann, Kraay, & Mastruzzi, 2009). Tatu Vanhanen has dedicated his life's research to the growth and development of democracy and his main measure the Index of Democratization is comprised of 'objective' measures of his key dimensions of democracy that draw on Dahl (1971): participation and competition. For participation, he uses the official electoral turnout of the population. For competition, he uses the size of the smallest political party in the legislative chamber. His index then multiplies these two dimensions, which he argues captures the essence of democracy (see, e.g. Vanhanen, 2003).

More interestingly, Foweraker & Krznaric (2003) use different official statistics to differentiate democratic performance of established democracies in the West. They argue that extant measures of democracy like the Polity IV measure provide very little indication of the variation in established democracies, and lead to the conclusion that these democracies and their performance are both uniform and superior to other new and restored democracies (Foweraker & Krznaric, 2003, p. 314). They show that these established democracies are not necessarily uniform, and that there are deficiencies in civil and minority rights protections, such as women's representation, equal access to the law, and political discrimination against minorities, as well as disproportionately high incarceration rates (Foweraker & Krznaric, 2003, p. 327). These problem areas are particularly acute in the US, the UK and Australia (Foweraker & Krznaric, 2003, pp. 327–332), while their overall conclusions shed considerable insight into the variation in well-established democracies particularly in the realm of human rights (see below).

3.5. New Forms of Data

In addition to the continued development and refinement of these existing measurement strategies, there

are new trends in data collection that make use of the 'democratization of technology' that has taken place more or less during the first decade of the twenty-first century. The rise of social media and the increasing availability of smartphones and other mobile devices has led to a revolution in the ability of individual people to have a voice in ways that were hitherto not possible. User-generated content on the Internet, in the form of 'tweets', YouTube videos, SMS alert networks, and other platforms of information dissemination, have created a volume of information on country conditions that is beginning to transform the ability of political scientists and other researchers to study human rights. The information that is now available is 'double edged': on the one hand, it provides the ability for grassroots reporting and narrative accounts of real time events as they unfold, and on the other hand, it provides 'meta data' on the events themselves, as smart technology often contains automatic functions that include the date, time and location that something has happened (typically through embedded 'global positioning system' technology, or GPS).

The combination of real time data and meta data allows for collection, fusion, and visualization of democracy and human rights events across space and time, often at the 'street corner' level of accuracy. The collection of these kinds of data occurs in two ways: (1) 'crowd sourcing' through specialized data collection 'portals' such as the platform made available through Ushahidi (www.ushahidi.com), or (2) collection of data from already existing 'open data' sources, such as Facebook, Twitter, news media, and NGO reporting, among others. In their raw form, the data are not particularly useful, but, through fusing different sources into well-structured data bases that conform to the 'who did what to whom' understanding of human rights violations, they can be used for human rights assessments of countries. Moreover, since the meta data may contain additional information about date, time, and location of events, it is possible to map violations on publicly available mapping programmes, such as Google Maps.

4. Empirical Relationships

The theoretical connections and overlaps set out above show that it is not unreasonable to expect strong empirical relationships between democracy and human rights. Democracies are meant to be based on the protection of fundamental rights and thus there is an expectation that human rights protections will be higher in democracies than non-democracies or that the protection of human rights will co-vary with the level of democracy. Large scale cross-national comparative analyses that specify civil and political rights protection as the dependent variable tend to use a narrow and procedural definition of democracy as an independent variable (see, e.g. Landman, 2005a, 2005b; Landman & Carvalho, 2016; Mitchell & McCormick, 1988; Poe & Tate, 1994) in an effort to minimise the problem of endogeneity. Such studies show that democracy and human rights are indeed positively correlated with one another but not perfectly so. From the first cross-national study by Mitchell & McCormick (1988) to the latest pooled-cross section time-series models on human rights protection, there is a significant relationship between democracy and human rights.

For example, Table 1 shows the correlations between democracy and human rights using a variety of measures across a different selection of country cases and time (Landman, 2005a, p. 110, 2016, p. 144). The first row in the table includes correlations for the Polity IV measure of democracy and different measures of human rights for a sample of 194 countries between 1976 and 2000, which are all statistically significant and indicate varying magnitudes in the relationship between democracy and human rights. The negative signs for these correlations are due to the fact the democracy scores are coded low for non-democracy and high for full democracy, while human rights scores (with the exception of Cingranelli and Richards) are coded low for good human rights protection (or low levels of violation) and high for bad human rights protection (or high levels of violation). The second row is a set of correlations for a sample of 21 Latin American countries between 1981 and 2010 (Landman, 2016, p. 144) using the same Polity IV variable and a slightly different set of human rights measures, including the Amnesty International derived Political Terror Scale, US State Department derived Political Terror Scale, and the Cingranelli and Richards (CIRI) physical integrity rights scale.[1] While all these correlations are significant at the $p < .001$ level of significance, the magnitude varies from relatively low values to high values suggesting that

Table 1. Correlations between democracy and human rights.

	PTS (Amnesty)	PTS (US State)	Civil Liberties (F House)	Political Rights (F House)	Torture (Hathaway)	Physical Integrity (CIRI)
World 1976–2000[1]	−.36*** (.000)	−.41*** (.000)	−.85*** (.000)	−.91*** (.000)	−.34*** (.000)	
Latin America 1981–2010[2]	−.27*** (.000)	−.38*** (.000)				.35*** (.000)

Notes: Pairwise correlations, p-values in parentheses, *$p < .10$; **$p < .05$; ***$p < .001$; Landman (2005a, 2016, p. 144).

[1] The coding for this scale is the inverse of what is used across the other scales and thus is positively correlated with the Polity IV measure of democracy.

democracy and human rights certainly co-vary, but not perfectly so.

This variation in the magnitude of the relationship between the particular measure of democracy and these different human rights measures suggests several things. First, these measures are indeed measuring different things. The Polity IV measure primarily captures the institutional dimension of democracy, while the human rights measures focus on a narrow set of physical integrity rights (the Political Terror Scale and the CIRI scale) and torture (Hathaway, 2002) or on broader sets of civil and political rights (Freedom House), where there is much more conceptual (and therefore empirical) overlap between democracy and human rights. Indeed, the coding checklists for the Freedom House measures include attributes most commonly associated with democracy. Second, the gap between democracy and human rights evident in correlations that are less than a perfect 1 capture the notion of what Larry Diamond (1999) and Fareed Zakaria (2003) have called ‘illiberal democracies’, where it is perfectly possible for democracies to hold elections, have peaceful transfers of power between civilian leaders, and functioning legislatures while at the same time being unable to prevent the violation of certain human rights (see Beetham et al., 2008; Landman, 2016). To demonstrate this point further, it is possible to combine standards-based measures of civil and political rights seen in Table 1 above into a single factor score and then plot this factor score against the Polity IV measure of democracy. There are strong and significant factor loadings ranging from .684 to .909 for each of five measures of human rights on a single extracted factor component that is common to all measure (see Landman & Larizza, 2009, p. 721). Figure 2 shows a scatter plot between the Polity IV measure of democracy and this human rights factor (see Landman, 2013, p. 39; Landman, Kernohan, & Gohdes, 2012; Landman & Larizza, 2009), where it is clear that there is a positive and significant re-

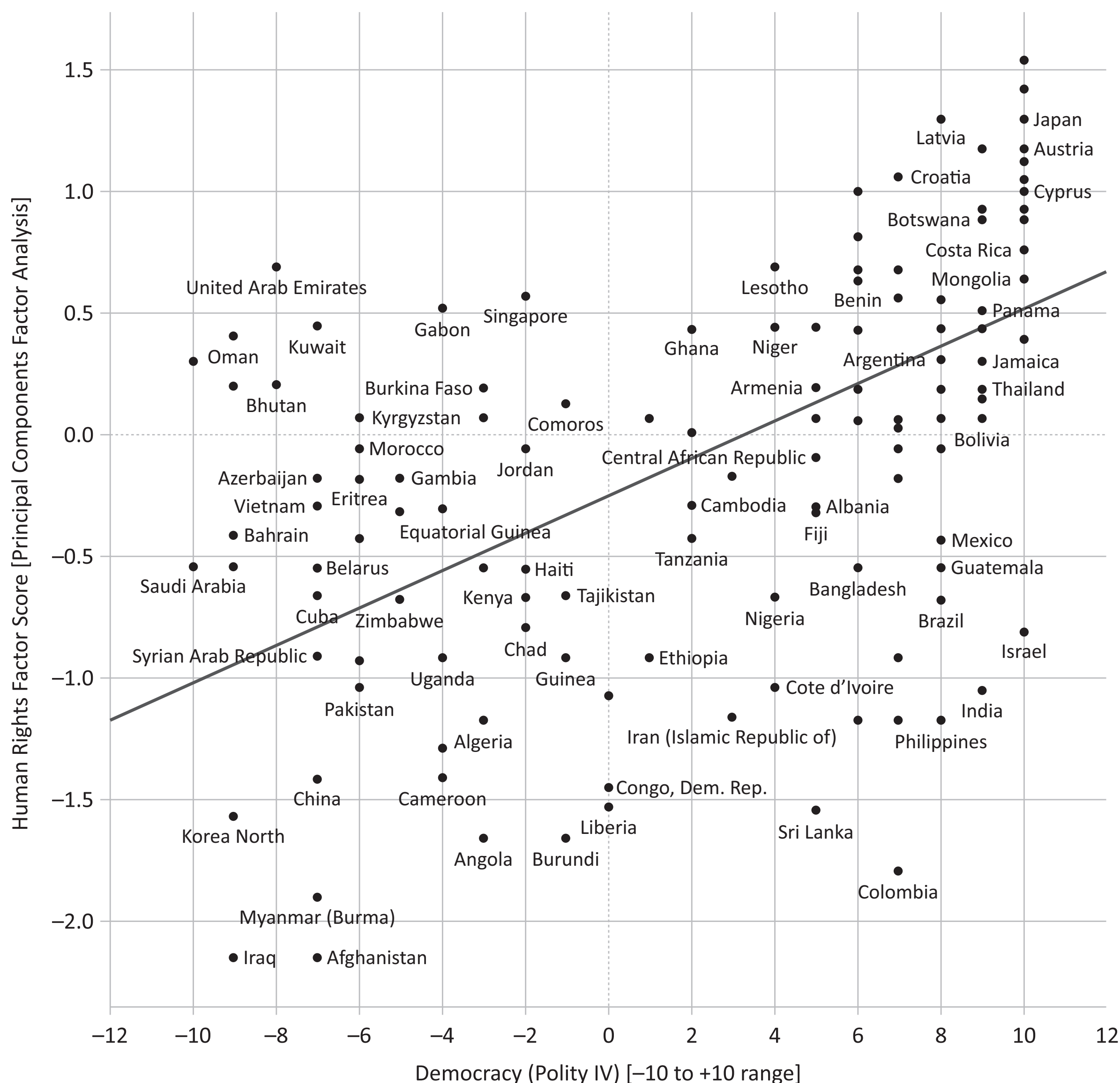


Figure 2. Democracy and Human Rights (2000).

lationship between the two measures (captured by the fitted line). It is also evident from Figure 2 that there is a significant number of countries that would qualify as 'illiberal democracies' sitting in the lower right quadrant (e.g. Brazil, India, the Philippines, and Colombia). These countries score relatively high on democracy but relatively low on their ability to protect human rights. Third, the significant relationships can also be down to an element of human rights sitting within measures of democracy. Indeed, Hill (2016) has shown that democracy measures such as Polity IV have certain limited elements of human rights in them, rendering some empirical analysis between democracy and human rights spurious.

Beyond the relationship between democracy and civil and political rights, Fukuda-Parr et al. (2015, pp. 131–135) show that democracies have a much better record of fulfilling social and economic rights. Their Social and Economic Rights Fulfilment (SERF) Index ranges from 0 (no fulfilment) to 100 (expected fulfilment). They show that the 5th Quintile democracies (using the Polity IV measure of democracy) have a mean score on fulfilling social and economic rights of 80.92 with a low of 56.06 and a high of 94.05, where this range is significantly better than for lower scoring democracies and autocracies (Fukuda-Parr et al., 2015, p. 132). These positive relationships for Polity IV and SERF are also upheld for the World Bank's Worldwide Governance Indicators on 'Voice and Accountability' and 'Rule of Law', and the Freedom House scales of political rights and civil liberties (Fukuda-Parr et al., 2015, pp. 132–133). For my own work on Latin America across 21 republics for the period 1980–2010, the SERF index is positively correlated with the Polity IV measure of democracy (Kendall's Tau B = .241, $p < .000$) (Landman, 2016, pp. 144). Again, as in the relationships between democracy and civil and political rights, the fulfilment of social and economic rights is a function of more than just democracy, and that any relationship is not perfectly correlated. Rather, variation in democracy accounts for some of the variation in social and economic rights fulfilment.

The empirical relationship between democracy and human rights is highly variegated and dependent more on the definitions of democracy that are adopted than human rights, since human rights have been formally articulated through international and domestic law in ways that democracy has not. While there are no agreed philosophical foundations for the existence of human rights, the law of human rights across domestic, regional and international jurisdictions, as well as the jurisprudence that accompany it have provided what human rights lawyers call 'core content' of rights and their obligations. It is this core content and articulation of state obligations that in my view represent a 'systematized concepts' (Adcock & Collier, 2001) that can be operationalized through the different types of data discussed here. In contrast, the concept of democracy relies only on political theory and political philosophy for its core content and has not been 'legalized' in the same way as human rights (Meckled-Garcia, 2005). As we have seen, definitions of democracy vary considerably and variously include different sets of human rights. The positive and significant relationship between democracy and human rights attests to their complementarity, while the remaining gap in the relationship between them confirms that they are different from one another.

The utilization of measures for empirical analysis needs to be consistent in setting out what is (or is to be) measured, compared, and analysed; where any use of measures must be as closely linked to the concepts that they purport to measure. This linkage between concepts and measures involve significant trade-offs between complexity, viability, and validity (Landman & Carvalho, 2009, pp. 24–30). It can be argued that there is a direct and negative relationship between conceptual complexity and measurement viability. Complex conceptual frameworks for measuring democracy and human rights might reduce viability, as complexity raises cost and faces challenges of data availability and accessibility. The four different types of data outlined here—events, standards, surveys, socio-economic and administrative—can and have been variously to capture part or most of each concept depending on the purpose of the empirical analysis (Landman & Carvalho, 2009, p. 29).

5. Challenges and Opportunities

Democracy and human rights are complex, multi-faceted and multi-dimensional concepts that are not mutually exclusive from one another. Definitions of democracy variously include both *institutional* dimensions that constrain executives, separate power and authority, and provide mechanisms for accountability, as well as rights dimensions that provide fundamental protections that allow individuals and groups to aggregate their interests, articulate those interests, shield themselves from arbitrary abuses of power, and enjoy the ability to exercise freedom and agency in their public and private lives. The first and crucial step in any systematic effort to compare, measure, and analyse democracy and human rights is to provide precise and coherent definitions of the concepts to be measured and analysed, the boundary conditions for them, and the attributes that comprise them.

It for the reasons of complexity, multi-dimensionality, and variable overlap between democracy and human rights that measurement strategies have been difficult, challenging, and evolving. Different attributes of democracy and human rights can be delineated through different indicators, which can yield different 'scores on units' (Adcock & Collier, 2001, p. 201) that vary across space and over time. Many of these attributes and dimensions are observable, while many are not, where lateral methods, proxy measures, and 'latent class' analytical techniques and probabilistic inferential statistics (such as multiple systems estimation, or MSE) are required. Overt elements of democracy and human rights such as elections and violations can be observed and

counted, while many aspects suffer from what the late Will Moore calls 'the fundamental problem of unobservability', where practices, actions, choices, and interpersonal interactions take place behind closed doors and in secret locations.

The scholarly and practitioner communities working on democracy and human rights have made great strides in developing increasingly nuanced and effective measurement strategies that have captured more of the inherent complexity and multi-dimensionality of democracy and human rights. Events-based data, standards-based data, survey-based data, and socio-economic and administrative statistics are being used in increasingly creative and systematic ways to capture the temporal and spatial variation in democracy and human rights. From Lipset's (1959) original polychotomous coding to the latest release of the Varieties of Democracy (V-Dem) data set, there have been great strides made in the measurement and analysis of democracy, the quality of democracy, and democratic performance. From the early work of Gastil to the latest analysis from Fariss (2014) and the Human Rights Data Analysis Group (HRDAG), there have been significant advances in the measurement and analysis of human rights (see www.hrdag.org).

Despite these many advances, however, many challenges remain. First, there is still the need to work on how democracy and human rights are defined and how those aspects that are unique to each are circumscribed, while greater attention is given to the different ways in which democracy and human rights overlap with one another and how they are related to one another. Second, the specification of systematic definitions of both concepts is directly linked to the ways in which they are measured. Third, there continues to be an over-reliance on subjective coding of subjective information collected on democracy and human rights. Now more than ever, there are increasing types of data being generated that can be harnessed and analysed in ways that can enhance our understanding and explanation of the variation in democracy and human rights. Big data techniques, machine learning and supervised machine learning, web scraping and corpus linguistic analytical techniques offer new ways of measuring, mapping, and understanding democracy and human rights.

Conflict of Interests

The author declares no conflict of interests.

References

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

Anderson, C. J., Paskeviciute, A., Sandovici, M. E., & Tverdova, Y. V. (2005). In the eye of the beholder? The foundations of subjective human rights conditions in East-Central Europe. *Comparative Political Studies*, *38*(7), 771–798.

Ball, P. B., Spirer, H. F., & Spirer, L. (Eds.). (2000). *Making the case: Investigating large scale human rights violations using information systems and data analysis*. Washington, DC: American Association for the Advancement of Science (AAAS).

Banks, A. S. (1971). *Cross-polity time series archive*. Cambridge, MA: MIT Press.

Beetham, D. (1999). *Democracy and human rights*. Cambridge: Polity.

Beetham, D., Carvalho, E., Landman, T., & Weir, S. (2008). *Assessing the quality of democracy: A practical guide*. Stockholm: International IDEA.

Brandal, N., Bratberg, Ø., & Thorsen, D. (2013). *The Nordic model of social democracy*. London: Palgrave Macmillan.

Buena de Mesquita, B., Downs, G. W., Smith, A., & Cherif, F. M. (2003). Thinking inside the box: A Closer look at democracy and human rights. *International Studies Quarterly*, *49*(3), 439–457.

Collier, D., & Levitsky, S. (1997). Democracy with adjectives: Conceptual innovation in comparative politics. *World Politics*, *49*(3), 430–451.

Coppedge, M., Lindberg, S., & Skaaning, S.-E. (2016). Measuring high level democratic principles using the V-Dem data. *International Political Science Review*, *37*(5), 580–593.

Coppedge, M., & Reinicke, W. (1988). A scale of polyarchy. In R. D. Gastil (Ed.), *Freedom in the world: Political rights and civil liberties, 1987–1988* (pp. 101–121). New York, NY: Freedom House.

Coppedge, M. (2012). *Democratization and research methods*. Cambridge: Cambridge University Press.

Dahl, R. A. (1971). *Polyarchy: Participation and opposition*. New Haven, CT: Yale University Press.

Diamond, L. (1999). *Developing democracy: Toward consolidation*. Baltimore, MD: Johns Hopkins University Press.

Doorenspleet, R. (2005). *The fourth wave of democratization: Identification and explanation* (Doctoral dissertation). University of Leiden, Leiden, The Netherlands.

Fariss, C. J. (2014). Respect for human rights has improved over time: Modelling the changing standard of accountability. *American Political Science Review*, *108*(2), 297–318.

Foweraker, J., & Kznaric, R. (2000). Measuring liberal democratic performance: An empirical and conceptual critique. *Political Studies*, *48*(2), 759–787.

Foweraker, J., & Kznaric, R. (2003). Differentiating the democratic performance of the West. *European Journal of Political Research*, *42*(3), 313–340.

Foweraker, J., & Landman, T. (1997). *Citizenship rights and social movements: A comparative and statistical analysis*. Oxford: Oxford University Press.

Freeman, M. (2017). *Human rights* (3rd ed.). Cambridge: Polity.

Fukuda-Parr, S., Lawson-Remer, T., & Randolph, S. (2015). *Fulfilling social and economic rights*. Oxford: Oxford University Press.

Gallie, W. B. (1956). Essentially contested concepts. *Proceedings of the Aristotelian Society*, *56*(1), 167–198.

Gastil, R. D. (1980). *Freedom in the world: Political rights and civil liberties*. Westport, CT: Greenwood Press.

Hathaway, O. (2002). Do treaties make a difference? Human rights treaties and the problem of compliance. *Yale Law Journal*, *111*(4), 1932–2042.

Helliwell, J. (1994). Empirical linkages between democracy and economic growth. *British Journal of Political Science*, *24*(2), 225–248.

Hill, D. (2016). Democracy and the concept of personal integrity rights. *The Journal of Politics*, *78*(3), 822–835.

Huntington, S. (1991). *The third wave: Democratization in the late twentieth century*, Norman, OK: University of Oklahoma Press.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2009). *Governance matters VIII: Aggregate and individual governance indicators* (Policy Research Working Paper). Washington, DC: World Bank.

Koch, I. E. (2005). Dichotomies, trichotomies, or waves of duties? *Human Rights Law Review*, *5*(1), 81–103.

Landman, T. (2002). Comparative politics and human rights. *Human Rights Quarterly*, *43*(4), 890–923.

Landman, T. (2005a). Review article: The political science of human rights. *British Journal of Political Science*, *35*(3), 549–572.

Landman, T. (2005b). *Protecting human rights: A global comparative study*. Washington, DC: Georgetown University Press.

Landman, T. (2006). *Studying human rights*. Oxford and New York, NY: Routledge.

Landman, T. (2009). *Human rights (volumes I–IV)*. London: Sage.

Landman, T. (2013). *Human rights and democracy: The precarious triumph of ideals*. London: Bloomsbury Academic.

Landman, T. (2016). Democracy and human rights: Explaining variation in the record. In J. Foweraker & D. Trevizo (Eds.), *Democracy and its discontents in Latin America* (pp. 121–132). Boulder, CO: Lynne Rienner.

Landman, T., & Carvalho, E. (2009). *Measuring human rights*. London and Oxford: Routledge.

Landman, T., & Carvalho, E. (2017). *Issues and methods in comparative politics*. London and Oxford: Routledge.

Landman, T., & Häusermann, J. (2003). *Map-making and analysis of the main international initiatives in developing indicators of democracy and good governance*. Luxembourg: Eurostat.

Landman, T., Kernohan, D., & Gohdes, A. (2012). Relativising human rights. *Journal of Human Rights*, *1*(4), 460–485.

Landman, T., & Kersten, L. (2016). Measuring human rights. In M. Goodhart (Ed.), *Human rights: Politics and practice* (3rd ed., pp. 127–144). New York, NY: Oxford University Press.

Landman, T., & Larizza, M. (2009). Inequality and human rights: Who controls what when and how. *International Studies Quarterly*, *53*(3), 715–736.

Lerner, D. (1958). *The passing of traditional society: Modernizing the Middle East*. Glencoe, IL: The Free Press of Glencoe.

Lijphart, A. (1994). Democracies: Forms, performance, and constitutional engineering. *European Journal of Political Research*, *25*(1), 1–17.

Lijphart, A. (1999). *Patterns of democracy: Government forms and performance in thirty-six countries*. New Haven, CT: Yale University Press.

Lijphart, A. (2012). *Patterns of democracy: Government forms and performance in thirty-six countries* (2nd ed.). New Haven, CT: Yale University Press.

Lindberg, S. (2006). *Democracy and elections in Africa*. Baltimore, MD: Johns Hopkins University Press.

Lipset, S. M. (1959). Some social requisites for democracy: Economic development and political legitimacy. *The American Political Science Review*, *53*(1), 69–105.

Lührmann, A., Lindberg, S., & Tannenberg, M. (2017). *Regimes in the world (RIW): A robust regime type measure based on V-Dem* (Working Paper. Series 2017: 47). Gothenburg: Varieties of Democracy Institute.

Macpherson, C. B. (1973). *Democratic theory: Essays in retrieval*. Oxford: Clarendon Press.

Marshall, M., Gurr, T. R., & Jaggers, K. (2016). *Polity IV project: Political characteristics and transitions, 1800–2015*. Retrieved from http://www.systemicpeace.org/inscrdata.html

Meckled-Garcia, S. (2005). *The legalisation of human rights: Multidisciplinary approaches*. Oxford: Routledge.

Mitchell, N. J., & McCormick, J. M. (1988). Economic and political explanations of human rights violations. *World Politics*, *40*(3), 476–498.

Muller, E. N., & Seligson, M. A. (1987). Inequality and insurgency. *American Political Science Review*, *81*(2), 425–451.

Munck, G. L. (2009). *Measuring democracy: A bridge between scholarship and politics*. Baltimore, MD: Johns Hopkins University Press.

Norris, P. (2017). *Strengthening electoral integrity*. New York, NY: Cambridge University Press.

Payne, L. (2000). *Uncivil movements: The armed right wing and democracy in Latin America*. Baltimore, MD: Johns Hopkins University Press.

Poe, S. C., & Tate, C. N. (1994). Repression of human rights to personal integrity in the 1980s: A global analysis. *American Political Science Review*, *88*(4), 853–872.

Przeworski, A. (1985). *Capitalism and social democracy*. Cambridge: Cambridge University Press.

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

Seybolt, T. B., Aronson, J. D., & Fischoff, B. (Eds.). (2013). *Counting civilian casualties: An introduction to recording and estimating nonmilitary deaths in conflict*. Oxford: Oxford University Press.

Sørensen, G. (1993). *Democracy and democratization*. Cambridge: Cambridge University Press.

United Nations. (1993). *UN General Assembly, Vienna Declaration and Programme of Action*. Retrieved from http://www.refworld.org/docid/3ae6b39ec.html

Vanhanen, T. (2003). *Democratization: A comparative analysis of 170 countries*. London: Routledge.

Zakaria, F. (2003). *The future of freedom: Illiberal democracy at home and abroad*. New York, NY: W.W. Norton.

About the Author

Todd Landman is Professor of Political Science and Pro Vice Chancellor of the Faculty of Social Sciences at the University of Nottingham. He is author of *Issues and Methods in Comparative Politics* (4th ed., Routledge 2017, with Edzia Carvalho), *Democracy and Human Rights: The Precarious Triumph of Ideals* (Bloomsbury 2013), *Measuring Human Rights* (Routledge 2009, with Edzia Carvalho), *Studying Human Rights* (Routledge 2006), *Protecting Human Rights: A Comparative Study* (Georgetown 2005), and *Citizenship Rights and Social Movements: A Comparative and Statistical Analysis* (Oxford 1997, with Joe Foweraker). He has carried out a large number of international consultancies on development, democracy and human rights.

Politics and Governance (ISSN: 2183–2463)
2018, Volume 6, Issue 1, Pages 60–77
DOI: 10.17645/pag.v6i1.1214

Article

Regimes of the World (RoW): Opening New Avenues for the Comparative Study of Political Regimes

Anna Lührmann *, Marcus Tannenberg and Staffan I. Lindberg

V-Dem Institute, Department of Political Science, University of Gothenburg, 405 30 Gothenburg, Sweden;
E-Mails: anna.luehrmann@v-dem.net (A.L.), marcus.tannenberg@gu.se (M.T.), xlista@gu.se (S.I.L.)

* Corresponding author

Submitted: 3 October 2017 | Accepted: 12 January 2018 | Published: 19 March 2018

Abstract
Classifying political regimes has never been more difficult. Most contemporary regimes hold *de-jure* multiparty elections with universal suffrage. In some countries, elections ensure that political rulers are—at least somewhat—accountable to the electorate whereas in others they are a mere window dressing exercise for authoritarian politics. Hence, regime types need to be distinguished based on the *de-facto* implementation of democratic institutions and processes. Using V-Dem data, we propose with Regimes of the World (RoW) such an operationalization of four important regime types—closed and electoral autocracies; electoral and liberal democracies—with vast coverage (almost all countries from 1900 to 2016). We also contribute a solution to a fundamental weakness of extant typologies: The unknown extent of misclassification due to uncertainty from measurement error. V-Dem's measures of uncertainty (Bayesian highest posterior densities) allow us to be the first to provide a regime typology that distinguishes cases classified with a high degree of certainty from those with "upper" and "lower" bounds in each category. Finally, a comparison of disagreements with extant datasets (7%–12% of the country-years), demonstrates that the RoW classification is more conservative, classifying regimes with electoral manipulation and infringements of the political freedoms more frequently as electoral autocracies, suggesting that it better captures the opaqueness of contemporary autocracies.

Keywords
autocracy; democracy; democratization; regime; typology

Issue
This article is part of the issue "Why Choice Matters: Revisiting and Comparing Measures of Democracy", edited by Heiko Giebler (WZB Berlin Social Science Center, Germany), Saskia P. Ruth (German Institute of Global and Area Studies, Germany), and Dag Tanneberg (University of Potsdam, Germany).

© 2018 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Classifying political regimes has never been more difficult. Most regimes in the world hold *de-jure* multiparty elections with universal suffrage. In some countries, elections ensure that political rulers are—at least somewhat—accountable to the electorate whereas in others they are a mere window dressing exercise for authoritarian politics. Therefore, we need to base regime classification on the *de-facto* implementation of democratic institutions and processes. This is key to being able to make a meaningful distinction between electoral democracies and electoral autocracies. Such data is provided by the Varieties of Democracy (V-Dem) Project, which covers 177 countries from 1900 to 2016 (Coppedge et al., 2017a, 2017b). While V-Dem primarily provides interval measures, many important research questions require crisp regime measures. For instance, categorical measures of regimes have been used in studies on democracy aid effectiveness (Lührmann, McMann, & Van Ham, 2017), inquiries of democratic diffusion (Gleditsch & Ward, 2006), backsliding (Erdmann, 2011),

sequencing (Wang et al., 2017), characteristics of authoritarian regimes (Schedler, 2013), and regime survival (e.g. Bernhard, Hicken, Reenock, & Lindberg, 2015; Przeworski, Alvarez, Cheibub, & Limongi, 2000; Svolik, 2008).

We use the V-Dem data to classify countries into four regime categories. In closed autocracies, the chief executive is either not subjected to elections or there is no meaningful, *de-facto* competition in elections. Electoral autocracies hold *de-facto* multiparty elections for the chief executive, but they fall short of democratic standards due to significant irregularities, limitations on party competition or other violations of Dahl's institutional requisites for democracies. To be counted as electoral democracies, countries not only have to hold *de-facto* free and fair and multiparty elections, but also—based on Robert Dahl's famous articulation of "Polyarchy" as electoral democracy (Coppedge, Lindberg, Skaaning, & Teorell, 2016; Dahl, 1971, 1998)—achieve a sufficient level of institutional guarantees of democracy such as freedom of association, suffrage, clean elections, an elected executive, and freedom of expression. A liberal democracy is, in addition, characterized by its having effective legislative and judicial oversight of the executive as well as protection of individual liberties and the rule of law.

Although the typology is widely accepted (e.g. Diamond, 2002; Rössler & Howard, 2009; Schedler, 2013), comprehensive, longitudinal measures have not been available until now. Regimes of the World (RoW) closes this gap by classifying virtually all country-years from 1900 to today based on this typology. In addition, we provide an innovative method to address a key weakness in extant typologies: identifying ambiguous cases close to the thresholds between regime types using V-Dem's measures of uncertainty. This additional information can be integrated into quantitative analyses, for instance by allowing scholars to conduct robustness checks which exclude more ambiguous cases.

Section two discusses prior approaches to regime types while the third section details the RoW typology. Section four compares our regime typology to several of the most frequently used extant measures.

2. Prior Approaches to Drawing the Line between Regime Types

Longstanding conceptual and methodological discussions include whether democracy is a best understood as a multidimensional (Coppedge et al., 2011; Dahl, 1971; Vanhanen, 2005), continuous (Bollen & Jackman, 1989; Lindberg, 2006), polychotomous (Collier & Levitsky, 1997), or a dichotomous concept (Alvarez, Cheibub, Limongi, & Przeworski, 1996; Cheibub, Gandhi, & Vreeland, 2010), as well as debate the precise differentiation between democratic and various types of autocratic regimes (Geddes, Wright, & Frantz, 2014; Kailitz, 2013; Wahman, Teorell, & Hadenius, 2013), including the existence of a "grey zone" (Diamond, 2002). We agree with Collier and Adcock (1999) that the appropriate type of regime measure depends on the nature of the research question at hand. We seek here to provide a robust and comprehensive regime type measure for research requiring an ordinal or a dichotomous measure.

There are two main approaches to conceptualizing and crafting dichotomous measures of democracy and autocracy: as a difference in kind or as a difference in degree, which are associated with qualitative and quantitative approaches to measurement, respectively (Lindberg, 2006, pp. 22–27). The in-kind/qualitative approach typically proceeds in a Sartorian fashion by setting a number of necessary conditions that a regime must fulfill in order to be coded as a democracy. For example, that there are competitive, multiparty elections with suffrage extended to a certain share of the population. The degree/quantitative strand usually introduces a cut-off on a continuous measure of democracy, coding countries above the threshold as democratic and countries below the threshold as being autocratic. In the following, we provide details regarding how six of the most influential datasets on regimes distinguish between democracies and autocracies.

2.1. In-Kind/Qualitative Approaches

Cheibub et al. (2010) apply three criteria to distinguish democracies from autocracies: uncertainty, irreversibility, and repeatability.[1] Operationally, they identify democracies as regimes in which there are, first, more than one legal party; second, a legislature elected by popular elections, and a chief executive that is either directly, or indirectly popularly elected; and finally, an alternation of power must have occurred under the same electoral rules that brought the incumbent into office. While these clear and parsimonious coding rules minimize the need for subjective judgments, they also come at a cost. Two of these criteria raise concerns of conceptual validity. The mere existence of two legal parties hardly guarantees contestation, as understood in established democratic theory (Dahl, 1971), and the alternation rule leads to both type I and type II errors. First, as Wahman (2014) shows, it underestimates the number of democracies since incumbents often enjoy an electoral advantage even in established democracies. Second, even manipulated and un-democratic elections are sometimes lost, which leads to the alternation rule overestimating the number of autocracies (Wahman, 2014, p. 222). These errors have consequences. For example, Knutsen and Wig (2015) demonstrate that the alternation rule leads to the underestimation of democracy's effect on economic growth.

Geddes et al. (2014) sort all cases into either the democratic or autocratic bin (before proceeding to clas-

[1] Their "Democracy and Dictatorship" dataset builds on earlier work by an overlapping group of authors (Cheibub, Przeworski, Limongi Neto, & Alvarez, 1996; Przeworski et al., 2000).

sify sub-categories of the latter). They stipulate the following coding rules: a case is coded as democratic if the executive achieves power through "reasonably fair competitive" direct or indirect elections with suffrage exceeding at least 10% of the population (Geddes, Wright, & Frantz, 2013, p. 6). This requires a fair amount of judgment by the coder. For example, relying on reports from election observers to determine if an election was reasonably "fair and competitive" can be problematic since such organizations lack shared standards (Kelley, 2009). It is not clear what a "competitive" election or "large" party is by Geddes et al. (2014)'s standards (see Geddes et al., 2013, p. 6), nor is it clear how Geddes et al. (2014) estimate the size of parties which did not enjoy legal rights (Wahman et al., 2013).

Boix, Miller and Rosato (2013) provide a dichotomous measure of democracy/autocracy from 1800 to 2010. Similar to Cheibub et al. (2010) and Geddes et al. (2014), Boix et al. (2013) rely on a set of necessary conditions. For a country to be coded as democratic, the executive must either be directly or indirectly elected in "popular" elections and the legislature in "free and fair" elections. They also require that a majority of the male population has the right to vote. Boix et al. (2013) suffers from a similar weakness as Geddes et al. (2014)—they asses the freedom and fairness of elections without minimizing bias due to the potentially erroneous judgment of the coder.

2.2. Degree/Quantitative Approaches

Other scholars apply a threshold on a continuous measure to distinguish between political regimes (Lindberg, 2016; Schedler, 2013; Wahman et al., 2013). The most apparent difficulty with this approach is deciding where to draw the line between democracies and autocracies, which is inevitably, an arbitrary decision (Bogaards, 2012). Even for the most commonly used large-N data sets—Freedom House and Polity—there is no consensus in the literature on where to draw the line. Bogaards (2012) identifies at least 14 different ways to use Freedom House ratings and at least 18 different ways to use the Polity scores to classify democracies.

Freedom House itself uses its political rights and civil liberty scores to label countries as "free", "partly free", and "not free" (Freedom House, 2017). However, this three-level ordinal scale evades the question of which "partly free" country is a democracy and which not. Furthermore, it neglects any necessary conditions—such as free and fair elections—that are commonly found in the literature. Similarly, the Polity project (Marshall, Gurr, & Jaggers, 2014) provides various detailed assessments of different aspects of regime quality, but refrain from identifying an unambiguous cut-off point between democracy and autocracy. Polity suggests using the combined Polity score to cut the regime spectrum into three parts: autocracies (−10 to −6), democracies (6 to 10), and anocracies, with anocracies being between the first two categories.[2]

Wahman et al. (2013) identify the cut-off point on a combined Freedom House and Polity scale that best represents five qualitative democracy measures, such as the ones we discussed above. They proceed by estimating the mean score on the combined scale for the year before democratic breakdown and the year after transition, as coded by the five measures. They then use the grand mean of seven of these years as their empirical cut-off point for democracy, while advising users to run robustness checks using both the 6.5 and the 7.5 levels.

Scholars addressing the whole regime spectrum have come to distinguish, typically, between closed and electoral autocracies on one hand and liberal and electoral democracies on the other hand (e.g. Diamond, 2002; Rössler & Howard, 2009; Schedler, 2013) which has become one of the most prolific typologies in the discipline, as well as in the policy-practitioners' world. Nevertheless, we lack comprehensive, longitudinal measures of this four-fold regime typology.[3] Below, we suggest a way to fill this gap while simultaneously avoiding the weaknesses of the current measures which have been outlined above.

3. The RoW Typology

Following this brief review of some of the extant regime typologies, we endeavor to classify regimes into four categories: closed autocracy, electoral autocracy, electoral democracy and liberal democracy (Table 1). First, we separate along the democratic and the autocratic regime spectrum and then develop the democratic and autocratic subtypes.[4] In a minimalist, Schumpeterian sense, democracies are regimes that hold *de-jure* multiparty elections. However, many would agree with Pastor (1999, p. 123) that "the essence of democratic government is accountability". Such accountability can only evolve if incumbents fear retribution at the ballot box (Mechkova, Lührmann, & Lindberg, 2017a), and to this end, mere *de-jure* multiparty elections are not enough (e.g., Levitsky & Way, 2010; Schedler, 2013). We claim that Dahl's theory of polyarchy (1971, 1998) provides the most comprehensive and most widely accepted theory of what distinguishes a democracy based on six (1998, p. 85—originally p. 8 in his 1971 book) institutional guarantees (elected officials, free and fair elections, freedom of expression, alternative sources of information, associational autonomy, and inclusive citizenship). This conception requires not only free and fair elections but also the freedoms that make them meaningful, and thus avoids the electoral fallacy (Diamond, 2002; Karl, 1986). This allows for demarcation between electoral autocracies and

[2] See http://www.systemicpeace.org/polityproject.html
[3] Typically, scholars use data from sources such as Freedom House and/or the Database on Political Institutions, which only starts in the 1970s.
[4] This strategy follows common advice for concept formation (e.g., Collier & Adcock, 1999, pp. 548–549; Goertz, 2006; Sartori, 1970).

Table 1. Regime classification.

Closed Autocracy	Electoral Autocracy	Electoral Democracy	Liberal Democracy
No *de-facto* multiparty, or free and fair elections, or Dahl's institutional prerequisites not minimally fulfilled		*De-facto multiparty, free and fair elections, and* Dahl's institutional prerequisites minimally fulfilled	
No multiparty elections for the chief executive or the legislature	*De-jure* multiparty elections for the chief executive and the legislature	The rule of law, or liberal principles not satisfied	The rule of law, and liberal principles satisfied

democracies, unlike minimalist definitions. In short, in democracies rulers are *de-facto* accountable to citizens through periodic elections and in autocracies they are not.[5] Therefore, we approach *de-facto* multiparty and free and fair elections as necessary, qualitative criteria for labelling a regime as a democracy.

We distinguish between electoral democracies that only achieve the basic criteria above, and liberal democracies. We focus on this distinction because it is the most common within the democratic regime spectrum (e.g. Diamond, 1999, 2002; Merkel, 2004; Munck, 2009). In addition to fulfilling the criteria for electoral democracy, liberal democracies are characterized by an additional set of individual and minority rights beyond the electoral sphere, which protect against the "tyranny of the majority"; thus having limits on government is intrinsic to democracy itself (e.g. Dahl, 1956; Hamilton, Madison, & Jay, 1787/2009; cf. Coppedge, Gerring, Lindberg, Skaaning, & Teorell, 2017d, p. 21; Lindberg, Coppedge, Gerring, & Teorell, 2014). This is in Dahl's words "Madisonian" democracy (Dahl, 1956, p. 4). Core components thus include legislative and judicial oversight over the executive providing checks and balances, as well as the protection of individual liberties, including access to, and equality before, the law. In particular, the rule of law is a fundamental prerequisite for the implementation of the liberal principle as it ensures that decisions are implemented (Merkel, 2004).

Autocracies are regimes where rulers are not accountable to citizens by Dahl's standards. The key differences along the authoritarian spectrum are whether the office of the chief executive and seats in the national legislature are subject to direct or indirect multiparty elections (Schedler, 2013, p. 2). In closed autocracies, the chief executive and the legislature are either not subject to elections, or there is no *de-facto* competition in elections such as in one-party regimes. Regimes with elections that do not affect who is the chief executive (even if somewhat competitive) also fall into this category (following Brownlee, 2009; Donno, 2013; Rössler & Howard, 2009, p. 112).

In electoral autocracies, on the other hand, the chief executive is dependent on a legislature that is itself elected in *de-jure* multiparty elections (in parliamentary systems), directly elected alongside a separately elected legislature (in presidential systems), or a combination of both (in semi-presidential systems). In an electoral autocracy, these institutions are *de-facto* undermined such that electoral accountability is evaded (Diamond, 2002; Gandhi & Lust-Okar, 2009; Levitsky & Way, 2010; Schedler, 2002, 2013). They thus fall short of democratic standards due to significant irregularities, limitations on party competition, or other violations of Dahl's institutional requisites. This conceptualization builds on Schedler's influential work on electoral authoritarianism (2002, 2006, 2013) and the notion of competitive authoritarianism developed by Levitsky and Way (2010).

3.1. Operationalization with V-Dem Data

We operationalize the RoW regime typology using data from Varieties of Democracy (V-Dem).[6] Version 7.1 covers 178 countries from 1900 to 2016 (Coppedge et al., 2017a, 2017b, 2017c; Marquardt & Pemstein, 2017; Pemstein et al., 2017). Figure 1 portrays the stepwise decision rules. To qualify as an (electoral) democracy, regimes must fulfil three necessary conditions. (1) *De-facto* multiparty elections as indicated by a score above 2 on the V-Dem indicator for multiparty elections (v2elmulpar_osp); (2) free and fair elections where mistakes and irregularities did not affect the outcome, as indicated by a score above 2 on the respective V-Dem indicator (v2elfrfair_osp);[7] and (3) following Lindberg (2016, p. 90), a score larger than 0.5 on the V-Dem Electoral Democracy Index (EDI, v2x_polyarchy) which explicitly measures Dahl's institutional *de-facto* guarantees, based on 41 indicators (Coppedge et al., 2016, 2017a).[8] The index runs from 0 (not democratic) to 1 (fully democratic).

These coding rules strike a balance between two principles in operationalization: substitutability and necessity. In line with Coppedge et al. (2011), we treat Dahl's list of institutions as partly substitutable. A score larger

[5] This reflects the electoral principle of democracy (Coppedge et al., 2016, p. 3).

[6] This operationalization will be included in the V-Dem dataset version 8 under the variable name "v2x_regime" (to be released in Spring 2018).

[7] The V-Dem measurement model converts expert scores to interval-level point estimates (Pemstein et al., 2017). We use a version of the data in which these interval-level estimates were converted to the original 0–4 scale, which is indicated by the suffix _osp.

[8] The aggregation rule for the EDI allows for one strong sub-component to partially compensate for weaknesses in others, but also penalizes countries weak in one sub-component according to the "weakest link" argument. Thus, the index is formed in one half by the weighted average of its component indices and in the other half by the multiplication of those indices (Coppedge et al., 2016, 2017a).

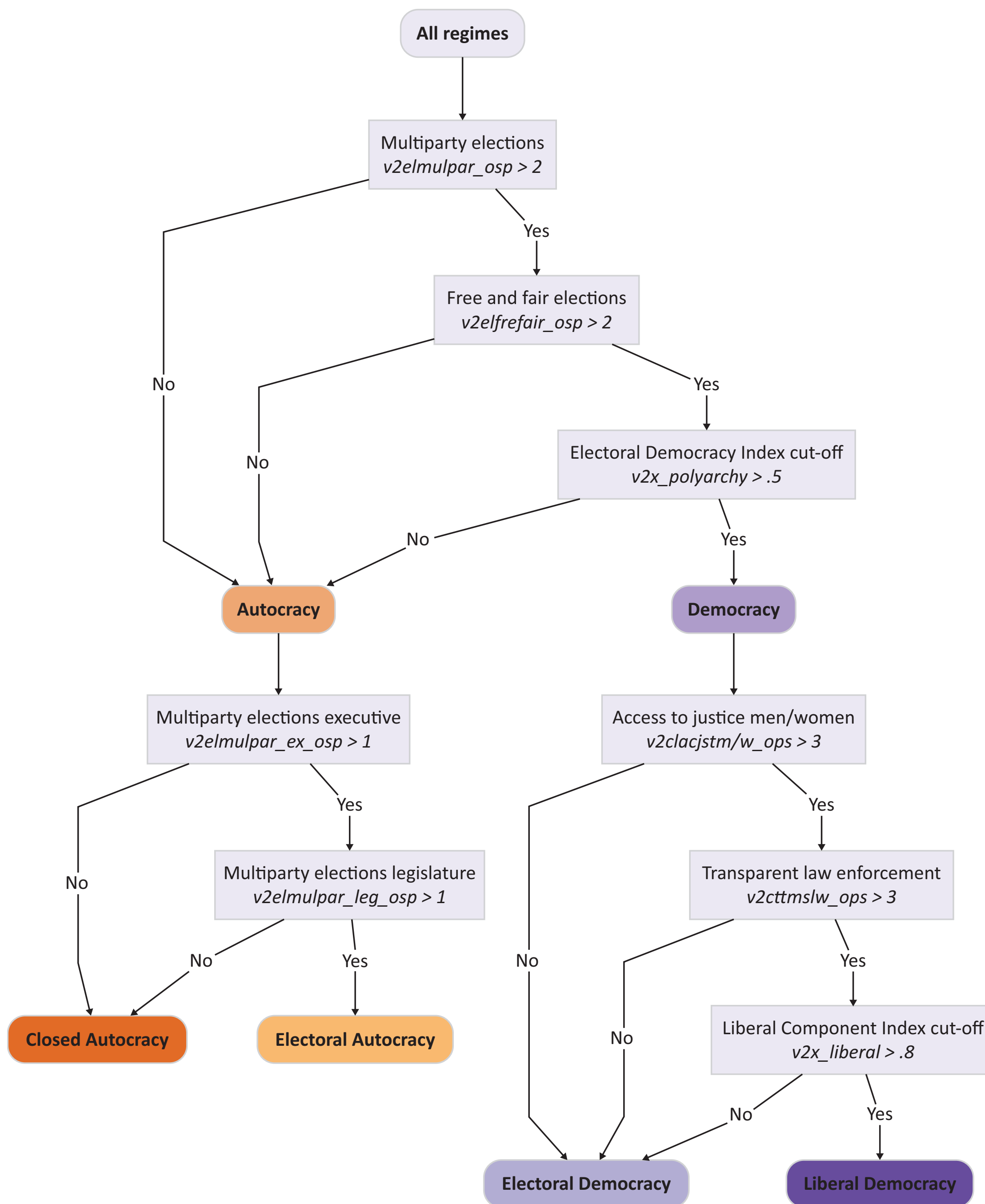


Figure 1. Coding schema for the RoW typologies (for descriptions of variables see Coppedge et al., 2017a).

than 0.5 on the EDI, demonstrates that the balance of potential weaknesses in one area is partly compensated for by strengths in other areas to such a degree that the regime may be classified as being more democratic than not.[9] Yet, in the conceptualization above, given the aggregation of 41 indicators, it remains possible that a country could reach a level above 0.5 on the index while still lacking two critical aspects: *de-facto* multiparty elections and the ability of such an election result to be resistant to the effect of irregularities and unintentional mistakes. We approach *de-facto* free and fair and multiparty elections as necessary conditions. Hence, even while these two indicators are also part of the EDI among the 39 other indicators, we ensure that the—admittedly arbitrary—cut-off point on a continuous scale does not lead to the misclassification of regimes as democracies by combining it with the two key qualitative democracy indicators.

Among dictatorships identified by their failure to meet one or more of the criteria of democracies, electoral autocracies are distinguished from closed dictatorships in that they subject the chief executive to elections

[9] For examples underscoring the empirical validity of these cut-off points, see the detailed discussion in section 4.

at least *de-jure* multiparty competition as indicated by a score above 1 on the applicable V-Dem multiparty elections indicator (v2elmulpar_osp; see Appendix 1).

What distinguishes electoral and liberal democracies is that the latter guarantee the three key aspects of the liberal dimension of democracy discussed above. We operationalize this notion with three necessary criteria. First, liberal democracies need to satisfy three qualitative criteria focusing on the ultimate guarantees of individual liberty: Scores above "3" on the V-Dem indicators transparent and predictable law enforcement (v2cltrnslw_osp), and secure and effective access to justice for men (v2clacjstm_osp) and women (v2clacjstw_osp).[10] While the non-arbitrary enforcement of laws is a prerequisite for the implementation of rules in the first place, access to justice gives individuals the chance to challenge arbitrary enforcement patterns (Botero & Ponce, 2011). In order to further guarantee that no country is undeservedly classified as a liberal democracy, we require a liberal democracy to overall satisfy the liberal principles of respect for personal liberties and the rule of law, and judicial as well as legislative constraints on the executive, as indicated by a score above 0.8 on the summary V-Dem Liberal Component Index (v2x_liberal).[11] Corresponding to the distinction between democracies and autocracies, the inclusion of an aggregated index in the operationalization rules allows for the substitution of weaknesses in one area with strength in others. The threshold is naturally arbitrary but setting it high, at the upper quartile of the scale, seeks to ensure that the criterion adheres to the fairly strict demands expressed in the literature in the liberal tradition.

3.2. Accounting for Ambiguity: Lower and Upper Bounds of Regime Categories

A principal objection leveraged against quantitative approaches to measuring regime types is that countries close to thresholds between categories may be misclassified due to measurement error and uncertainty (e.g. Boix et al., 2013). However, qualitative approaches face the same issue (e.g. Alvarez et al., 1996). The only difference is that we do not know how close or far away a case if from the threshold since they are based on assessments of—often individual—coders with unknown thresholds and unreported uncertainty. The thresholds in quantitative approaches are often more transparent and the consequences of varying them can be tested (Lindberg, 2016, p. 81), but without confidence intervals around point estimates, we still do not know which cases may be misclassified regardless of threshold.

We suggest a major advance on current categorizations in this regard by incorporating into the RoW typology "grey-zone" categories of ambiguously classified cases as indicated by confidence intervals from the underlying Bayesian aggregation methods.[12] The V-Dem dataset provides not only point estimates for indices and variables but also demarcates the interval in which V-Dem's custom-designed Bayesian item-response theory measurement model places 68% of the probability mass for each country-year score. These are calculated slightly different at the indicator and index-level but provide the rough equivalent of one standard deviation confidence interval on either side of the points estimates.[13]

We use these intervals to identify cases that are close to the thresholds between categories, those which are ambiguously classified. If the V-Dem Bayesian highest posterior density interval for an indicator or an index used for the categorization into the four main regime types, overlaps with the threshold of an adjacent category, then the case is classified as ambiguous.

For example, Macedonia's score on the EDI was 0.53 in 2016, slightly above the threshold for electoral democracy (0.5). The values for the two qualitative criteria for elections (multiparty, 3.9; and free and fair, 3.0) are also above the thresholds electoral democracy. However, the lower bound of the EDI score (v2x_polyarchy_codelow) for Macedonia is 0.48 and thus falls within the range of electoral autocracy. Hence, in the RoW typology, we label the country as being "Electoral Democracy Lower Bound", to reflect this ambiguity. Our classification is corroborated by credible reports that freedom of expression has been restricted in Macedonia in recent years.[14]

[10] In an earlier version of this article (and in the V-Dem Data Set v7), we did not include these additional qualitative criteria for liberal democracies. As a consequence, some countries with dubious respect for liberal principles met the threshold for liberal democracies (e.g. Hungary and Tunisia in 2016; the United States prior to the improvements in civil rights in 1968). In order to make our measure of liberal democracy a more ambitious reflection of democratic "completeness"—to borrow from Welzel (2013, p. 255)—we opted to include additional criteria.

[11] This index gives the average of following indices on a scale from 0 (not at all satisfied) to 1 (satisfied): equality before the law and individual liberties (v2xcl_rol), judicial constraints on the executive (v2x_jucon), and legislative constraints on the executive (v2xlg_legcon) (Coppedge et al., 2017a, p. 47).

[12] This operationalization is included in the V-Dem Data Set v8 under the variable name "v2x_regime_amb" (to be released in Spring 2018). We are grateful to Valeriya Mechkova for suggesting this approach to using the Bayesian highest posterior density intervals.

[13] For each *indicator*, V-Dem provides upper and lower bound estimates, which represent 68% of the highest posterior densities (distribution mass), i.e., a range of most probable values for a given observation. The intervals increase with the degree of ambiguity in the raw, expert-coded data. At the indicator-level, mainly three factors influence the size of the intervals: high levels of disagreement between expert coders, a low number of coders, and the presence of coders with relatively low estimated reliability (i.e., high stochastic error variance). V-Dem uses Bayesian Factor Analysis (BFAs implemented with the R package MCMCpack) to aggregate indicators to mid-level indices, such as the Clean Election Index. In the BFA framework, the size of the area covered by the 68 highest posterior densities of mid-level *indices* increases in size if underlying indicators show low levels of correlation. The BFAs are run over 900 posterior draws from the indicators. As a result, uncertainty about indicators also influences the size of the interval in which the modeling places 68% of the probability mass of the mid-level indices. Similar logic applies for top-level indices (such as the Electoral Democracy Index, and the Liberal Component Index), which combine several mid-level indices (Marquardt & Pemstein, 2017; Pemstein et al., 2017).

[14] On the recent developments in Macedonia see BBC (http://www.bbc.com/news/world-europe-36031417) and the European Digital Rights Association (https://edri.org/huge-protest-against-corruption-surveillance-in-macedonia).

Similarly, Poland lost its status as a liberal democracy in 2013 when the point estimate for one of the qualitative indicators (the transparent law enforcement indicator; v2cltrnslw_osp) dropped below the threshold of 3.0, while the upper bound remains above the threshold. It is therefore classified as an "Electoral Democracy Upper Bound".

The RoW typology thus represents a more transparent ordinal measure of regime types than prior approaches, reflecting the estimates of uncertainty of the underlying data calculated by state-of-the-art Bayesian models. We argue that this is a major advance on extant regime typologies—quantitative or qualitative—which do not report how certain we should be about each classification. This brings together conceptual validity and precision with transparent and systematic incorporation of uncertainty with four "pure" regime types and six upper and lower bound regime categories.

4. Opening New Perspectives in Regime Studies

RoW also provides us with unique opportunities to answer new questions. For example, we can analyse changes over time with regard to the share of countries in these grey zones between the "pure" regime types. Figure 2 demonstrates that almost all countries were unambiguously classified at the beginning of the last century (light-grey line on top of the graph). The level of ambiguity (black dashed line) started increasing from around 1960 and peaked during the 1990s—coinciding with the height of the third wave of democratization identified by Huntington (1992). By 2016, almost 30% of all countries were in one of the ambiguous categories, and 12% fell in the critical grey zone between democracy and autocracy. This is a significant result in itself.

There are two but distinct developments driving this trend. First, an increasing number of countries are in the ambiguous regime categories because they have *de-jure* democratic institutions, but simultaneously undermine their effectiveness. The share of unambiguous regimes dropped from above 94% in 1950 to 70% in 2016. Second, the average distance between the upper and lower bounds of the V-Dem indicators and indices have increased in recent decades, reflecting among other things greater disagreement among coders. This also suggests that it has become harder to unambiguously assess countries' states of affairs, even on the very discreet issues that V-Dem ask country experts to rate. The world is becoming opaquer in terms of regimes.

Figure 3 shows the development over time of the RoW regime types (tinted colors indicate ambiguous categories). Our data allow us to show how the number of regimes in the ambiguous categories has increased over recent decades. Several commentators have also recently expressed concerns about potential backsliding among liberal democracies. This is captured in the RoW measure with the number of liberal democracies declining in the last few years. Furthermore, we can identify two pronounced developments associated with the second half of the third wave of democratization, from around 1990. First, the two intermediate categories between electoral democracies and autocracies have both grown wider. Second, the number of closed autocracies declined sharply. Electoral autocracies and electoral democracies have replaced this regime type. An important implication of these two developments is that an increasingly greater number of countries risk being misclassified by extant measures, thus opening the spectre of biased, or even misleading results. RoW can help us solve that problem.

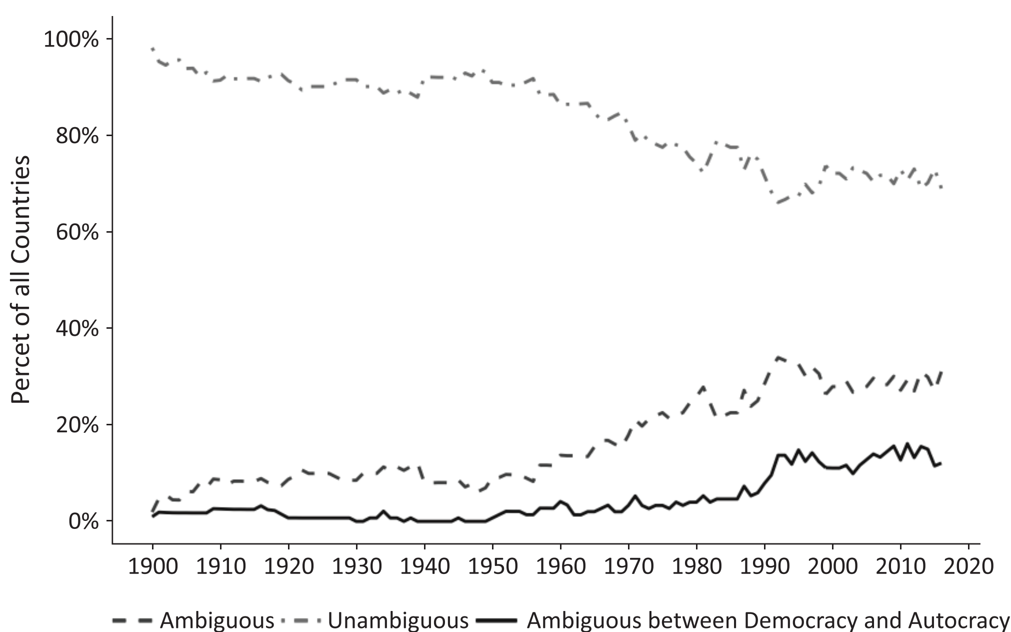


Figure 2. The development of regime ambiguity in the world from 1900 to today.

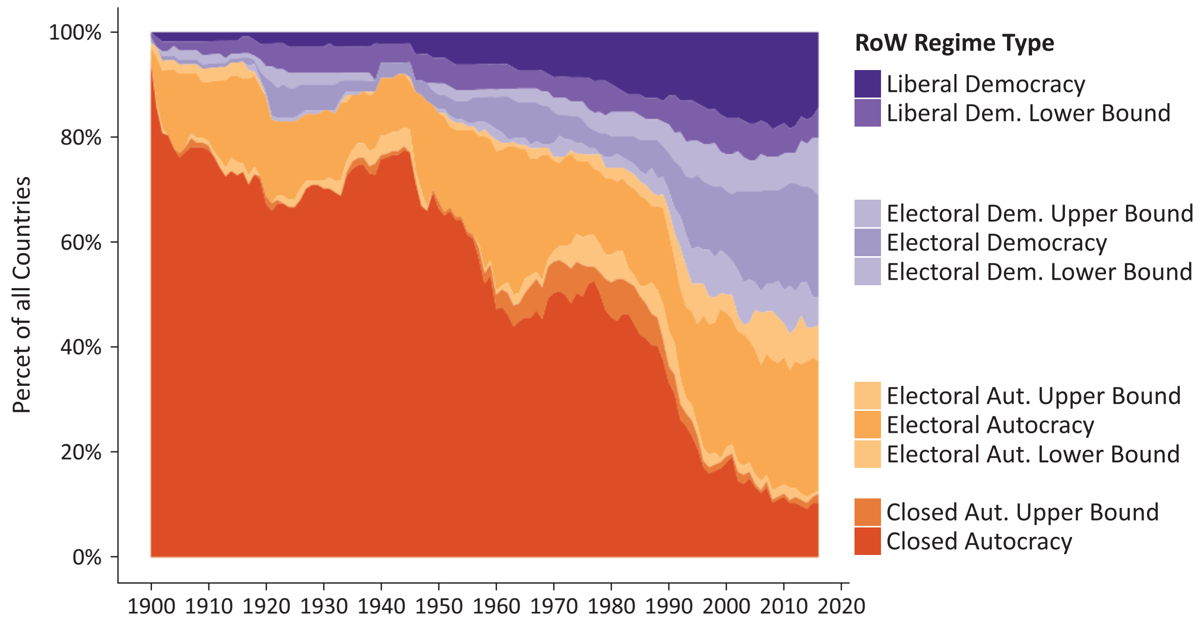


Figure 3. Regimes of the World (RoW) 1900–2016. Source: Coppedge et al. (2017b).

In addition to such descriptive analysis, we recommend the use of ambiguous categories for robustness checks in quantitative analysis. For instance, in democratic survival analyses it makes sense to repeat the analysis varying the in- and exclusion of the ambiguous cases from both the democratic and autocratic regime categories.

5. Comparing RoW to Dichotomous Measures of Democracy

The distinction between democracy and autocracy is arguably the most important aspect of a regime typology. The RoW typology lends itself also to research that requires a dichotomous measure since both the two categories of democracy and autocracy can be collapsed. The lower bounds of (electoral) democracy and the upper bounds of (electoral) autocracy still apply and can be used in combination with the pure regime types in the same fashion as discussed above. In this section, we compare RoW's distinction between democracy and autocracy to the most relevant extant measures, namely those provided by Boix et al. (2013), Cheibub et al. (2010), and Geddes et al. (2014); and Freedom House (Freedom House, 2017), Polity (Marshall et al., 2014), and Wahman et al. (2013).

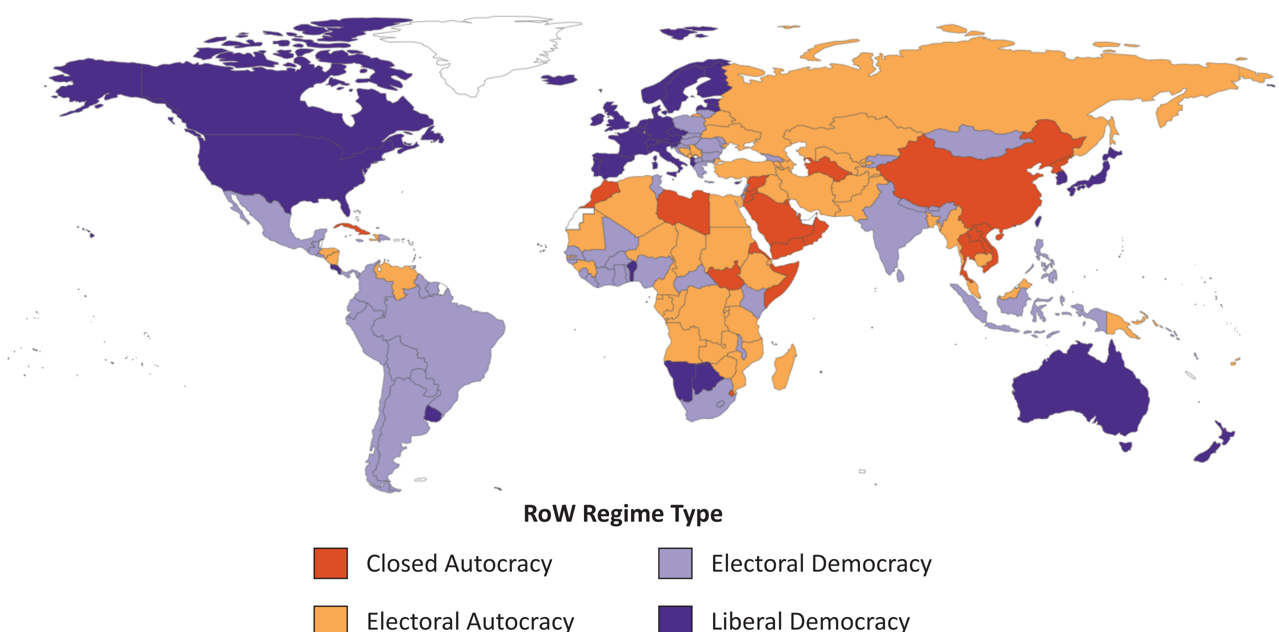


Figure 4. Regimes of the World (RoW) 2016. Source: Coppedge et al. (2017b).

The purpose of this appraisal is two-fold: First, it helps to assess the convergent validity of the RoW measure, one of the most commonly used strategies for new regime type measures (Adcock & Collier, 2001, p. 540). Second, and in line with the spirit of this thematic issue, we seek to make clear what the empirical consequences of the measurement choices are for users of the RoW typology.

A comparison with extant measures, unfortunately, means we have to disregard the ambiguous and pure regime classifications in the RoW measure. Figure 3 first illustrates the distribution of regime types in the world in 2016. Most regimes are in the democratic spectrum (56%): 62 countries qualify as electoral democracies and 35 as liberal democracies (of 174 countries). 56 countries (32%) are electoral autocracies and 21 (12%) are closed autocracies. For a complete list see Appendix 2.

Table 2 compares RoW to the most commonly used measures in the literature. One striking difference is the coverage of RoW, which is matched only by Boix et al. (2013) and Polity, in that it includes all countries and semi-independent territories (including most colonies) from 1900 until the present, and will continue to be updated annually. Cheibub et al. (2010) and Geddes et al. (2014) start in 1946, Wahman et al. (2013) in 1970 and none of them provide data after 2010. While Boix et al. (2013) starts in 1800, it is not updated so the last seven years are not covered. Polity has a longer time series and is updated annually but does not cover semi-independent countries and territories. The fourth column shows that the rate of agreement is relatively high, varying between 88.5% (Cheibub et al., 2010) and 93.1% (Wahman et al., 2013). Excluding the cases which our typology qualifies as ambiguous, the level of agreement varies between 91.7% (Cheibub et al., 2010) and 93.5% (Geddes et al., 2014). When there is disagreement between RoW and other measures, our classification tends to be more conservative and sets a higher bar for what counts as a democracy, i.e., by classifying certain countries as autocracies whereas others place them in the democratic regime spectrum.

The four different panels in Figure 5 display the RoW count of democracies over time compared with the other measures. Three other measures have data prior to 1970—Boix et al. (2013), Cheibub et al. (2010), and Geddes et al. (2014).

The overlap of observations between our RoW measure with Boix et al. (2013) (11, 262 cases) is the second largest, with a rate of agreement on classification in 90.8% of these observations (Figure 5 [A]). The level of agreement increases to 93.5% if we exclude observations that fall into the ambiguous categories according to RoW. In general, Boix et al. (2013) has a lower threshold for democracy: 84% of disagreements on the classification of country-years are due to Boix et al. (2013) classifying them as democracies while they are coded as autocracies by RoW. For example, Boix et al. (2013) code Chile between 1909 and 1949 as democratic even though only 25 to 35% of the adult population were enfranchized due to a lack of female suffrage. Another example is Guatemala, which Boix et al. (2013) and Cheibub et al. (2010) code as democratic following the general election in 1958 up until the onset of civil war in 1981. RoW classifies it as an electoral autocracy. We think the RoW classification has greater face validity since it captures the absence of *de-facto* minimum level of institutional requirements of democracy in this case: Illiterate women were banned from voting up until 1966 (Organi-

Table 2. Comparison of six dichotomous measures to the RoW democracy threshold.

	Country Years	Coverage	Country-Year Overlap with RoW	Agree with RoW	RoW Autocracy Other Democracy	RoW Democracy Other Autocracy
RoW	17140	1900–2016				
Boix et al. (2013)	16988	1800–2010	11,262	90.8% (93.5%)	7.7% (5.8%)	1.5% (0.7%)
Polity	16826	1800–2016	11,394	92.1% (94.3%)	5.4% (4.1%)	2.5% (1.6%)
Cheibub et al. (2010)	9117	1946–2008	8,187	88.5% (91.7%)	8.5% (6.4%)	2.9% (1.9%)
Geddes et al. (2014)	7956	1946–2010	7,688	90.2% (92.8%)	7.4% (5.7%)	2.4% (1.4%)
Wahman et al. (2013)	6279	1970-2010	6,277	93.1% (96.8%)	2.8% (1.3%)	4.0% (1.9%)
Freedom House	6277	1973–2016	6,275	88.8% (93.3%)	1.7% (0.9%)	9.6% (4.6%)

Note: Numbers in brackets are calculated excluding the cases we identified as ambiguous (see section 2). Source: RoW, Coppedge et al. (2017b); Boix et al. (2013); Cheibub et al. (2010); Geddes et al. (2014); Freedom House (2017), Polity (Marshall et al., 2014) and Wahman et al. (2013).

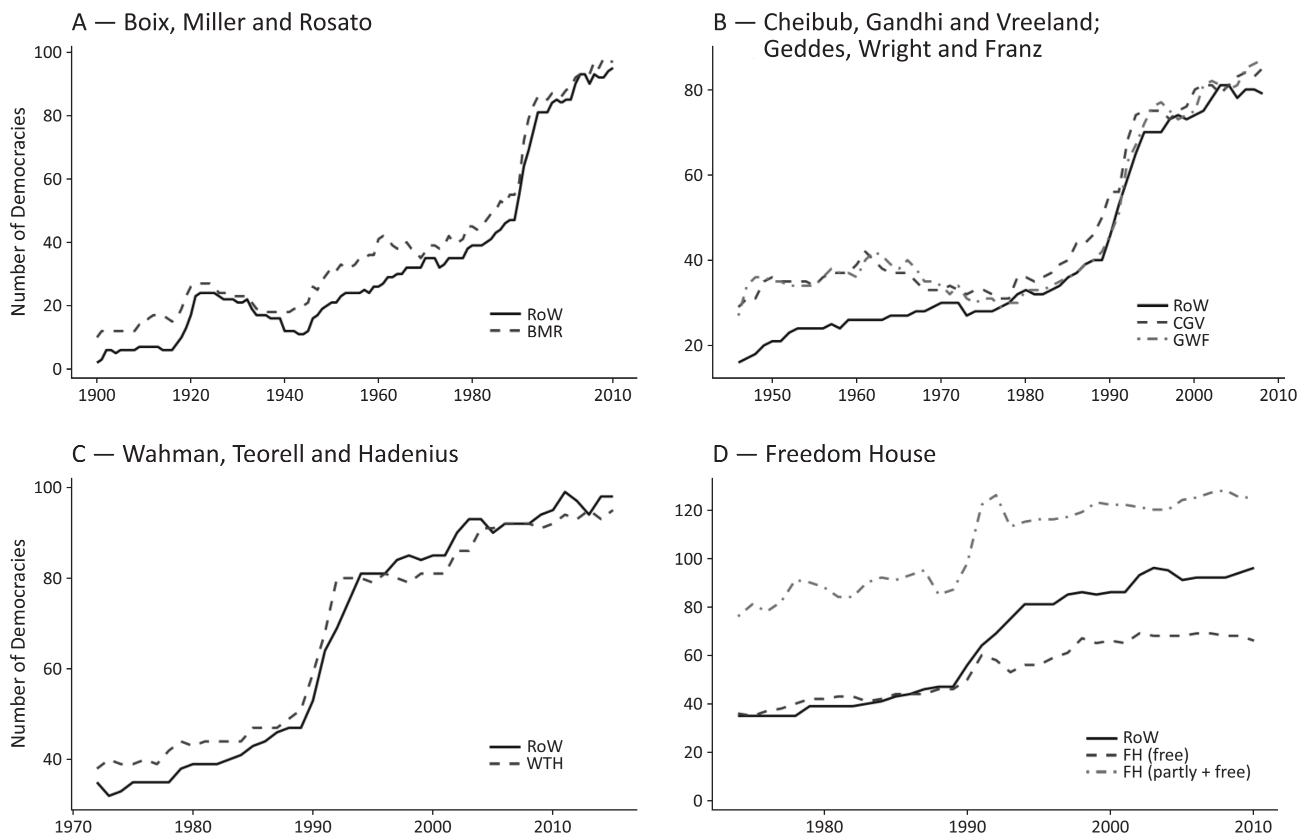


Figure 5. Visual comparison of other measures to RoW. Note: Each panel is limited to the time period and cases of the dataset with least coverage.

zation of American States, 2008), and parties faced severe obstacles to establish themselves and to their participation in elections. Furthermore, electoral intimidation was common throughout the period, and civil society organizations were not free to form or operate. Boix et al. (2013) also codes Czechoslovakia (1939–1945), Norway, Belgium, the Netherlands (1940–1945), and Denmark (1943–1944), as democratic during the years of German occupation whereas RoW does not. Out of the few Boix et al. (2013) autocracies that are coded as democracies in RoW, half are classified as ambiguous cases. Among the unambiguous democracies, RoW captures the dramatic shift from an absolute monarchy to democracy that took place in Bhutan following its first parliamentary elections in 2007/2008, and coded the country as being democratic from 2009 onwards. This is in line with case study evidence (Turner & Tshering, 2014).

Cheibub et al. (2010) provide regime classifications for 9,117 country-years; 8,187 observations overlap with RoW, and the rate of agreement is 88.5% (Figure 5 [B]).[15] Out of the 933 disagreements, 75% (N = 696) are country-years that Cheibub et al. (2010) code as democratic and RoW classifies as autocratic. This discrepancy may be due to Cheibub et al. (2010) using a lower threshold for democracy than RoW. For instance, Cheibub et al. (2010) classifies Kyrgyzstan (2005–2008) and Armenia (1995–2008) as democracies whereas the V-Dem expert-coded *de-facto* indicators capture severe shortcomings in terms of the most basic requirements of democracy such as the freedom and fairness of elections. When RoW classifies countries as democracies which Cheibub et al. (2010) codes as autocratic, it is due to Cheibub et al. (2010)'s controversial alternation rule. For example, Botswana, South Africa, and Namibia are autocracies according to Cheibub et al. (2010) for all years covered, because only one party has been in power since the introduction of multi-party elections. The V-Dem indicators build on indicators that do not require alternation in power, leading to their classification as liberal democracies in the late 1990s. This result is in line with the conclusions of prominent observers (e.g., Diamond, 1999, pp. ix–xxvi).

The RoW measure covers all but 168 of Geddes et al. (2014)'s 7,956 observations,[16] and out of the overlapping country years, the level of agreement of the two measures is 90.2% (92.8% excluding ambiguous cases).

[15] Cheibub et al. (2010) cover 973 observations that are not in the RoW measure. These are mainly microstates not included in the V-Dem Data set: Luxemburg; Andorra, Antigua and Barbuda, Bahamas, Bahrain, Belize; Brunei; Grenada; Kiribati; Lichtenstein; Malta; Marshall Island; Micronesia; Nauru; Palau; Samoa; San Marino; St. Kitts and Nevis; St. Lucia; St. Vincent and the Grenadines; Tonga; Tuvalu; United Arab Emirates. Additionally, Cheibub et al. (2010) covers Oman 1970–1999; Cameroon 1961–1963; and Mozambique 1975–1977, which are not included in V-Dem.

[16] These are the two small states Luxemburg and United Arab Emirates and individual years in Oman (1946–1999), Cameroon (1961–1963), and Mozambique (1976–1977), which are not included in the V-Dem data.

The measures diverge in particular prior to the 1970s and in the early 2000s. Again, when classifying a country as democratic, RoW is more demanding than Geddes et al. (2014) (Figure 5 [B]). For example, Geddes et al. (2014) code Sierra Leone (1999–2002) as democratic even though the then ongoing civil war drastically undermined rules and procedures (Harris, 2014). Similarly, they rate the Central African Republic (1993–2003), Burundi (2005–2010), and Nepal (1991–2002) as democratic, whereas V-Dem's expert-based indicators indicate severe violations or grave deficiencies in the institutional requisites of democracy. In contrast, a number of countries in which V-Dem experts report relatively strong democratic institutions, both *de-jure* and *de-facto*, are coded as autocracies by Geddes et al. (2014): Botswana (1967–2010), Burkina Faso (1993–2010), Ghana (1996–2000), Namibia (1991–2010), and Senegal (1983–2000). The V-Dem coding is consistent with academic assessments of the state of democracy in Ghana (Abdulai & Crawford, 2010), Botswana and Namibia (Diamond, 1999), although some observers of Senegal denote the time period as one of "transition to a fully democratic state" and prefer to label the country as "semi-democratic" (Coulon, 1988; Vengroff, 1993, p. 23). RoW reflects this ambiguity, classifying Senegal as an unambiguous democracy only after the improvements following the 1993 election.

RoW covers all observations in the Wahman et al. (2013) data set with the exception of Mozambique from 1975–1977. Out of all measures compared in this article, Wahman et al. (2013)'s has the highest level of concordance with RoW (Figure 5 [C]; 93.1% or 96.8% excluding ambiguous cases). Wahman et al. (2013) is based on the Freedom House and Polity ratings (see discussion in section 1). When defining only countries that Freedom House codes "free" as democracies the agreement falls to 88.8% or 93.3% when excluding the cases classified as ambiguous in RoW (see Figure 5 [D]). The bulk of the disagreements stem from countries that we classify as democracies but that are "partly free" according to Freedom House. However, lowering the dichotomous threshold to include all "partly free" countries as democracies reduces the concordance to 75.5%, indicating that a majority of countries that Freedom House code as partly free are coded as autocracies in RoW. Hence, overall the agreement between the RoW and Wahman et al. (2013) datasets is greater than when comparing RoW to either Freedom House or Polity separately.

Polity IV is the data source with the greatest number of country-years overlapping with RoW (11,394). Following Marshall et al. (2014)'s suggestion to treat countries above and equal to 6 on the combined Polity scale as democracies, classification agreement with RoW is 92.1%, or 94.3% when excluding ambiguous cases. Most disagreements are once again due to RoW autocracies being coded as democracies in Polity. For example, Polity codes Sweden (1916–1919) and the United States (1900–1919) as perfect democracies (score 10) when women were disenfranchised. RoW classifies these cases as electoral autocracies. In recent years, Burundi (2005–2013), Malawi (1994–2013), and Malaysia (2008–2012) are democracies according to Polity while V-Dem experts observe severe obstacles to democracy. There are also disagreements of other sorts. For example, while Polity codes Suriname as just short of being a democracy (score of 5) since 1991, V-Dem's indicators rate the country as a liberal democracy for the same time period. Figure 6 shows the share of countries coded as RoW democracies for each value on the combined Polity scale. In general, higher values on the Polity scale correspond to a higher share of RoW democracies. The spike in the polity score of 0 is driven by Burkina Faso (2001–2014) and Uruguay (1939–1951), which are classified as democracies in RoW. According to V-Dem coders Burkina Faso had

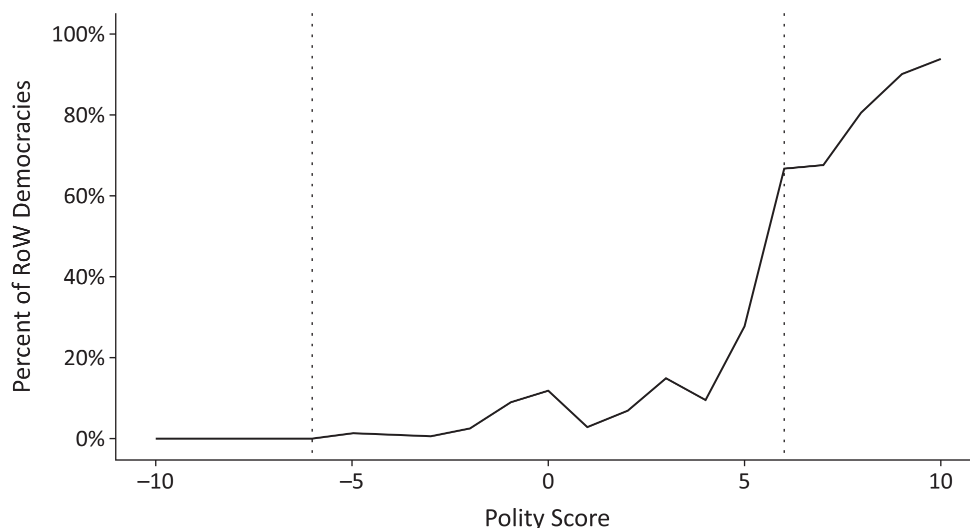


Figure 6. Percentage of RoW democracies by Polity score. Note: The dotted vertical lines mark Polity's suggested thresholds of autocracy (≤ −6) and democracy (≥ 6), with anocracy in between (Marshall et al., 2014). Cases of foreign interruption, interregnum or anarchy, and transitions (polity codes: −66, −77, and −88) are excluded from this comparison.

relatively strong democratic institutions, both *de-jure* and *de-facto* during those years. While Uruguay did not guarantee full political freedoms in the first three years following the dictatorship of Gabriel Terra (1933–1938), the country can indeed be considered a democracy following the introduction of its new constitution in 1942. Similarly, the relatively high value of -1 on the Polity scale is driven largely by Senegal (1983–2000), which according to V-Dem coders had both *de-jure* and *de-facto* democratic institutions.

Overall the new RoW typology relatively closely tracks the classification of country-years as either democratic or autocratic by major extant binary measures of democracy. This is a good sign of convergent validity. However, there are substantial differences concerning a significant number of cases, primarily where *de-facto* practices deviate from *de-jure* standards. We argue that the RoW typology does a better job than others in these instances, discriminating “real” from “fake” democracies.

For example, most other measures code Kenya as democratic[17] in the years following the crisis that erupted after president Kibaki was accused of stealing the December 2007 election (Rutten & Owuor, 2009). Politically-motivated (Kagwanja & Southall, 2009)—and allegedly state-sponsored—violence left more than 1,000 people dead and up to 500,000 people displaced (Human Rights Watch, 2008). RoW picks up this political turbulence, with Kenya being classified as autocratic from 2007 until the freer and fairer elections of 2013. While Cheibub et al. (2010), Geddes et al. (2014), and Boix et al. (2013) code Sri Lanka as democratic between 2005 and 2009, RoW captures the limitations to democracy that existed before and during the 2008/2009 civil war and classify it as autocratic. Similarly, Cheibub et al. (2010), Geddes et al. (2014), Boix et al. (2013), and Polity code Burundi as democratic following the presidential election of 2005, whereas RoW classifies Burundi as autocratic reflecting, among other things, that there was no *de-facto* multiparty competition and president Nkurunziza ran unopposed. RoW also categorizes Nigeria as an electoral autocracy prior to 2011, a reflection of the widespread electoral manipulation that marred all Nigerian elections until 2011 (Lewis, 2011), whereas Cheibub et al. (2010) and Geddes et al. (2014) classify Nigeria as a democracy from 2000. While Polity, Freedom House and Wahman et al. (2013) also place Nigeria on the autocratic spectrum prior to 2011, the democratic improvements in 2011 are not noted in their coding, which is static up until 2015. Another example is Albania, which Boix et al. (2013), Cheibub et al. (2010), and Geddes et al. (2014) code as democratic from 1991 or 1992 and onwards even though the main opposition leader Fatos Nano was jailed from 1993 to 1999 on politically motivated charges (Abrahams, 1996). In contrast, RoW codes Albania as democratic only from 2002 onwards.

Many fewer country-years are classified as democracies in RoW when most or all other measures code them as autocratic. This applies for example to Namibia (1991–2010) where free and fair multiparty elections in combination with freedom of expression and association qualify Namibia as a democracy based on the assessment of the V-Dem experts, whereas most other data sets (Boix et al., 2013; Cheibub et al., 2010; Geddes et al., 2014; and Polity) disagree. In line with RoW, Freedom House classifies Namibia as “free” from the 1990’s onwards, and Larry Diamond (1999) describes it as a “liberal democracy”. Similar disagreement can be observed for Zambia (1994–2007), Burkina Faso (1993–2010) and Mexico (1995–1999).

6. Conclusions

Many research questions require that scholars use a discrete regime variable, either on the right- or left-hand side. Extant approaches to this task are laudable, but are often either limited in their temporal or geographical coverage, not fully transparent in their coding procedures, or have questionably low thresholds for democracy. None of them provide measures of measurement error or other sources of uncertainty to help identify ambiguous cases situated close to thresholds. In this article, we propose a new regime typology—RoW—covering almost all countries from 1900 to 2016. We build on theory conceptualizing democracy as embodying the core value of making rulers responsive to citizens, achieved through electoral competition for the electorate’s approval under circumstances when suffrage is extensive; political and civil society organizations operate freely; elections are clean and not marred by fraud or systematic irregularities; and where elections affect the chief executive of the country. In between elections, democracy requires freedom of expression and an independent media capable of presenting alternative views on matters of political importance. We hence classify countries only as democratic if a minimum level of Dahl’s (1971) famous institutional requisites are fulfilled in terms of freedom of expression and alternative sources of information, freedom of association, universal suffrage, free and fair elections, and the degree to which power is *de-facto* vested in elected officials. Furthermore, we distinguish between democratic (liberal and electoral democracy) and autocratic subtypes (closed and electoral autocracy). Earlier versions of our typology have already been used in scholarly work on democratic backsliding (Mechkova et al., 2017b), the Sustainable Development Goals (Tosun & Leininger, 2017), and political culture (Welzel, 2017), which further underscores the usefulness of the new RoW measures.

Our threshold for democracy is more demanding than in all extant data sets because we base our typology not only the existence and quality of elections but also on Dahl’s notion of Polyarchy. Some extant data sets are limited to *de-jure* rules and other indicators that are directly observable. Other data sets only focus on the implementation of elections in a narrow sense, such as

[17] Except for Freedom House (2017), which classifies Kenya as “Partly free” since 2002.

their *de-jure* competitiveness. RoW is based on V-Dem's high standards in the aggregation of expert-coded data and recruitment of expert coders, which make the data more reliable and allows us to assess the *de-facto* implementation of institutions as opposed to simply their *de-jure* existence.

Finally, RoW is the only measure of discrete regime types that explicitly addresses a fundamental challenge for all typologies: classifying political regimes involves some amount of measurement error and other sources of uncertainty. Therefore, we have designed RoW to incorporate V-Dem's Bayesian intervals indicating where 68% of the probability mass for each country-year score is located. We use these intervals to identify the cases which are close to the thresholds between categories and which are as a result ambiguously classified. Thus, the main RoW measure puts country-years in categories reflecting either certain regime types (closed dictatorships, electoral autocracies, electoral democracies, and liberal democracies), or ambiguous cases in lower and upper bounds of these regime types. This innovation opens up research avenues for incorporating such uncertainty in empirical analyses, thus avoiding biased and potentially misleading results.

Acknowledgements

We are grateful to Valeriya Mechkova for the idea of building a RoW typology capturing ambiguously classified cases. For helpful comments, we also thank Philip Keefer, Beth Simmons, Ariel I. Ahram, Josh Krusell, Kyle Marquardt, Rick Morgan, Dag Tanneberg, the editors and anonymous reviewers; and participants of the V-Dem Research Conference (May 2017), the APSA General Conference (August 2017), the Varieties of Autocracy workshop at the University of Gothenburg (June 2017; funded by Riksbankens Jubileumsfond) and the Empirical Study of Autocracy workshop at the University of Konstanz (September 2017) where earlier versions of this article was discussed. This research project was supported by Riksbankens Jubileumsfond, Grant M13-0559:1, PI: Staffan I. Lindberg, V-Dem Institute, University of Gothenburg, Sweden; by Knut and Alice Wallenberg Foundation to Wallenberg Academy Fellow Staffan I. Lindberg, Grant 2013.0166, V-Dem Institute, University of Gothenburg, Sweden; as well as by internal grants from the Vice-Chancellor's office, the Dean of the College of Social Sciences, and the Department of Political Science at University of Gothenburg.

Conflict of Interests

The authors declare no conflict of interests.

References

Abdulai, A.-G., & Crawford, G. (2010). Consolidating democracy in Ghana: Progress and prospects? *Democratization*, *17*(1), 26–67.

Abrahams, F. (1996). *Human rights in post-communist Albania*. New York, NY: Human Rights Watch.

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

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

Bernhard, M., Hicken, A., Reenock, C. M., & Lindberg, S. I. (2015). *Institutional subsystems and the survival of democracy: Do political and civil society matter?* (Working Paper (4)). Gothenburg: Varieties of Democracy Institute.

Bogaards, M. (2012). Where to draw the line? From degree to dichotomy in measures of democracy. *Democratization*, *19*(4), 690–712.

Boix, C., Miller, M., & Rosato, S. (2013). A complete data set of political regimes, 1800–2007. *Comparative Political Studies*, *46*(12), 1523–1554.

Bollen, K. A., & Jackman, R. W. (1989). Democracy, stability, and dichotomies. *American Sociological Review*, *54*(4), 612–621.

Botero, J. C., & Ponce, A. (2011). *Measuring the rule of law* (Working Paper Series). Washington, DC: The World Justice Project.

Brownlee, J. (2009). Harbinger of democracy: Competitive elections before the end of authoritarianism. In S. I. Lindberg (Ed.), *Democratization by elections* (Vol. 2009, pp. 128–47). Baltimore, MD: Johns Hopkins University Press.

Cheibub, J. A., Gandhi, J., & Vreeland, J. R. (2010). Democracy and dictatorship revisited. *Public Choice*, *143*(1/2), 67–101.

Cheibub, J. A., Przeworski, A., Limongi Neto, F. P., & Alvarez, M. M. (1996). What makes democracies endure? *Journal of Democracy*, *7*(1), 39–55.

Collier, D., & Adcock, R. (1999). Democracy and dichotomies: A pragmatic approach to choices about concepts. *Annual Review of Political Science*, *2*(1), 537–565.

Collier, D., & Levitsky, S. (1997). Democracy with adjectives: Conceptual innovation in comparative research. *World Politics*, *49*(3), 430–451.

Coppedge, M., Lindberg, S., Skaaning, S.-E., & Teorell, J. (2016). Measuring high level democratic principles using the V-Dem data. *International Political Science Review*, *37*(5), 580–593.

Coppedge, M., Gerring, J., Altman, D., Bernhard, M., Fish, S., Hicken, A., . . . Teorell, J. (2011). Conceptualizing and measuring democracy: A new approach. *Perspectives on Politics*, *9*(2), 247–267.

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S.-E., Teorell, J., Altman, D., . . . Staton, J. (2017a). *V-Dem codebook v7.1*. Gothenburg: Varieties of Democracy (V-Dem) Project.

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S.-E., Teorell, J., Altman, D., . . . Wilson, S. (2017b). *V-Dem*

[country-year/country-date] dataset v7.1. Gothenburg: Varieties of Democracy (V-Dem) Project.

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S.-E., Teorell, J., Krusell, J., . . . Wilson, S. (2017c). *V-Dem methodology v7*. Gothenburg: Varieties of Democracy (V-Dem) Project.

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S.-E., & Teorell, J. (2017d). *V-Dem comparisons and contrasts with other measurement projects* (Working Paper (45)). Gothenburg: Varieties of Democracy (V-Dem) Project.

Coulon, C. (1988). Senegal: The development and fragility of semidemocracy. *Democracy in Developing Countries*, *2*, 141–178.

Dahl, R. A. (1956). *A preface to democratic theory*. Chicago, IL: University of Chicago Press.

Dahl, R. A. (1971). *Polyarchy: Participation and opposition*. New Haven, CT: Yale University Press.

Dahl, R. A. (1998). *On democracy*. New Haven, CT: Yale University Press.

Diamond, L. (1999). *Developing democracy: Toward consolidation*. Baltimore, MD: Johns Hopkins University Press.

Diamond, L. (2002). Thinking about hybrid regimes. *Journal of Democracy*, *13*(2), 21–35.

Donno, D. (2013). Elections and democratization in authoritarian regimes. *American Journal of Political Science*, *57*(3), 703–716.

Erdmann, G. (2011). Decline of democracy: Loss of quality, hybridisation and breakdown of democracy. In G. Erdmann & M. Kneuer (Eds.), *Regression of democracy?* (pp. 21–58). Wiesbaden: Springer.

Freedom House. (2017). *Freedom in the world 2016*. Washington, DC: Freedom House.

Gandhi, J., & Lust-Okar, E. (2009). Elections under authoritarianism. *Annual Review of Political Science*, *12*, 403–422.

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

Geddes, B., Wright, J., & Frantz, E. (2013). Autocratic regimes code book, version 1.2. *Perspective on Politics*, *12*(2). Retrieved from http://sites.psu.edu/dictators/wp-content/uploads/sites/12570/2014/06/GWF-Codebook.pdf

Gleditsch, K. S., & Ward, M. D. (2006). Diffusion and the international context of democratization. *International Organization*, *60*(4), 911–933.

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

Hamilton, A., Madison, J., & Jay, J. (2009). *The federalist papers*. New York, NY: Palgrave Macmillan. (Original work published 1787)

Harris, D. J. (2014). *Sierra Leone: A political history*. Oxford: Oxford University Press.

Human Rights Watch. (2008). *Ballots to bullets: Organized political violence and Kenya's crisis of governance*. New York, NY: Human Rights Watch.

Huntington, S. P. (1992). *The third wave: Democratization in the late twentieth century*. Norman, OK: University of Oklahoma Press.

Kagwanja, P., & Southall, R. (2009). Kenya's uncertain democracy: The electoral crisis of 2008. *Journal of Contemporary African Studies*, *27*(3), 257–461.

Kailitz, S. (2013). Classifying political regimes revisited: Legitimation and durability. *Democratization*, *20*(1), 39–60.

Karl, T. (1986). Imposing consent? Electoralism versus democratization in El Salvador. In P. Drake & E. Silva (Eds.), *Elections and democratization in Latin America* (pp. 9-36). San Diego, CA: University of California San Diego.

Kelley, J. (2009). D-minus elections: The politics and norms of international election observation. *International Organization*, *63*(4), 765–787.

Knutsen, C. H., & Wig, T. (2015). Government turnover and the effects of regime type: How requiring alternation in power biases against the estimated economic benefits of democracy. *Comparative Political Studies*, *48*(7), 882–914.

Levitsky, S., & Way, L. A. (2010). *Competitive authoritarianism: Hybrid regimes after the cold war*. Cambridge: Cambridge University Press.

Lewis, P. M. (2011). Nigeria votes: More openness, more conflict. *Journal of Democracy*, *22*(4), 60–74.

Lindberg, S. I. (2006). *Democracy and elections in Africa*. Baltimore, MD: Johns Hopkins University Press.

Lindberg, S. I. (2016). Ordinal versions of V-Dem's indices: When interval measures are not useful for classification, description, and sequencing analysis purposes. *Geopolitics, History, and International Relations*, *8*(2), 76–111.

Lindberg, S. I., Coppedge, M., Gerring, J., & Teorell, J. (2014). V-Dem: A new way to measure democracy. *Journal of Democracy*, *25*(3), 159–169.

Lührmann, A., McMann, K. M., & Van Ham, C. (2017). *The effectiveness of democracy aid to different regime types and democracy sectors* (Working Paper (40)). Gothenburg: Varieties of Democracy Institute.

Marquardt, K. L., & Pemstein, D. (2017). *IRT models for expert-coded panel data* (Working Paper (41)). Gothenburg: Varieties of Democracy Institute.

Marshall, M. G., Gurr, R. T., & Jaggers, K. (2014). *Political regime characteristics and transitions, 1800–2015. Data set users' manual*. Vienna, VA: Center for Systemic Peace.

Mechkova, V., Lührmann, A., & Lindberg, S. I. (2017a). *From de-jure to de-facto: Mapping dimensions and sequences of accountability* (World Development Report 2017). Washington, DC: World Bank.

Mechkova, V., Lührmann, A., & Lindberg, S. I. (2017b). How much democratic backsliding? *Journal of Democracy*, *28*(4), 162–169.

Merkel, W. (2004). Embedded and defective democracies. *Democratization*, *11*(5), 33–58.

Munck, G. L. (2009). *Measuring democracy: A bridge be-*

tween scholarship and politics. Baltimore, MD: Johns Hopkins University Press.

Organization of American States. (2008). *Hemispheric struggle for women's suffrage, Inter-American Commission on Women Organization of American States*. Retrieved from www.oas.org

Pastor, R. A. (1999). The third dimension of accountability: The international community in national elections. In A. Schedler (Ed.), *The self-restraining state: Power and accountability in new democracies* (pp. 123–142). Boulder, CO: Lynne Rienner Publishers.

Pemstein, D., Marquardt, K. L., Tzelgov, E., Wang, Y.-T., Krusel, J., & Miri, F. (2017). *The V-Dem measurement model: Latent variable analysis for cross-national and cross-temporal expert-coded data* (Working Paper (21)). Gothenburg: Varieties of Democracy Institute.

Przeworski, A., Alvarez, M. E., Cheibub, J. A., & Limongi, F. (2000). *Democracy and development: Political institutions and well-being in the world, 1950–1990* (Vol. 3). Cambridge: Cambridge University Press.

Rössler, P. G., & Howard, M. M. (2009). Post-cold war political regimes: When do elections matter? In S. I. Lindberg (Ed.), *Democratization by elections* (Vol. 2009, pp. 101–128). Baltimore, MD: Johns Hopkins University Press.

Rutten, M., & Owuor, S. (2009). Weapons of mass destruction: Land, ethnicity and the 2007 elections in Kenya. *Journal of Contemporary African Studies*, *27*(3), 305–324.

Sartori, G. (1970). Concept misformation in comparative politics. *American Political Science Review*, *64*(4), 1033–1053.

Schedler, A. (2002). The menu of manipulation. *Journal of Democracy*, *13*(2), 36–50.

Schedler, A. (2006). *Electoral authoritarianism: The dynamics of unfree competition*. Boulder, CO: Lynne Rienner Publishers.

Schedler, A. (2013). *The politics of uncertainty: Sustaining and subverting electoral authoritarianism*. Oxford: Oxford University Press.

Svolik, M. (2008). Authoritarian reversals and democratic consolidation. *American Political Science Review*, *102*(2), 153–168.

Tosun, J., & Leininger, J. (2017). Governing the Interlinkages between the Sustainable Development Goals: Approaches to attain policy integration. *Global Challenges*, *1*(9), 1–12.

Turner, M., & Tshering, J. (2014). Is democracy being consolidated in Bhutan? *Asian Politics & Policy*, *6*(3), 413–431.

Vanhanen, T. (2005). *Measures of democracy 1810–2004. Version, 2., Dataset*. Tampere: Finnish Social Science Data Archive.

Vengroff, R. (1993). Transition to democracy in Senegal: The role of decentralization. In I. Kim & J. Shapiro (Eds.), *Establishing democratic rule* (pp. 159–188). Washington, DC: Paragon Press.

Wahman, M. (2014). Democratization and electoral turnovers in Sub-Saharan Africa and beyond. *Democratization*, *21*(2), 220–243.

Wahman, M., Teorell, J., & Hadenius, A. (2013). Authoritarian regime types revisited: Updated data in comparative perspective. *Contemporary Politics*, *19*(1), 19–34.

Wang, Y.-T., Lindenfors, P., Sundstöm, A., Jansson, F., Paxton, P., & Lindberg, S. I. (2017). Women's rights in democratic transitions: A global sequence analysis, 1900–2012. *European Journal of Political Research*, *56*(4), 735–756.

Welzel, C. (2013). *Freedom rising: Human empowerment and the quest for emancipation*. Cambridge: Cambridge University Press.

Welzel, C. (2017). *A tale of culture-bound regime evolution: The centennial democratic trend and its recent reversal* (Users Working Paper (11)). Gothenburg: Varieties of Democracy (V-Dem) Project.

About the Authors

Anna Lührmann is a Senior Research Fellow at the Varieties of Democracy (V-Dem) Institute (University of Gothenburg). She received her PhD in 2015 from Humboldt University (Berlin). Her articles on regime change, autocracies, democracy aid, and elections have appeared in the *Journal of Democracy*, *Electoral Studies*, and *International Political Science Review*.

Marcus Tannenberg is a PhD candidate at the Varieties of Democracy (V-Dem) Institute at the department of political science, University of Gothenburg. Marcus's work focuses on politics and public opinion in autocratic regimes, and survey methodology.

Staffan I. Lindberg is the director of the Varieties of Democracy (V-Dem) Institute and Professor of Political Science at the University of Gothenburg. He is a Wallenberg Academy Scholar, a member of the Young Academy of Sweden and recipient of an ERC consolidator grant as well as numerous other grants and awards for his research. He is the author of *Democracy and Elections in Africa* (John Hopkins University Press, 2006) and editor of *Democratization by Elections* (John Hopkins University Press, 2009).

Appendix

1. Threshold between Closed and Electoral Autocracies in Detail

The V-Dem data set includes specific indicators for legislative and executive elections (v2elmulpar_osp_leg/v2elmulpar_osp_ex). To identify which of the two should be used for assessing if the Head of the Executive is subject to *de-jure* multiparty elections, we need to take the relative power of the Head of State (HoS) and the Head of Government (HoG) and the appointment procedures into account. The V-Dem variable v2ex_hosw identifies if the HoS (v2ex_hosw = 1) or HoG (v2ex_hosw < 1) is the chief executive. If the HoG is the chief executive, the variable v2expathhg indicates whether the HoG is directly (8) or indirectly (7) elected or appointed by the HoS (6). In the first case, we take the multiparty variable for executive elections (v2elmulpar_osp_ex), in the second case for legislative elections (v2elmulpar_osp_leg) and in the third case the score for HoS as follows. If the HoS is the chief executive, the variable v2expathhs indicates whether the HoS is directly (7) or indirectly (6) elected. In the first case, we take the multiparty variable for executive elections (v2elmulpar_osp_ex), in the second case for legislative elections (v2elmulpar_osp_leg). (see Coppedge et al., 2017a).

2. RoW by Country for 2016.

Liberal Democracy		Electoral Democracy		Electoral Autocracy		Closed Autocracy	
Albania		Bhutan	+	Comoros	+	Turkmenistan	+
Australia		Cape Verde	+	Fiji	+	Kuwait	+
Austria		Chile	+	Guinea	+	Vietnam	+
Belgium		Ghana	+	Haiti	+		
Canada		Guyana	+	Honduras	+	China	
Costa Rica		Israel	+	Iraq	+	Cuba	
Cyprus		Lithuania	+	Madagascar	+	Eritrea	
Denmark		Mauritius	+	Mozambique	+	Jordan	
Estonia		Moldova	+	Niger	+	Laos	
Finland		Panama	+	Papua New Guinea	+	Libya	
France		Poland	+	Serbia	+	Morocco	
Germany		Senegal	+	Somaliland	+	North Korea	
Iceland		Seychelles	+			Oman	
Ireland		Slovakia	+	Afghanistan		Palestine/Gaza	
Japan		South Africa	+	Algeria		Qatar	
Netherlands		São Tomé & Príncipe	+	Angola		Saudi Arabia	
New Zealand		Trinidad & Tobago	+	Armenia		Somalia	
Norway		Tunisia	+	Azerbaijan		South Sudan	
Portugal		Vanuatu	+	Bangladesh		Swaziland	
Spain				Belarus		Syria	
Sweden		Argentina		Bosnia & Herzegovina		Thailand	
Switzerland		Bolivia		Burma/Myanmar		Yemen	
United Kingdom		Brazil		Burundi			
United States		Bulgaria		Cambodia			
Uruguay		Burkina Faso		Cameroon			
		Colombia		Chad			
Barbados	–	Croatia		Dem. Rep. of Congo			
Benin	–	Dominican Rep		Djibouti			
Botswana	–	Ecuador		Egypt			
Czech Republic	–	El Salvador		Equatorial Guinea			
Italy	–	Georgia		Ethiopia			
Latvia	–	Greece		Gabon			
Namibia	–	Guatemala		Gambia			
Slovenia	–	Hungary		Iran			
South Korea	–	India		Kazakhstan			
Taiwan	–	Indonesia		Malaysia			
		Ivory Coast		Maldives			
		Jamaica		Mauritania			

Liberal Democracy	Electoral Democracy		Electoral Autocracy		Closed Autocracy
	Lesotho		Montenegro		
	Liberia		Nicaragua		
	Mali		Pakistan		
	Mexico		Palestine/West Bank		
	Mongolia		Rep. of the Congo		
	Nepal		Russia		
	Nigeria		Rwanda		
	Paraguay		Singapore		
	Peru		Sudan		
	Philippines		Tajikistan		
	Romania		Tanzania		
	Solomon Islands		Turkey		
	Sri Lanka		Uganda		
	Suriname		Ukraine		
	Timor-Leste		Venezuela		
	Togo		Zambia		
			Zanzibar		
	Central African Rep.	–	Zimbabwe		
	Guinea-Bissau	–			
	Kenya	–	Uzbekistan	–	
	Kosovo	–			
	Kyrgyzstan	–			
	Lebanon	–			
	Macedonia	–			
	Malawi	–			
	Sierra Leone	–			

Note: The "+" and "–" denotes an ambiguous case. "+" indicates that some evidence suggests that the country might be better placed in the next higher category. "–" indicates that the country might be better placed in the next lower category. For more detail see section 3.

References

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S.-E., Teorell, J., Altman, D., . . . Staton, J. (2017a). *V-Dem codebook v7.1*. Gothenburg: Varieties of Democracy (V-Dem) Project.

Politics and Governance (ISSN: 2183–2463)
2018, Volume 6, Issue 1, Pages 78–91
DOI: 10.17645/pag.v6i1.1200

Article

Making Trade-Offs Visible: Theoretical and Methodological Considerations about the Relationship between Dimensions and Institutions of Democracy and Empirical Findings

Hans-Joachim Lauth * and Oliver Schlenkrich

Chair of Comparative Politics and German Government, University of Wuerzburg, 97074 Wuerzburg, Germany; E-Mails: hans-joachim.lauth@uni-wuerzburg.de (H.-J. L.), oliver.schlenkrich@uni-wuerzburg.de (O.S.)

* Corresponding author

Submitted: 30 September 2017 | Accepted: 7 December 2017 | Published: 19 March 2018

Abstract
Whereas the measurement of the quality of democracy focused on the rough differentiation of democracies and autocracies in the beginning (e.g. Vanhanen, Polity, Freedom House), the focal point of newer instruments is the assessment of the quality of established democracies. In this context, tensions resp. trade-offs between dimensions of democracy are discussed as well (e.g. Democracy Barometer, Varieties of Democracy). However, these approaches lack a systematic discussion of trade-offs and they are not able to show trade-offs empirically. We address this research desideratum in a three-step process: Firstly, we propose a new conceptual approach, which distinguishes between two different modes of relationships between dimensions: mutual reinforcing effects and a give-and-take relationship (trade-offs) between dimensions. By introducing our measurement tool, Democracy Matrix, we finally locate mutually reinforcing effects as well as trade-offs. Secondly, we provide a new methodological approach to measure trade-offs. While one measuring strategy captures the mutual reinforcing effects, the other strategy employs indicators, which serve to gauge trade-offs. Thirdly, we demonstrate empirical findings of our measurement drawing on the Varieties of Democracy dataset. Incorporating trade-offs into the measurement enables us to identify various profiles of democracy (libertarian, egalitarian and control-focused democracy) via the quality of its dimensions.

Keywords
control-focused democracy; democracy; Democracy Matrix; egalitarian democracy; libertarian democracy; measurement of democracy; profile of democracy; quality of democracy; trade-off; Varieties of Democracy

Issue
This article is part of the issue “Why Choice Matters: Revisiting and Comparing Measures of Democracy”, edited by Heiko Giebler (WZB Berlin Social Science Center, Germany), Saskia P. Ruth (German Institute of Global and Area Studies, Germany), and Dag Tanneberg (University of Potsdam, Germany).

© 2018 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

One unresolved question of the measurement of democracy is the existence of trade-offs between dimensions, that is to say, whether their relationship is characterized by tensions and conflicting goals, which result in trade-offs between them. Even though newer indices of democracy (Democracy Barometer, Varieties of Democracy/V-Dem) mention the idea of trade-offs, they are, however, not able to demonstrate trade-offs empirically. Giebler and Merkel (2016, p. 602) state, based on the Democracy Barometer data, that in contrast to the “traditional libertarian fear of a trade-off between freedom and equality…, we find that the two core principles of democracy (freedom and equality) possess a mutually reinforcing association”. Similarly, V-Dem mentions the idea of trade-offs in their conceptual paper (Coppedge, Gerring, Altman, & Bernhard, 2011), but they seem to

not be able to detect these trade-offs empirically, e.g. cases can be identified with the highest rating in the freedom dimension and in the equality dimension simultaneously (Coppedge, Lindberg, Skaaning, & Teorell, 2015, p. 9). Why is this the case?

We argue that there are at least two reasons: on the one hand, these measures lack a deep discussion of the conceptual foundations of trade-offs missing not only the detection of concrete realization of trade-offs but also their interconnectedness with different abstract conceptions of democracy. This means that current measures of democracy content themselves with only a short remark about trade-offs on the highest aggregated level (dimensions or principles) but do not consider these conceptual consequences for lower or mid-level components of democracies (institutions). In fact, no definite characterization or, to be more precise, definition of trade-offs has ever been made, even in the more theoretical discussions about the quality of democracy (see for a general discussion Diamond & Morlino, 2005). On the other hand, they lack an adequate empirical measurement strategy by not adapting their measurement and aggregation stage to capture the different "nature" of trade-off relationships. Current measures of democracy use unidimensional indicators to measure an actual two-dimensional relationship resulting in a blind spot concerning trade-offs. This article tackles these two conceptual and methodological problems: how can we understand trade-offs conceptually and how can we successfully incorporate them in a measurement of the quality of democracy?[1]

Thus, to close this research gap, this article proceeds in three steps: firstly, we propose a new conceptual approach, which is able to define and distinguish between two different modes of relationships between dimensions (section 2): mutual reinforcing effects between dimensions and a give-and-take relationship (trade-offs). By introducing our measurement tool, Democracy Matrix, which combines three dimensions (political freedom, political equality and political and constitutional control) with five central democratic functions, we locate trade-offs. On the basis of these three dimensions, we propose three ideal typical profiles of democracy: libertarian, egalitarian and a control-focused profile of democracy.

Secondly, we provide a new methodological approach to measure trade-offs (section 3): two independent measurements are combined to assess the quality of democracy. While one measuring strategy applies indicators commonly used in other indices (such as Freedom House or Varieties of Democracy) relying on a unidimensional interpretation, the other strategy employs indicators which serve to assess trade-offs by incorporating and expressing the two-dimensional relationship which is characteristic for trade-offs. We call the former type of indicators "quality measuring indicators" and the latter "trade-off indicators".

Thirdly, we demonstrate empirical findings of our measurement drawing on the Varieties of Democracy dataset (section 4). Incorporating trade-offs into the measurement enables the identification of various profiles of democracy via the quality of its dimensions.

2. Conceptual Considerations: Quality and Profiles of Democracies

2.1. The Democracy Matrix: A New Measurement Tool Which Combines Mutual Reinforcing Effects and Trade-Offs between Dimensions

The Democracy Matrix is based on the 15-Field-Matrix (Lauth, 2004, 2015). The 15-Field-Matrix combines three dimensions with five central democratic functions: Whereas the dimension of freedom captures the extent of the free self-determination of the citizens based on civil and political rights, the equality dimension encompasses legal egalitarianism and the actual realization of those rights (input-egalitarianism). The control dimension takes into account the protection of the two other dimensions through legal control performed by judiciaries and political control performed by intermediary institutions, media and parliament. On the one hand, this democracy conception is primarily rooted in Dahl's (1971) widely acknowledged distinction between "contestation" and "participation" which is resembled in the dimensions of freedom and equality. On the other hand, it adds a third dimension, control, to capture the deficient functioning of horizontal accountability and the rule of law.[2] This extension of the conception is due to the basic conviction that democracy is a type of limited rule. The analysis of third wave democracies, which often have shown significant deficits regarding horizontal accountability and rule of law (O'Donnell, 1994), demonstrates the relevance of this third dimension of control.

In addition, five central functions cut across these three dimensions concretizing the quality of democracy. The "procedures of decision" function analyzes the democratic quality of representative elections and direct democracy. The "regulation of the intermediate sphere" captures the democratic performance of interest aggregation and interest articulation by parties, interest organizations and civil society. "Public communication" evaluates the functioning of the media system and the public realm. The "guarantee of rights" function analyzes the democratic quality of the court system, whereas the last function, "rules settlement/implementation", focuses on the democratic quality of the work carried out by the executive and legislature. This unfolds 15 matrix-fields

[1] We are not convinced that the theoretical assumption of the existence of trade-offs could be wrong, although we will stress this possibility in our discussion as well.

[2] In a sense, this third dimension reflects the binding or limiting mechanism of democratic rule, which Dahl (1956) highlights under the term Madisonian democracy.

which supports the analysis of the quality of democracy in an elaborate manner.

The Democracy Matrix enhances the concept of the 15-Field-Matrix by distinguishing between two basic types of relations between dimensions: mutual reinforcing effects and trade-offs. While the mutual reinforcing effects are already sufficiently captured by the 15-Field-Matrix, the inclusion of the concept and measurement of trade-offs is the additional feature of the Democracy Matrix. Figure 1 presents trade-offs inside the Democracy Matrix, which we have identified.

The idea of mutual reinforcing effects between dimensions can usually be found in all measures of democracy: The Variety of Democracy-project describes this type of relationship for freedom and equality by stating "that to some extent the contribution of one attribute depends on the presence of the other. If, say, oppositional candidates are not allowed to run for election, or the elections are fraudulent, it does not matter much for the level of electoral democracy that all adults have voting rights" (Coppedge et al., 2015, p. 6). The concept of the Democracy Barometer is based on "the assumption of necessary and sufficient conditions for being a member of the category democracy" (Merkel et al., 2016, p. 8). In addition, Diamond and Morlino (2004, pp. 28–29) suppose that the "dimensions are closely linked and tend to move together, either toward democratic improvement and deepening or toward decay". This means that dimensions are not only necessary to understand democracy, but they are also mutually dependent. One dimension cannot exist without the other. The close relationship between freedom and equality has been emphasized by Dworkin (1996, p. 57): "So we have come, by different routes, beginning in different traditions and paradigms, to conceptions of liberty and equality that seem not only compatible, but mutually necessary". Dahl (1971) emphasizes that a democracy (polyarchy) is only present if both attributes—contestation and participation—are fulfilled. In terms of democratic theory, freedom without a minimum level of equality is as difficult to conceive as equality without freedom. Control, which is required for their protection and enforcement, is checked by constitutionally-set standards of freedom and equality, thereby constraining the unlimited exercise of power. Campbell, Carayannis and Scheherazade (2015) refer to the three dimensions of freedom, equality and control, but add with "sustainable development" a fourth dimension to their *Quadruple dimensional structure of democracy*, which is likewise constructed in a reinforcing perspective.

This mutual reinforcing effect between the dimensions expresses the baseline concept of the Democracy Matrix: all dimensions and thus all 15 matrix fields must work to a sufficient degree for a country to be classified as a democracy. Insofar the Democracy Matrix shares the assumptions of the other indices, it differs in the way it conceptualizes and incorporates trade-offs. This conception implies—as it will be shown below—the combination of two procedures of measurement.

2.2. Conception and Identification of Trade-Offs

Despite the complementary relationship structure, potential tensions between dimensions are impossible to ignore according to political philosophy. Hidalgo (2014) speaks of antinomies within the democracy concept meaning a "contradictoriness of two propositions which both at the same time are reasonable, justified, and valid" (Hidalgo, 2014, p. 29, own translation). More generally but focused on freedom and equality as well, Berlin's value pluralism claims that the "world…is one in which we are faced with choices between equally ultimate ends, and claims that are equally absolute, the realization of some must inevitably involve the sacrifice of others" (Berlin, 1969, p. 168). Diamond and Morlino (2004, p. 21) describe the idea of trade-offs within the realm of democracies: "it is impossible to maximize all [dimensions] at once. [Every] democratic country must make an inherently value-laden choice about what kind of democracy it wishes to be".

Applied to our dimensional framework, relationships become increasingly strained, especially when a dimension is rigidly manifested. If we look at the features of the dimensions on a scale, a convincing case can be made for the following thesis: whereas in most parts of the scale the dimensions are mutually dependent and support one another, seeking the maximum value results in a trade-off. A choice for one side of the trade-off must be made. What is a trade-off and how can we understand a trade-off in democracies?

A relevant trade-off in democracies satisfies the following conditions:

- A trade-off is settled in the *political* sphere of *democracies*: just as democracy is solely defined in political and procedural terms (Munck, 2016, pp. 16–18), trade-offs are only relevant for the quality of democracy if they are in the political sphere. Thus, trade-offs in the economic sphere (e.g. between policy goals) are excluded.
- A trade-off occurs because *only one institution fulfills a specific political function in one dimension*. At the same time, however, this institution *produces necessarily opposed or inverse effects in another dimension linked to the same function*. This relationship means that a choice is forced between different institutional designs accepting the advantages but also the disadvantages of this specific realized institutional solution.
- *Contrasting but interrelated democracy conceptions* offer different institutional solutions for the same function: On the one hand, these conceptions carry equal normative weight, and can be reasonably justified. The same level of quality of democracy is accredited to them, which implies that they and their institutional choices are neutral in relation to the comprehensive quality of democracy. On the other hand, every democracy

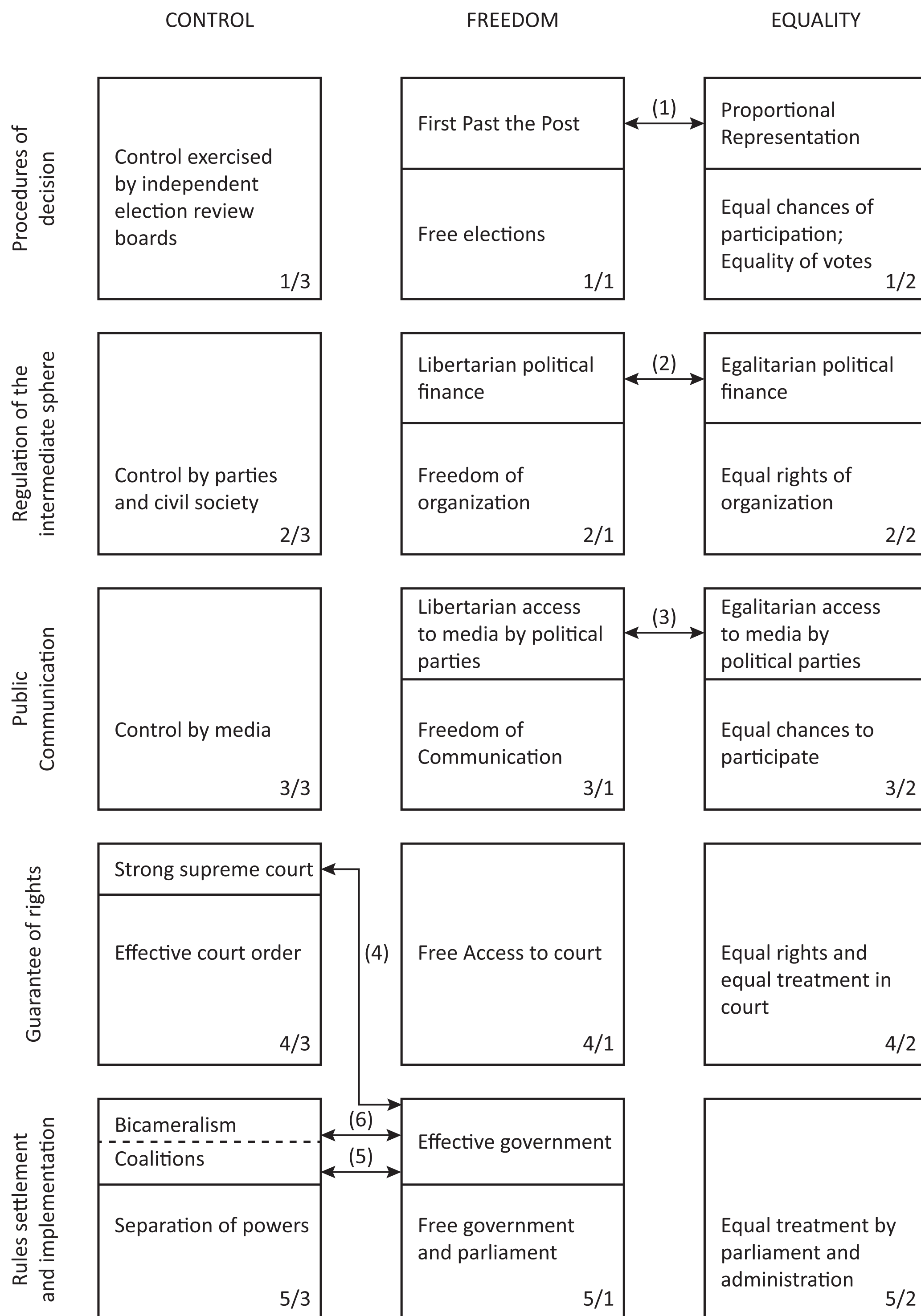


Figure 1. Democracy Matrix. The numbers in parentheses refer to the trade-offs described in the article, the arrows represent the connectedness in which an increase of one dimension determines a decrease of the related dimension. Source: Lauth (2004, 2015).

conception ultimately emphasizes different political values while disregarding others (e.g. freedom over equality). This means they stress a different structuring of the same democratic quality. Therefore, institutions due to their linkage to democracy conceptions highlight different democracy dimensions.

- If an institution overemphasizes one pole of a trade-off by neglecting the other pole completely, an *overstretching* of a trade-off occurs, which damages the baseline concept: between the two poles of a trade-off, there is a normative legitimate space described by the democracy conceptions, in which a democracy can place itself (Hidalgo, 2014). While

> a trade-off accentuates dimensions differently, it still leaves the democracy dimension fully intact: we would not speak of a trade-off anymore when a democracy leaves this democratic space (e.g. overemphasizing the control dimension at the cost of the freedom dimension: a supreme court which acts as a super-legislature). In this case, it damages the baseline concept respective the mutual reinforcing effects between the dimensions.

These explanations distinguish two levels of abstraction: institutions and dimensions. The basic statement is that it is not possible to realize all three dimensions of the Democracy Matrix in comprehensive manner because they are unavoidably linked to trade-offs. This assumption does not mean, however, that each democratic conception as a liberal democracy or a republican democracy must show trade-offs themselves. The reason is trivial, such conceptions have already decided on their preferred dimensions. If you want to transfer the idea of trade-offs to different democratic conceptions, one must maintain that it is not possible to realize two different conceptions at the same time comprehensively. The narrow connection between institutions and dimensions allows the measurement of dimensional trade-offs. The tensions between the dimensions are manifested in institutional choices.

To sum up these considerations, a trade-off in democracies is defined as follows: a trade-off is an irresolvable connectedness between two inverse effects of one institution regarding two dimensions. This trade-off expresses two contrasting but normative, equally weighted democracy conceptions to which the selected institutions belong.

The next step is to identify the relevant trade-offs, keeping in mind that they exist between dimensions but are measured on the corresponding institutional level. We cannot discuss all the different conceptions of democracy in this article. Therefore, we consider as our starting point, the basic democracy principles, which are identified by V-Dem. They derive from six different fundamental conceptions of democracy from democracy theory: liberal, participatory, deliberative, egalitarian, majoritarian and consensus democracy (Coppedge et al., 2011, 2015).[3] These six conceptions are considered as normative equally justified. Thus, even though we agree that "no single conception can reasonably purport to embody all the meanings of democracy" (Coppedge et al., 2011, p. 253), we are convinced that with the help of the trade-off concept, it is possible to create a single, overarching and theoretical justified framework which is able to capture those different notions of democracy. Four concepts are especially helpful to discover relevant trade-offs on the institutional level.

Majoritarian and *consensus* democracy (Lijphart, 2012) are obviously contrary democracy concepts (see Table 1). The former focuses on majority rule, the latter on a vast system of checks and balances. Thus, while consensus democracy stresses multiple structures of veto points constraining the actions of governments (e.g. strong second chamber, federalism, coalitions), the ideal setup of majoritarian democracies favors structures with lesser control abilities. Consensus democracy can also be understood as a constitutional democracy which is characterized by a government decision-making that "mandates a system of checks and balances, that includes, as a key element, courts with the power of judicial review" (Munck, 2016, p. 14). The possible collision between freedom and control is clearly reflected in the regulation of constitutional control, that is the judicial review of the decisions of government and parliament—the constitutional limitation of political majority rule in the "constitutional debate" (Elster & Slagstadt, 1988). We describe these contrasting models as a trade-off: a constitutional court increases the values for the control dimensions in the function "guarantee of rights" and reduces the values of the freedom dimension in the function "rules settlement and implementation" (see Figure 1).

The second, and perhaps even more principal opposing set is the divide between *libertarian* and *egalitarian* conceptions of democracy which relate to the tension between freedom and equality (see Table 2). Whereas egalitarian democracy would highlight political equality, libertarian democracy would focus on political liberty. This

Table 1. Trade-off between majoritarian and consensus democracy. Source: Own presentation.

	Majoritarian Democracy	**Consensus Democracy**
Function	Effective government	
	High	Weak
Institution	Single party governments Unicameral system Unitarism No supreme courts	Oversized coalitions (5) Bicameral system (6) Federalism Supreme courts (4)
Dimension	Freedom	Control

[3] Electoral democracy, another concept of democracy, is considered as a baseline concept by V-Dem and thus, is combined with the other conceptions.

Table 2. Trade-off between libertarian and egalitarian democracy. Source: Own presentation.

	Libertarian Democracy	Egalitarian Democracy
Function	Access to government; influence	
	Free	Equal
Institution	Plurality Voting Unregulated party finance Unregulated media access	PR (1) Equal party finance (2) Equal media access (3)
Dimension	Freedom	Equality

trade-off has triggered profound ideological and philosophical clashes (Dworkin, 1996). We can illustrate the trade-offs between the two dimensions (freedom and equality) by the following examples of institutions with different characteristics.

Electoral systems can be arranged along two "representation principles" (Nohlen, 2014, pp. 243–244). Proportional representation (PR) increases the chances of parties being represented in parliament. In parliamentary systems, this often leads to compromises in the formation of a government (coalitions), which would be less necessary in majority elections (plurality voting system or First Past the Post—FPTP). In the latter case, the electorate has a higher degree of freedom in determining the government than in PR-systems, where the coalition formation is mostly decisive. Beyond government selection, however, equal representation—as measured by the proportionality factor—is reflected most comprehensively in proportional electoral systems (Nohlen, 2014). While disproportional electoral systems stress the freedom dimension, proportional electoral systems emphasize the equality dimension.

A further trade-off can be found in the way of regulating political finance: "The way that political finance should be regulated needs to be the result of a country's political goals….To put it differently, since there is no form of democratic governance that is preferred everywhere, there is no ultimate method of regulating political finance" (Ohman, 2014, p. 16). Two ideal types of political finance can be distinguished. Whereas the egalitarian model of political finance emphasizes equal opportunities between candidates and/or parties through public finance, the libertarian model of political finance has a "lack of restrictions on expenditure and contributions, market principles of access to the media [and] no public funding" (Smilov, 2008, p. 3). This type conceives donations to parties or candidates as a freedom of expression. Therefore, the libertarian political finance model strengthens the freedom dimension within the function "regulation of the intermediate sphere", while the egalitarian political finance model focuses on the equality dimension.

Finally, the trade-off between libertarian and egalitarian media access follows the same considerations as the political financing of political parties. A libertarian media access "provides for market access to the media" (Smilov, 2008, p. 9) giving more economically powerful actors more possibilities, while the egalitarian model provides free media time for candidates and/or parties. This trade-off concerns the "public communication" function and the dimensions of "freedom" and "equality".

2.3. Profiles of Democracies: The Interplay of the Two Basic Relationships of Dimensions

The theoretical debate not only elucidates general trade-offs between all core dimensions of democracy; it also shows that they are interconnected and mutually supportive. By combining the mutual reinforcing effects with the trade-offs, we can obtain different profiles of democracy. Thereby, each basic relationship serves a specific task: the mutual reinforcing effects indicate the appropriate manifestation of a dimension. If a democracy is present, the trade-off-relationship becomes important, determining the final shape of the dimensions in the upper spectrum of a working democracy. In principle, democracy theory implies that an "optimal" or "perfect" democracy cannot be based upon the complete realization of all three dimensions, making the achievement of the highest level of democratic quality in every dimension impossible. Maximizing the quality of democracy on one dimension, necessarily sacrifices democracy quality in another dimension. The decision involving the preference given to which dimension(s) is a matter to be decided by the democratic sovereign, the people. They decide on which dimensions should be emphasized at the cost of others resulting in diverging profiles.

Based on the dimensional framework of the Democracy Matrix, three ideal typical profiles of democracies can be distinguished (see Figure 2): a libertarian profile of democracy, which maximizes the freedom dimension at the cost of the others, an egalitarian profile of democracy, which highlights the equality dimension and finally, a control-focused profile of democracy emphasizing the control dimension. Empirically, we will probably observe hybrid types with specific profiles (e.g. high egalitarian and control values combined with low freedom values).

3. A New Measurement Strategy: Quality Measuring Indicators and Profile Measuring Indicators

Despite the lack of a deep theoretical discussion of trade-offs by current measures, another problem, which is nei-

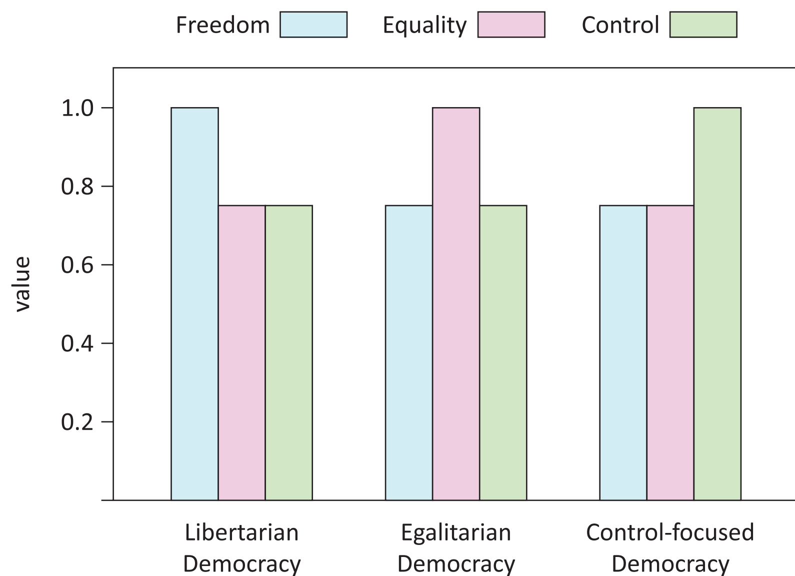


Figure 2. Three profiles of democracies Source: Own presentation.

ther discussed nor resolved, exists on the methodological level. If tensions among the dimensions exist, they need to be accounted for at the indicator-level as well. No approach of measuring democracy discusses trade-offs in its methodological foundation. The consequences of this oversight are reflected in the creation of indicators. As no inherent relationships between indicators vis-à-vis the tensions among dimensions are considered, the indicators are not a valid measurement of trade-offs. On the contrary, they systematically ignore them, because it is not possible to measure a two-dimensional give-and-take relationship with unidimensional indicators. Therefore, no current measurement of democracy can detect trade-offs. Recognized measurements—such as Polity and Freedom House—present findings in which all the indicators show the highest value (10 by Polity, 1 by Freedom House). Even the findings of the newer approaches (Democracy Barometer, V-Dem) are no exception.

The main problem lies partly in the selection of indicators, but also in their use—or more precisely, in the interpretation of the relevant indicators. The solution consists in the use of one indicator for the measurement of two dimensions, but to evaluate it differently with regard to each dimension. This procedure will now be explained.

First and foremost, we have to differentiate between two categories of indicators (see Table 3). The first category includes the usual indicators which are commonly found in the discipline (such as free elections). We call this kind of indicators "quality measuring indicators". The second category captures the realization of the trade-offs via a two-dimensional interpretation. They express the degree of tension between two normative equal conceptions of democracy by assessing the structural arrangement of a specific function (e.g. PR or plurality voting for the electoral system). These are neutral in the aggregate, but sensitive to the differentiation of the dimensions. The inherent tensions of these indicators are due to their contradictory preferences with regard to the dimensions: If one indicator allows the highest grading in a particular dimension, it necessarily prevents the highest grading in the corresponding dimension. We call this type of indicators "trade-off indicators".

The next task is to measure this difference with the trade-off indicators. In order to do so, collected data must be transformed in order to measure democracy, as is necessary with many other indicators (Lauth, 2016). The starting point of our consideration about election systems is that both election systems satisfy the criteria of a working democracy, even if they emphasize different priorities (e.g. freedom vs. equality). For that reason, the transformation of the data must take the threshold of a working democracy into account. In our concept, this

Table 3. Two types of indicators. Source: Own presentation.

Quality measuring indicators	Trade-off indicators
Mutual reinforcing relationship of dimensions	Conflicting relationship of dimensions
Universal: scale captures autocracies and democracies	Only democracies
Gradual differences in the quality of democracy	Equivalent differences in quality of democracy
Unidimensional interpretation	Two-dimensional interpretation
Level of the quality of democracy	*Extent of trade-off*

threshold is at 0.75, the data of the (dis)proportionality should be transformed into a scale from 0.75 to 1. This difference of 0.25 is sufficient to illustrate the hypothetical trade-offs in the empirical research.

For the final assessment of the democratic quality of elections, these values are multiplied with ratings from the commonly used indicators for measuring elections set at a 0 to 1 scale.[4] The multiplication strategy is necessary to respect the assumption that trade-offs are mainly pronounced in the higher areas of the dimensions. The multiplication of both variables shows the highest difference in the profiles (0.75–1.0) when the quality of democracy has the highest degree (1.0). The lower the degrees of the quality, the less pronounced the profile is. The effect, however, is also remarkable in the middle ranges. This is due to the fundamental decision about the institutional foundation of the profiles.

The proposed research strategy therefore combines two measurements, one of regime rating and one of trade-offs. Together they make it possible to assess the quality of democracy and the shaping of its dimensions. The main instrument is the dual interpretation of one indicator in relation to two dimensions. An additional methodological instrument is the setting of thresholds. By using both instruments together and linking them with conventional assessments, it is possible to construct a method of measuring democracy that accurately takes trade-offs into account. In applying this method, it is not possible for all empirical cases to be rated with the highest values in every dimension. Lower values, however, do not represent democratic deficiencies, but different profiles within the area of working democracies.

4. Empirical Findings: The Empirical Manifestations of Profiles of Democracy

In this section, we illustrate our new approach by presenting the results of a cluster analysis and, in addition, we show long-time developments of profiles for single countries. The Democracy Matrix uses the V-Dem dataset (Version 6.2, Coppedge et al., 2016) to measure every single matrix field and the trade-offs (see Figure A1 in the Annex for further details). We performed a hierarchical cluster (Ward's method) analysis with 94 cases (working democracies)[5] resulting in six clusters (see Figure 3). We find strong evidence for a control-focused profile (cluster 2 and with a somewhat lower quality of democracy cluster 3; e.g. The United States, Australia and Switzerland) and a libertarian profile (cluster 4: United Kingdom, New Zealand and Ireland). We find less evidence for a genuine egalitarian democracy type but there are two clusters with high equality values (cluster 1 and 5): while cluster 1 has high values for the equality and control dimension (e.g. Sweden, Norway and Germany), cluster 5 combines slightly higher values for the equality dimension than the control dimension (Austria, Belgium and Netherlands). Therefore, both clusters seem to be a mix between an egalitarian and control-focused democracy profile. Lastly, cluster 6 represents a profile which balances all three dimensions.

Figure 3. Results of the cluster analysis. Cluster size: 19 (1), 17 (2), 15 (3), 10 (4), 21 (5), 12 (6). Source: Own calculation based on the V-Dem-Dataset (Version 6.2; Coppedge et al., 2016).

[4] For example, the quality measuring indicators show for both countries (A and B) the highest democratic quality for all dimensions (1). Country A has a FPTP voting system emphasizing the freedom dimension over the equality dimension, but country B uses a PR electoral system highlighting the equality dimension in contrast to the freedom dimension. This trade-off indicator shows for country A the values 1 (freedom dimension) and 0.75 (equality dimension), this is vice versa for country B. Multiplying the quality measuring and trade-off indicators gives the values 1 (freedom dimension) and 0.75 (equality dimension) for country A and the opposite result for country B, showing the different structuring of the same quality of democracy.

[5] To increase the sample size and the varieties of profiles, we pooled the data for 1970, 1990 and 2010. Thus, some cases are included in the analysis more than once.

We now show the empirical development of profiles for single countries and contrast the results from the quality-measuring indicators alone with our combined measurement approach using trade-off-indicators as well. Excluding clusters 3 and 6 due to the somewhat lower quality of democracy, we selected those cases which are prototypical for the other four clusters (Germany, United States, United Kingdom and Belgium). The left side of Figure 4 shows the results of the quality measuring indicators alone, whereas the right side shows the results of the trade-off indicators multiplied with the quality measuring indicators. The difference between the left and the right side is noticeable: Overall, we gain only small differences between the dimensions using the quality measuring indicators alone (left side). The United States and, since 1990, the United Kingdom seem to be an exception showing considerably lower values for the equality dimension. When we add the trade-off indicators, we are able to gain more profound profiles of democracy (right side). Whereas high values in every dimension of the quality measuring indicators can be achieved; trade-offs are only visible using our new measurement strategy.

For example, Germany combines high control and equality values with lower values for freedom (mixed type of an egalitarian and control-focused democracy). This profile has not changed since 1950. The United States combines high control and freedom values with lower values for equality (mixed type of a libertarian and control-focused democracy). This profile is constant throughout history, but there was a short time increase of equality between 2000 and 2010. Compared to the left side of Figure 4, the control dimension of the United States reaches higher values than the freedom dimension, emphasizing that the differences on the left side are differences in the quality of democracy rather than "real" trade-offs between dimensions. The democracy profile of the United Kingdom consists of high values for freedom combined with lower values for control and equality (libertarian democracy). Since 1990, we can detect a significant increase of the control dimension for this case, which is congruent to the findings of qualitative studies (Strohmeier, 2011). Lastly, Belgium's profile consists of higher values for equality with intermediate values for the control dimension and low values for the freedom dimension (mixed type between egalitarian and control-focused democracy). Belgium changed to this profile since the 1980 by increasing the equality dimension.

5. Conclusions

This article shows that the differentiation between a mutual reinforcing relationship and a conflicting relationship between dimensions is fruitful: the former type of relationship expresses the interdependence of the dimensions, so that one dimension cannot exist without the other. The latter relationship captures the idea that not all dimensions can be maximized simultaneously: "ramping up" a dimension to the highest degree possible taking a loss in another dimension. This means that every democracy has to choose between a precarious balance of contentious dimensions.

To detect the relevant trade-offs, the starting point of our discussion are fundamental conceptions of democracy, mainly embedded in the Democracy Matrix. We propose a set of trade-offs for libertarian vs. egalitarian, and majoritarian vs. consensus conceptions of democracy. Finally, we construct three ideal types of profiles of democracy on the basis of the trade-offs: a libertarian, an egalitarian, and a control-focused profile.

We introduce a new measurement and aggregation strategy justified by a conceptual foundation. The two types of relationships are operationalized using on the one hand quality-measuring indicators and on the other hand trade-off indicators. While quality-measuring indictors follow a unidimensional interpretation—the current standard in this research field, the additional measurement with trade-off indicators uses a double evaluation linked to different dimensions. Using the V-Dem dataset, we showed preliminary empirical findings of our new measurement tool, the Democracy Matrix. These findings indicate that we can empirically discover our proposed ideal typical democracy profiles but that there is a wide variety of hybrid profiles as well.

This result differs from the findings of the Democracy Barometer and the Varieties of Democracy which both are not capable of detecting trade-offs. The choice of theoretical conception as well as the measurement strategy matters. Whereas the Democracy Barometer proposes one specific type of democracy as being the benchmark for the quality of democracy (the liberal democracy in the form of the embedded democracy),[6] V-Dem offers several types. But in contrast to the Democracy Matrix, it lacks a comprehensive meta theory of trade-offs which not only recognizes those types as normative equal but focuses on their interwoven relationship as well.

New research questions consist in the further theoretical and empirical examination of the different profiles of democracy: On the theoretical level, it seems that all trade-offs involve the freedom dimension. Why is the freedom dimension so significant in this regard? Is it more relevant than the other two dimensions? On the empirical level, we can analyze the change of profiles throughout time in countries (e.g. Belgium's change to more egalitarian values since the 1980s) and their respective causes. Another question concerns the interplay of democracy profiles and governance structures as well as policy-outputs (similar to Lijphart, 2012). Does the control-focused profile have a higher risk of a reform grid-

[6] Similar to the Democracy Barometer, Munck's (2016) well informed reconceptualization of the quality of democracy ranks one specific democracy model higher than the others. He proposes a majoritarian democracy—mixed with PR elections—as a benchmark for the quality of democracy. This seems, however, highly problematic.

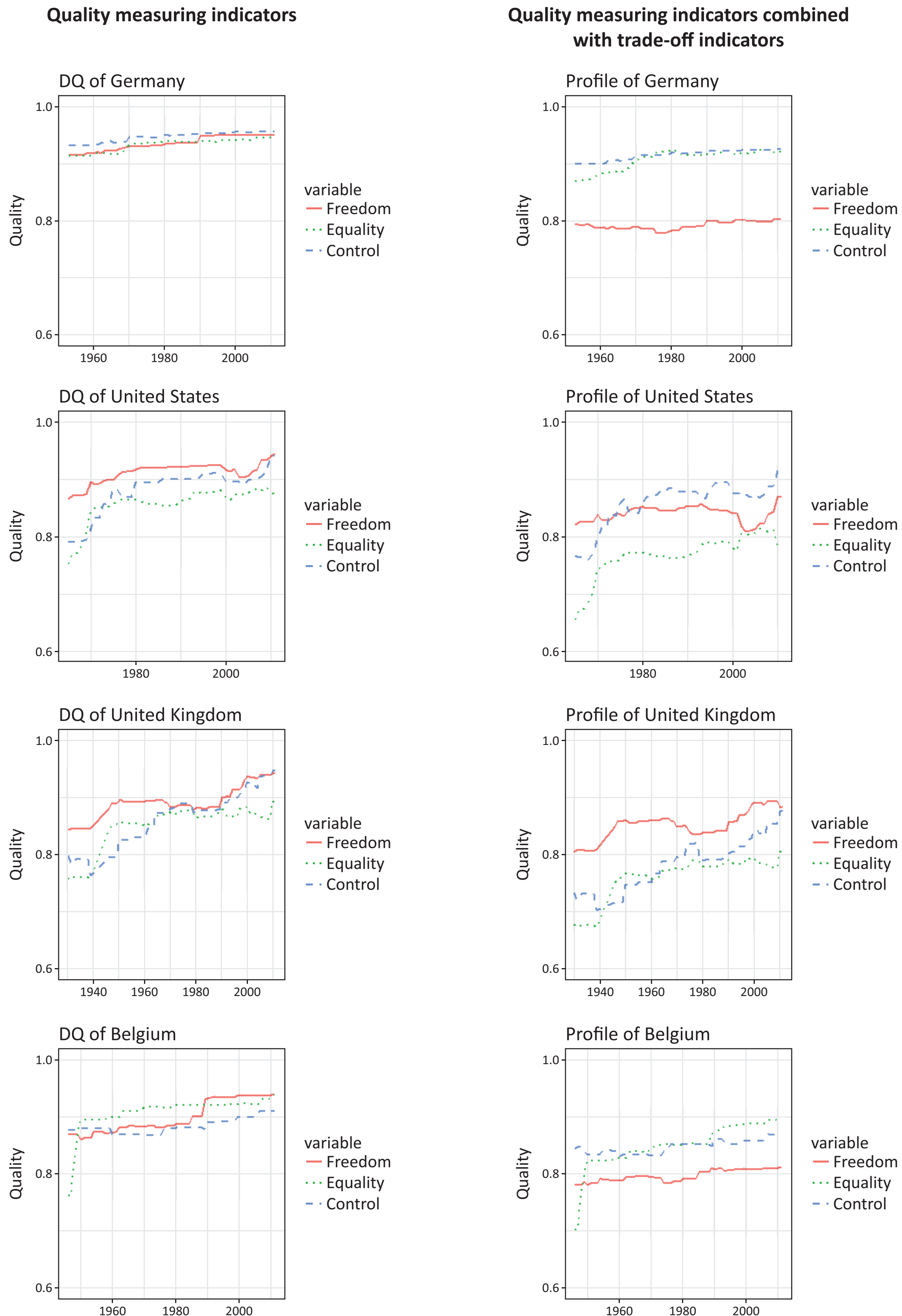


Figure 4. Empirical results. Source: Own calucation based on the V-Dem-Dataset (Version 6.2; Coppedge et al., 2016). The left side shows the results for the quality measuring indicators; the right side presents the findings of the new measurement approach which combines the quality measuring and profile measuring indicators. The colors are as follows: red = freedom dimension, blue = control dimension, green = equality dimension.

lock? Does the egalitarian profile coincide with a strong welfare state? These research questions are opened up by the conceptualization and measurement of trade-offs and should be approached in future research projects.

Acknowledgments

The authors would like to thank the participants in the Workshop "Why Choice Matters—Revisiting and Comparing Measures of Democracy" (1–2 June 2017, in Zurich). They would also like to thank Gary S. Schaal as well as the anonymous reviewers for their careful reading and insightful comments.

Conflict of Interests

The authors declare no conflict of interests.

References

Berlin, I. (1969). *Four essays on liberty*. Oxford: Oxford University Press.

Campbell, D. F. J., Carayannis, E. G., & Scheherazade, S. R. (2015). Quadruple helix structures of quality of democracy in innovation systems: The USA, OECD countries, and EU member countries in global comparison. *Journal of the Knowledge Economy*, *6*(3), 467–493.

Coppedge, M., Gerring, J., Altman, D., & Bernhard, M. (2011). Conceptualizing and measuring democracy: A new approach. *Perspectives on Politics*, *9*(1), 247–267.

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S., Teorell, J., Altman, D., . . . Zimmermann, B. (2016). *V-Dem [Country-year/country-date] Dataset v6.2*. Gothenburg: Varieties of Democracy (V-Dem) Project.

Coppedge, M., Lindberg, S. I., Skaaning S., & Teorell, J. (2015). *Measuring high level democratic principles using the V-Dem Data* (Working Paper No. 6). Gothenburg: The Varieties of Democracy Institute.

Dahl, R. A. (1956). *A preface to democratic theory*. Chicago, IL: University of Chicago Press.

Dahl, R. A. (1971). *Polyarchy: Participation and opposition*. New Haven, CT: Yale University Press.

Diamond, L., & Morlino, L. (2004). The quality of democracy. An overview. *Journal of Democracy*, *15*(4), 20–31.

Diamond, L., & Morlino, L. (2005). *Assessing the quality of democracy*. Baltimore, MD: John Hopkins University Press.

Dworkin, R. (1996). Do liberty and equality conflict? In P. Barker (Ed.), *Living as equals* (pp. 39–57). Oxford: Oxford University Press.

Elster, J., & Slagstadt, R. (1988). *Constitutionalism and democracy*. Cambridge: Cambridge University Press.

Giebler, H., & Merkel, W. (2016). Freedom and equality in democracies: Is there a trade-off? *International Political Science Review*, *37*(5), 594–605.

Hidalgo, O. (2014). *Antinomien der Demokratie* [*Antinomies of democracy*]. Frankfurt and New York, NY: Campus.

Lauth, H. (2004). *Demokratie und Demokratiemessung* [*Democracy and the measurement of democracy*]. Wiesbaden: VS Verlag.

Lauth, H. (2015). *The matrix of democracy: A three-dimensional approach to measuring the quality of democracy and regime transformations* (Würzburg Working Paper Series in Political Science and Sociology, WAPS no. 6). Würzburg: University of Würzburg.

Lauth, H. (2016). The internal relationships of the dimensions of democracy: The relevance of trade-offs for measuring the quality of democracy. *International Political Science Review*, *37*(5), 606–617.

Lijphart, A. (2012). *Patterns of democracy. Government forms and performance in thirty-six countries* (2nd ed.). New Haven, CT and London: Yale University Press.

Merkel, W., Bochsler, D., Bousbah, K., Bühlmann, M., Giebler, H., Hänni, M., . . . Wessels, B. (2016). *Democracy barometer. Methodology (Version 5)*. Aarau: Zentrum für Demokratie.

Munck, G. L. (2016). What is democracy? A reconceptualization of the quality of democracy. *Democratization*, *23*(1), 1–26.

Nohlen, D. (2014). *Wahlrecht und Parteiensystem. Zur Theorie und Empirie der Wahlsysteme* [*Electoral law and party system. Theory and empirical findings of electoral systems*] (7th ed.). Opladen and Toronto: Leske & Budrich.

O'Donnell, G. (1994). Delegative democracy. *Journal of Democracy*, *5*(1), 55–69.

Ohman, M. (2014). Getting the political finance system right. In E. Falguera, S. Jones, & M. Ohman (Eds.), *Funding of political parties and election campaigns—A handbook on political finance* (pp. 13–34). Stockholm: IDEA.

Smilov, D. (2008). *Dilemmas for a democratic society: Comparative regulation of money and politics* (DISC Working Paper Series no. 4). Budapest: Center for the Study of Imperfections in Democracy.

Strohmeier, G. (2011). Westminster im Wandel [The changing nature of the Westminster system]. *Aus Politik und Zeitgeschichte*, *61*(4), 32–40.

About the Authors

Hans-Joachim Lauth is Professor for Comparative Politics and German Government at the Institute for Political Science and Sociology, University of Wuerzburg since 2008. He received his PhD at the University of Mainz in 1991. In 2002 he finished his second book (*Habilitation*) about Democratic Theory and Measuring Democracy. His current topics of research concern democracy, informal institutions, Rule of Law and separation of powers (accountability), civil society and comparative methods.

Oliver Schlenkrich is a PhD student at the Chair of Comparative Politics and German Government, University of Wuerzburg. He currently works on the DFG research project led by Prof. Dr. Hans-Joachim Lauth "The democracy matrix as an alternative to the democracy indices of Freedom House and Polity: Implementation of the Varieties-of-Democracy-Data by using the 15-Field-Matrix of Democracy". His research interests concern democracy, political culture, political participation, quality of statehood and quantitative methods.

Annex

1. Aggregation Method of the Democracy Matrix

1.1. Aggregation Method for Quality Measuring Indicators

Here, we use a very similar aggregation technique as V-Dem (Coppedge et al., 2016).

- The lowest parts of our concept trees (indicators: V-Dem relative scale): mean;
- CDF of the mean (normal distribution with mean 0 and standard deviation 1) resulting in a value between 0 and 1;
- We use the following formula (see V-Dem) for aggregation up to the matrix field level:

$$Component_A * Component_B * 0.5 + Component_A * 0.25 + Component_B * 0.25$$

  The left side of this formula ($Component_A * Component_B * 0.5$) incorporates the necessary condition (no compensation) whereas the right part of the formula ($Component_A * 0.25 + Component_B * 0.25$) allows some compensation. If our concept demands a strict necessary condition, we weight the left part of the formula more heavily (> 0.5) in contrast to the right part (< 0.5). If our concept does require a "softer" necessary condition, more weight to the right part is given ($> 0,5$) at the cost of the left part (< 0.5). In addition, the right side of the formula allows for a precise weighting of the different components;
- Dimensions (transformational status): multiplication of the related matrix fields to the power of (1/5):

$$\text{Dim}_{\text{qual}} = (\text{Field}_{\text{qual1}} * \text{Field}_{\text{qual2}} * \text{Field}_{\text{qual3}} * \text{Field}_{\text{qual4}} * \text{Field}_{\text{qual5}})\hat{}(1/5)$$

1.2. Aggregation Method for Profile Measuring Indicators

- We use only a subset of the V-Dem-Data: every country which has a value > 0.5 in every matrix field measured by the quality measuring indicators is included;
- For each trade-off indicator the empirical minimum is set to 0.75, empirical maximum to 1;
- Dual use of these indicators, e.g. libertarian party finance models gain a 1 in the freedom dimension and a 0.75 in the equality dimension; egalitarian party finance models gain 0.75 in the freedom dimension and a 1 in the equality dimension;
- If more than one trade-off is located in a matrix field, weighting applies.

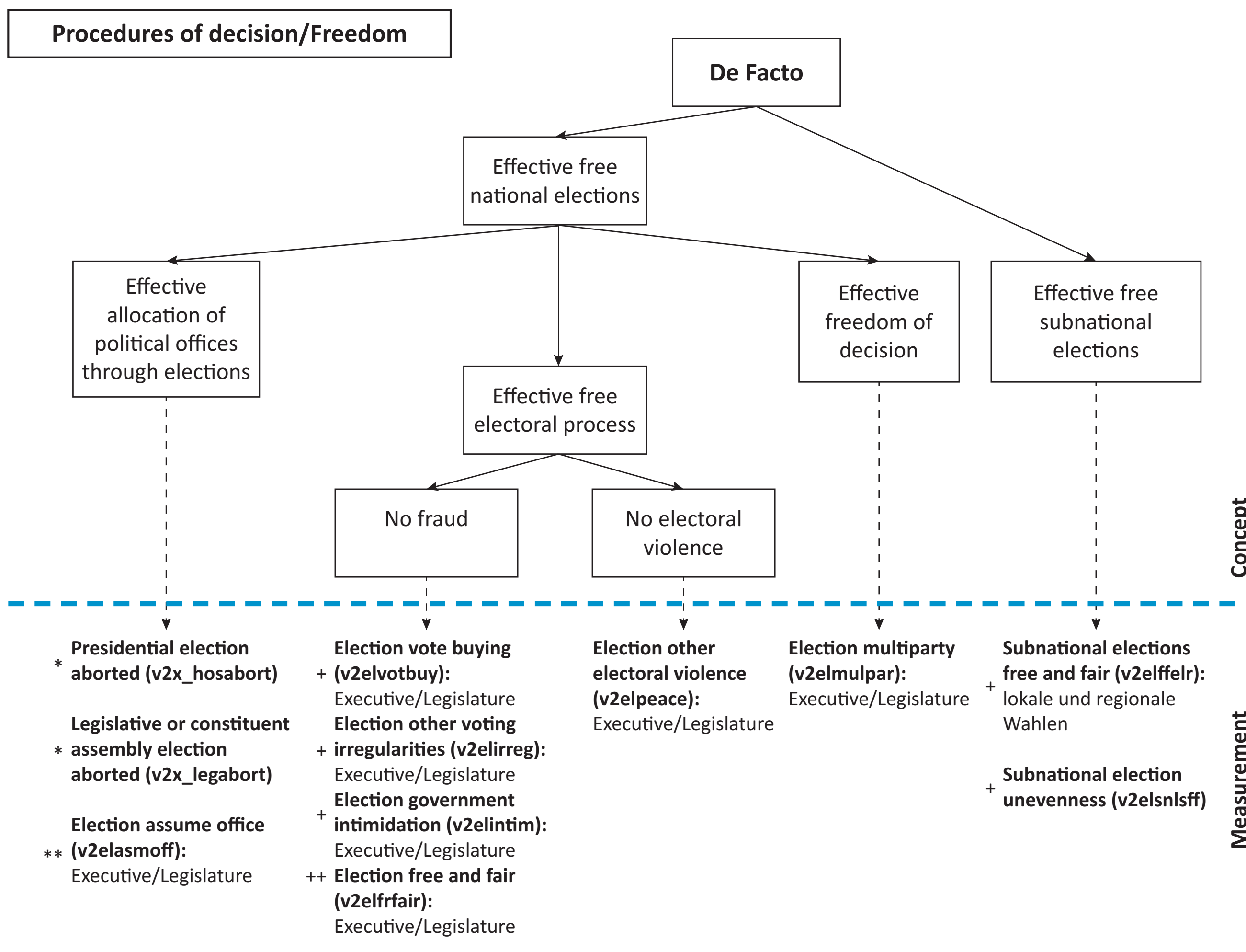
1.3. Combining Quality Measuring and Profile Measuring Indicators

- The empirical results of the matrix fields measured by the quality measuring indicators are multiplied with the related trade-offs:

$$\text{Field}_{\text{profile}} = \text{Field}_{\text{qual}} * \text{trade_off}$$

- Dimensions (Democracy profiles): Multiplication of the related matrix fields to the power of (1/5):

$$\text{Dim}_{\text{profile}} = (\text{Field}_{\text{profile1}} * \text{Field}_{\text{profile2}} * \text{Field}_{\text{profile3}} * \text{Field}_{\text{profile4}} * \text{Field}_{\text{profile5}})\hat{}(1/5)$$



Free procedures of decision include nationa and sub-national elections (i.e. regional and local elections). Local elections only if such a level in the political system exists and has appropriate decision-making powers. Both are necessary and together sufficient conditions, the weighting follows the strength of the relative decision-making power between these levels. National elections can be classified as free, if three necessary and sufficient conditions are met: first, effective allocation of political offices through elections in the sense that an election winner can actually take office. Second, effective freedom of decision must be present, i.e. that citizens can actually select from different parties. Finally, the electoral process must be free, so that neither serious fraud nor electoral violence occur. The first two conditions are very important: an election is not free in which the winners are hindered to take offices. Even an election in which only one party / candidate can be elected is not free as well.

Figure A1. Concept tree for matrix field "Procedures of decision/Freedom". Source: Own presentation; the indicators can be found in the V-Dem-Codebook (Coppedge et al., 2016).

References

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S., Teorell, J., Altman, D., . . . Miri, F. (2016). *V-Dem codebook v6*. Gothenburg: Varieties of Democracy (V-Dem) Project.

Politics and Governance (ISSN: 2183–2463)
2018, Volume 6, Issue 1, Pages 92–104
DOI: 10.17645/pag.v6i1.1235

Article

Method Factors in Democracy Indicators

Martin Elff [1] and Sebastian Ziaja [2,*]

[1] Faculty of Political & Social Sciences, Zeppelin University, 88045 Friedrichshafen, Germany; E-Mail: martin.elff@zu.de
[2] Research Center for Distributional Conflict and Globalization, Heidelberg University, 69115 Heidelberg, Germany; E-Mail: ziaja@uni-heidelberg.de

* Corresponding author

Submitted: 20 October 2017 | Accepted: 22 December 2017 | Published: 19 March 2018

Abstract
Method factors represent variance common to indicators from the same data source. Detecting method factors can help uncover systematic bias in data sources. This article employs confirmatory factor analysis (CFA) to detect method factors in 23 democracy indicators from four popular data sources: The Economist Intelligence Unit (EIU), Freedom House, Polity IV, and the Varieties of Democracy (V-Dem) project. Using three different multi-dimensional concepts of democracy as starting points, we find strong evidence for method factors in all sources. Method-specific factors are strongest when yearly changes in the scores are assessed. The sources find it easier to agree on long-term average scores. We discuss the implications for applied researchers.

Keywords
confirmatory factor analysis; democracy; democracy indicators; measurement; method bias

Issue
This article is part of the issue "Why Choice Matters: Revisiting and Comparing Measures of Democracy", edited by Heiko Giebler (WZB Berlin Social Science Center, Germany), Saskia P. Ruth (German Institute of Global and Area Studies, Germany), and Dag Tanneberg (University of Potsdam, Germany).

© 2018 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Are measures of political regime characteristics systematically influenced (or biased) by the institutions that created them? We answer this question by assessing whether democracy indicators coming from the same source exhibit common deviations not found in indicators originating from other sources. Kenneth Bollen (1993) referred to these common deviations as 'method factors'. That is, we assess whether democracy indicators are affected by method factors.

Method factors should be of concern for the applied researcher because method factors can be a sign of systematic bias in the data source, i.e., 'method bias'. But biased indicators do not just lead to improper descriptions of the issue to be measured. Even erroneous conclusions can be drawn from inferential studies if bias in a democracy indicator is correlated with an explanatory variable such as economic liberalization or ethnic fragmentation. In a field with immediate policy implications such as democratization research, false conclusions can cause actual harm. There is peril in informing policy with biased indicators.

This decade has seen a renewed interest in measuring democracy and explaining why it succeeds in some places and fails in others. The interest was sparked by initially promising signs of liberalization in the Middle East, now known as the Arab spring, as well as by formally democratized countries backsliding into autocratic practice. Such backsliding has occurred prominently in Russia, Turkey and Venezuela, but also in member states of the European Union, such as Poland and Hungary.

Responding to the desire to better track these developments, several new producers of democracy data have come forward. Certainly, the most notable addition to the group of democracy data producers is the Varieties

of Democracy (V-Dem) project. It boasts thousands of experts who have been involved in coding a new dataset on democracy for almost all countries since 1900. 'Most [experts] have lived in their countries of expertise for nearly thirty years, and at least 60 percent are nationals of that country' (Lindberg, Coppedge, Gerring, & Teorell, 2014, p. 162). The size and diversity of V-Dem's expert group contrasts with existing data sources, such as the Polity IV project or the FH data, which rely on a much smaller number of experts who are predominantly citizens of the United States. This difference in expert groups raises the question whether V-Dem data differs systematically from existing data sources.

The suspicion that traits of the expert group could influence indicators has already been voiced by Bollen (1993, p. 1213). He lists political attitudes of coders, incomplete information, and the aggregation of individual indicators into indices of democracy as potential sources of method bias. All of these issues are present to varying degrees in all social science indicators. One might assess these issues from varying perspectives, examining underlying concepts, data collection and aggregation procedures.

In this article, we focus on the indicator scores that the sources produce. We employ a 'convergent/divergent validation' approach (Adcock & Collier, 2001, p. 540): indicators from one source representing a particular conceptual dimension of democracy should be similar to indicators from other sources representing the same dimension—they should converge. Indicators from the same source representing different conceptual dimensions of democracy should not be as similar—they should diverge to some degree.

Our specific strategy is inspired by Bollen's seminal 1993 article 'Liberal Democracy: Validity and Method Factors in Cross-National Measures'. He employs confirmatory factor analysis (CFA). His model allows a range of indicators to load on two latent dimensions of democracy—*political liberties* and *democratic rule*—and on latent factors representing the sources of the indicators. This enables him to assess the amount of systematic error that indicators from the same source exhibit, i.e., the method factors. The sources he considers are Arthur Banks' *Cross-National Time-Series Data Archive* as well as Freedom House's *Freedom in the World* and *Freedom in the Press*. Bollen's analysis has not been updated with current democracy indices. Giebler (2012, p. 510) argues that most approaches comparing democracy measures focus excessively on conceptual differences, rather than on methodological differences. A study that has paid much attention to systemic bias in democracy ratings in the past decade is Pemstein, Meserve and Melton's (2010, pp. 444–446) presentation of their *Unified Democracy Scores*. What distinguishes their approach from ours is that they employ top-level indices that attempt to capture democracy as a whole. Also, Treier and Jackman (2008), in an attempt to detect measurement error in the Polity data, employ a unidimensional model.

Our contribution to the literature is an update and extension to Bollen's approach that evaluates various measures of regime characteristics that have been used as indicators of unidimensional or multidimensional conceptualisations of democracy. We update his approach by employing an updated set of four sources and 23 indicators and three different conceptual frameworks in our analysis. We extend his approach by using data with temporal variation over the period 2006 to 2014 (and 1972 to 2014 for an alternative dataset with three sources). Temporal variation provides us with the ability to provide a more nuanced assessment: do sources agree on average on grading countries by traits of democracy and do they agree about the timing and direction of the changes taking place over time?

The main result of our analysis is that most if not all measures of regime characteristics under study exhibit method factors. That is, these measures are influenced to a non-ignorable degree by the sources or institutes that produce them. The objectivity of these measures is thus limited. This is particularly salient when one considers changes over time in democracy indicators: different sources make very different assessments on changes in political and civil rights. In the cross-sectional view, however, sources show less pronounced method factors and more substantial agreement. Considering individual sources, method factors are largest for some of the Polity IV indicators.

2. Background: Three Concepts of Democracy

Our aim is to examine whether current measures of democracy are affected by the institutions or research groups by which they are produced or by the data sources from which they are derived. To ensure that our findings regarding the existence of method factors do not depend on a particular conception of democracy, we check for the existence of method factors against the backdrop of each of three different conceptualisations: a two-dimensional one, which has already been employed in Bollen's (1993) study, a three-dimensional conceptualisation put forward by the V-Dem project (Coppedge, Gerring, Lindberg, Skaaning, & Teorell, 2017a), and finally a conceptualisation with no less than four dimensions, inspired by Merkel (2004).

Bollen's two-dimensional scheme entails *political liberties* and *democratic rule*. This approach is rather minimalist and introduces only one distinction: between individuals' abilities to participate in the political system, and the functioning of the latter in the spirit of democracy. Using this approach has the advantage of making our analysis comparable to Bollen's. The dimensions are defined as follows: *Political liberties* 'exist to the extent that the people of a country have the freedom to express a variety of opinions in any media and the freedom to form or to participate in any political group' (Bollen, 1993, p. 1208). *Democratic rule* '(or political rights) exists to the extent that the national government is ac-

countable to the general population, and each individual is entitled to participate in the government directly or through representatives' (Bollen, 1993, p. 1209).

Since we have more indicators available than Bollen had in 1993, we can test more detailed models. V-Dem suggests a scheme which entails seven 'principles' of democracy: electoral, liberal, participatory, majoritarian, consensual, deliberative and egalitarian (Coppedge et al., 2017a, pp. 20–25). We employ only the first three principles here and exclude the latter four. We exclude majoritarian and consensual, as these principles do not have a more democratic and a less democratic pole (Lijphart, 2012)—rather they refer to variations of democracy which most measurement projects do not tap into. Measuring majoritarian and consensual principles across all countries is challenging, and even V-Dem abstains from quantifying these concepts at the moment (Coppedge et al., 2017a, p. 24). We exclude the deliberative principle because it refers to the rationality of political debates, which is not measured directly by other sources. We exclude the egalitarian principle because it refers to individual socio-economic prerequisites for political empowerment, which includes financial resources and thus seems to overstretch the concept of democracy.

The *electoral principle* of V-Dem captures the core idea of polyarchy, i.e., 'making rulers responsive to citizens through periodic elections' (Coppedge et al., 2017a, p. 20). The *liberal principle* refers to individual rights against state repression, guaranteed by 'constitutionally protected civil liberties, strong rule of law, and effective checks and balances that limit the use of executive power' (Coppedge et al., 2017a, p. 21). The *participatory principle* 'embodies the values of direct rule and active participation by citizens in all political processes' (Coppedge et al., 2017a, p. 22). 'Direct rule' entails problems similar to that encountered in the majoritarian and consensual principles: more direct rule does not necessarily imply more democracy, and may erode into a tyranny of the majority. Moreover, most sources abstain from measuring issues of direct democracy separately. We thus focus on 'active participation' as conceptual focus of this dimension.

A more detailed theoretical model which is independent of our data sources is Wolfgang Merkel's (2004) concept of 'embedded democracy'. Embedded democracy has an *electoral regime* at its core, which is complemented by *political liberties*, *civil rights*, *horizontal accountability*, and the *effective power to govern* (Merkel, 2004, p. 37). Referring to Dahl (1971), Merkel (2004, p. 38) describes the *electoral regime* to entail 'universal, active suffrage, universal, passive right to vote, free and fair elections and elected representatives'. *Political rights* 'complete the vertical dimension of democracy and make the public arena an independent political sphere of action, where organizational and communicative power is developed' (Merkel, 2004, p. 38). This requires freedom of association and freedom of expression. Political rights provide the input for the electoral regime, which would be lacking input without the former. *Civil rights* maintain the rule of law by containing state power. 'The actual core of the liberal rule of law lies in basic constitutional rights' (Merkel, 2004, p. 39). This requires an independent judiciary as a guarantor. Civil rights thus constitute negative rights in the political system, whereas political rights constitute positive rights. Merkel refers to Guillermo O'Donnell (1994) with the definition of *horizontal accountability*, which requires 'that elected authorities are surveyed by a network of relatively autonomous institutions' (Merkel, 2004, p. 40). This is necessary since the vertical forms of control provided by the three preceding institutions 'control the government only periodically'. The *effective power to govern* finally asserts that the elected representatives are actually in control of the state (Merkel, 2004, p. 41). We disregard this last requirement here, as only a few sources attempt to measure it.

3. Data

The data sources we consider beyond V-Dem are the Economist Intelligence Unit's (EIU) democracy index (Economist Intelligence Unit, 2014), Freedom House (FH; Freedom House, 2016) and the Polity IV project (Marshall, Gurr, & Jaggers, 2016). As Bollen (1993, p. 1210) does, we focus on subjective measures, i.e., indicators of *de facto* democratic quality, not *de jure* provisions. The former are also more susceptible to systematic biases and deserve a closer inspection in this regard. Both this focus and the limited coverage of other sources explain why we constrain our analysis to four sources at the present.

Other sources were excluded for providing discrete regime types instead of linear measures of regime status (e.g. the Democracy-and-Dictatorship data by Cheibub, Gandhi, & Vreeland, 2010), for measuring *de jure* instead of *de facto* regime traits (e.g., the Database of Political Institutions by Beck, Clarke, Groff, Keefer, & Walsh, 2001), or for providing insufficient spatial or temporal coverage (e.g., the Bertelsmann Transformation Index, 2016). The Appendix lists additional sources that were excluded and gives reasons for these decisions.

Each data source considered here provides at least two levels of indicators: those at the lowest level, which are coded directly (by judgement, observation or other means), and those at intermediate and higher levels, which are aggregated from lower-level indicators. We employ data at an intermediate level of aggregation, i.e., indicators that are supposed to measure rather general attributes of democratic rule and political liberties, such as fair elections and freedom of speech. Disregarding very detailed indicators such as those provided by V-Dem allows us to maintain a roughly equal level of aggregation across sources, although a perfect alignment is not possible. The selection of the indicators and their assignment to conceptual dimensions is documented at length in the Appendix.

The indicators are collated into two data sets: a 'longer' data set which includes a smaller range of indicators for which data are available for a relatively long period, from 1972 to 2014, and a 'shorter' but 'wider' data set which includes a wider range of indicators for which data are available only for the relatively short period from 2006 to 2014. The 'longer' data set, to which we refer to as *D1*, contains 5,864 observations from 160 countries for 15 indicators, while the 'shorter' data set, referred to as *D2*, contains 1,070 observations from 157 countries for 23 indicators. The EIU's democracy data and detailed indicators are only available for more recent years and appear in *D2*, but not in *D1*.

The following results focus on *D2*. Results pertaining to *D1* confirm the general findings from *D2* and can be found in the Appendix. A description of *D2* is given by Table 1, which indicates the source of the indicators and what dimensions they represent according to conceptualisations of political regimes with different numbers of dimensions. The Appendix also provides tables with summary statistics for both datasets. All data was obtained via the *Quality of Government* database (Teorell et al., 2017) and the V-Dem dataset version 7.1 (Coppedge et al., 2017b).

Table 1. Indicators and dimension assignment in data set *D2*.

Description	Source	2-Dimensional concept	3-Dimensional concept	4-Dimensional concept
Civil liberties	EIU	Political liberties	Liberal principle	Civil rights
Electoral process and pluralism	EIU	Democratic rule	Electoral principle	Electoral regime
Functioning of government	EIU	Democratic rule	Liberal principle	Horizontal accountability
Political participation	EIU	Political liberties	Participatory principle	Political rights
Associational and Organizational Rights	FH	Political liberties	Participatory principle	Political rights
Electoral Process	FH	Democratic rule	Electoral principle	Electoral regime
Freedom of Expression and Belief	FH	Political liberties	Liberal principle	Civil rights
Personal Autonomy and Individual Rights	FH	Political liberties	Liberal principle	Civil rights
Political Pluralism and Participation	FH	Democratic rule	Liberal principle	Political rights
Rule of Law	FH	Political liberties	Liberal principle	Civil rights
The competitiveness of participation (PARCOMP)	Polity	Democratic rule	Participatory principle	Political rights
Regulation of participation (PARREG)	Polity	Political liberties	Liberal principle	Political rights
Executive constraints (XCONST)	Polity	Democratic rule	Liberal principle	Horizontal accountability
Competitiveness of executive recruitment (XRCOMP)	Polity	Democratic rule	Electoral principle	Electoral regime
Openness of executive recruitment (XROPEN)	Polity	Political liberties	Participatory principle	Political rights
Civil society participation	V-Dem	Political liberties	Participatory principle	Political rights
Freedom of association (thick)	V-Dem	Political liberties	Electoral principle	Political rights
Freedom of expression (thick)	V-Dem	Political liberties	Liberal principle	Political rights
Judicial constraints on the executive	V-Dem	Democratic rule	Liberal principle	Horizontal accountability
Equality before the law and individual liberty	V-Dem	Political liberties	Liberal principle	Civil rights
Equal protection index	V-Dem	Political liberties	Liberal principle	Civil rights
Clean elections	V-Dem	Democratic rule	Electoral principle	Electoral regime
Legislative constraints on the executive	V-Dem	Democratic rule	Liberal principle	Horizontal accountability

4. Method

If measures of political regime characteristics are systematically influenced by the institutions that created them, then measures coming from the same institution should be more correlated than those coming from different institutions, at least after taking into account that these measures reflect certain substantive dimensions of regime characteristics. Equivalently, the variation of a measure of regime characteristics can be decomposed into a portion that can be attributed to a variation along conceptual dimensions such as, e.g. *democratic rule* and *political liberties*, and a portion that has to be attributed to the influence of the institution that created the measure. Confirmatory factor analysis (CFA) is the method of choice to examine whether such a decomposition is possible and adequate. In the context of confirmatory factor analysis this decomposition will take the form:

$$X_i = \alpha_i + \lambda_{ij} F_j + \kappa_{ik} G_k + U_i \quad (1)$$

where X_i is the value of the i-th regime characteristics indicator, F_i is the (unobserved) value of a common factor that represents the j-th conceptual dimension, G_k is the (unobserved) value of the common factor that represents the influence of the k-th institution, and U_i is a unique factor that represents random measurement error specific for the i-th indicator. For brevity, in the following we will refer to a common factor that represents a conceptual dimension of regime characteristics as a *conceptual factor*, while we refer to a common factor that represents the influence of the institution that created the indicator as a *method factor*.[1] The coefficient λ_{ij} of the conceptual factor F_j is referred to as the *loading* of indicator X_i on this factor. It is a parameter that is estimated in the context of a confirmatory factor analysis and represents how much the variation of the regime characteristics indicator is influenced by the conceptual factor. The coefficient κ_{ik} of the method factor G_k, which also can be referred to as a loading, is an estimated parameter that represents how much the indicator is influenced by the institute that has created it. Finally, α_j is an intercept that reflects the fact that, while the means of the common factors and the unique factor are assumed to be zero, the mean of X_i may be different from zero.[2]

Figure 1 illustrates a decomposition of the four indicators X_1, X_2, X_3, and X_4, into two conceptual factors F_1 and F_2, and into two method factors G_1 and G_2. The loadings in equation (1) are represented by arrows in Figure 1 as is the influence of the unique factors U_1, U_2, U_3, and U_4, which are represented by the empty circles. Figure 1 also illustrates some additional assumptions that we make in our analysis: firstly, that method factors are uncorrelated and that unique factors are uncorrelated, while conceptual factors may be correlated. These assumptions are motivated by the following considerations: the fact that one can distinguish between concepts such as *Democratic rule* and *Political liberties* does not imply that they are empirically uncorrelated. On the other hand, if the method factors are supposed to reflect influences that are *specific* to the institutions that create the indicators, this is best reflected by assuming the method factors to be uncorrelated. Otherwise, if we allowed the method factors to be correlated, such correlations would reflect commonalities among indicators of different institutions, which in turn could be attributed to the fact that they are supposed to measure the same phenomena or different aspects of the same phenomena. That is, allowing method factors to be correlated could contaminate them with correlations among indicators created by the substantial factors. As a consequence, this would lead to an overestimation of the relevance of method factors. Conversely, if fixing the correlations between method factors leads to an underestimation of the impact of method factors, then this means erring on the side of caution.

Confirmatory factor analysis does not only allow the estimation of factor loadings, variances, covariances, and correlations. More importantly, it allows the comparison of different models in terms of their fit to empirical data. Such model comparisons form the core of our research design. In order to assess the relevance of method factors, we conduct model comparison likelihood ratio tests with models which contain method factors against models which do not contain model factors. Such models can be obtained by deleting the term $\kappa_{ik} G_k$ from equa-

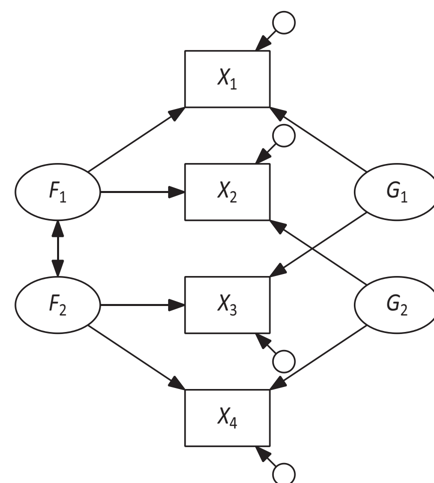


Figure 1. An Illustration of a factor model with conceptual factors and method factors.

[1] In factor analytic variants of multi-trait-multi-method analysis one would restrict the meaning of the term 'method factor' to common factors that represent the effects of a particular method of measurement. Since the influence of the creating institution on the regime indicators may be largely a consequence of the particular methods employed by this institution, we find it justifiable to use the term 'method factor' in a somewhat wider sense. If we had used a term like 'institutional factor' this might have led to confusion with conceptual factors that refer to institutional aspects of regime characteristics.

[2] Confirmatory factor analysis *per se* does not distinguish between different types of common factors. The distinction between two types of common factors, as reflected in different symbols used for the factors and their loadings, is a matter of interpretation that guides the construction of a factor model.

tion (1) and by deleting the nodes labelled G_1 and G_2 in Figure 1 as well as the arrows that connect them with the nodes labelled X_1, X_2, X_3, and X_4. The null hypothesis in these likelihood ratio tests is that a model without the method factors fits the data as well as a model that includes method factors. If a likelihood ratio test leads to the rejection of the null hypothesis, this means that an improvement of model fit brought about by the inclusion of method factors is more than a product of chance, which in turn provides evidence for the existence of method factors.

In order to make sure that our results are robust, we conduct our hypothesis tests based on the three different conceptualisations of the dimensionality of regime characteristics. That is, in the first variant the model that represents the null hypothesis has the two conceptual factors *Democratic rule* and *Political liberties*, in the second variant, the null model includes the three conceptual factors *Electoral principle*, *Liberal principle*, and *Participatory principle*, and in the third variant the null model includes the four conceptual factors *Civil rights*, *Electoral regime*, *Horizontal accountability*, and *Political rights*.

Apart from the identification problems that always lurk in complex CFA and structural equation models, our analysis is confronted with three related challenges: (1) non-normal and categorical indicators violate the standard assumptions on which likelihood-based inference in confirmatory factor analysis is based; (2) many indicators are conceptualised so that most democracies receive top scores—they are 'truncated'; (3) we observe the same countries at several points in time, which introduces dependencies not accounted for in standard CFA.

The first two challenges are illustrated by Figure 1: many indicators in our dataset place few countries in the centre of the empirical distribution and many at the extremes, in contrast to the shape of a normal distribution. Moreover, some of the indicators, in particular, the indicators from the Polity project, only have a small number of distinct values and therefore not have metric quality. In a situation like this maximum likelihood estimators may lose their asymptotic efficiency and test statistics may lead to false positives. For situations such as this Browne's (1984) asymptotically distribution-free estimator may retain asymptotic efficiency and provide relatively accurate test statistics. However, Browne's estimator requires a very large sample size, larger than the size of the data sets we employ in our analysis. For this reason, we stick to likelihood ratio test statistics and report Bollen-Stine bootstrap-based p-values (Bollen & Stine, 1992).

The second challenge is also illustrated by the histograms in Figure 2: contemporary democracies may be all too similar with respect to the regime indicators available in the data sets. Indicators such as Polity's *openness of executive recruitment* or V-Dem's *clean elections* show extreme peaks at the upper end of the scale. In order to address this problem, we repeat our analyses with subsets of the *D1* and *D2* data sets that contain only those regimes classified as non-democracies according to the democracy-and-dictatorship data (Cheibub et al., 2010; updated by Bjørnskov & Rode, 2017). We exclude all democracy regime types (codes 0 to 2) and retain all non-democracies (codes 3 to 5). Such a restriction of the data to non-democracies can eliminate some of the strongly peaked or U-shaped appearances in the histograms of the indicator variables, yet they still appear clearly non-normal.

The third challenge is that we use panel data, where measures are taken repeatedly from the same countries and therefore are not (conditionally) independent from one another. The methodology of CFA and structural equation modelling is mostly developed with cross-sectional data in mind, as is the available software to estimate such models. The (serial) dependence of measures taken from identical countries may or may not lead to biased estimates, but at least it will lead to inaccurate inference if standard errors are constructed based on the assumption of independence. We address this challenge with two approaches adopted from the econometrics of panel data (Baltagi, 2013): in a first approach, we fit our models to between country cross-sectional data constructed from the country-level means of the regime measures. This country-level aggregate data has a considerably smaller number of observations and thus a smaller power, but the serial dependence of the measures is eliminated. In a second approach, we keep the temporal information contained in the data and fit our models to within-country first differences of regime measures, i.e., to the amount of change compared to the previous year. This eliminates the between-country heterogeneity and reduces the serial dependence.

Considering four concepts with and without method factors means that we will have to estimate eight models for each of the two data sets we are employing (*D1* and *D2*), both for a full version of each data set and for the version reduced to non-democracies. In total, this gives 64 fitted models if we further distinguish between-country cross-sections and within-country first-differences. It is impossible to discuss the estimates based on this many model fits in a single article. Instead, we only discuss a series of chi-squared tests for model comparisons and present estimates only for models with four conceptual factors and four method factors fitted to data set *D2*. All models are estimated using the package *lavaan* (Rosseel, 2012) in the statistical environment *R* (R Core Team, 2017).

5. Results

As explained in the previous section, we conduct model comparison likelihood ratio tests to obtain evidence about the presence of method factors that represent the influence of the institutions on the regime indicators that they create. Each likelihood ratio test compares a model that contains only conceptual factors, common factors that represent only conceptual dimensions of regime

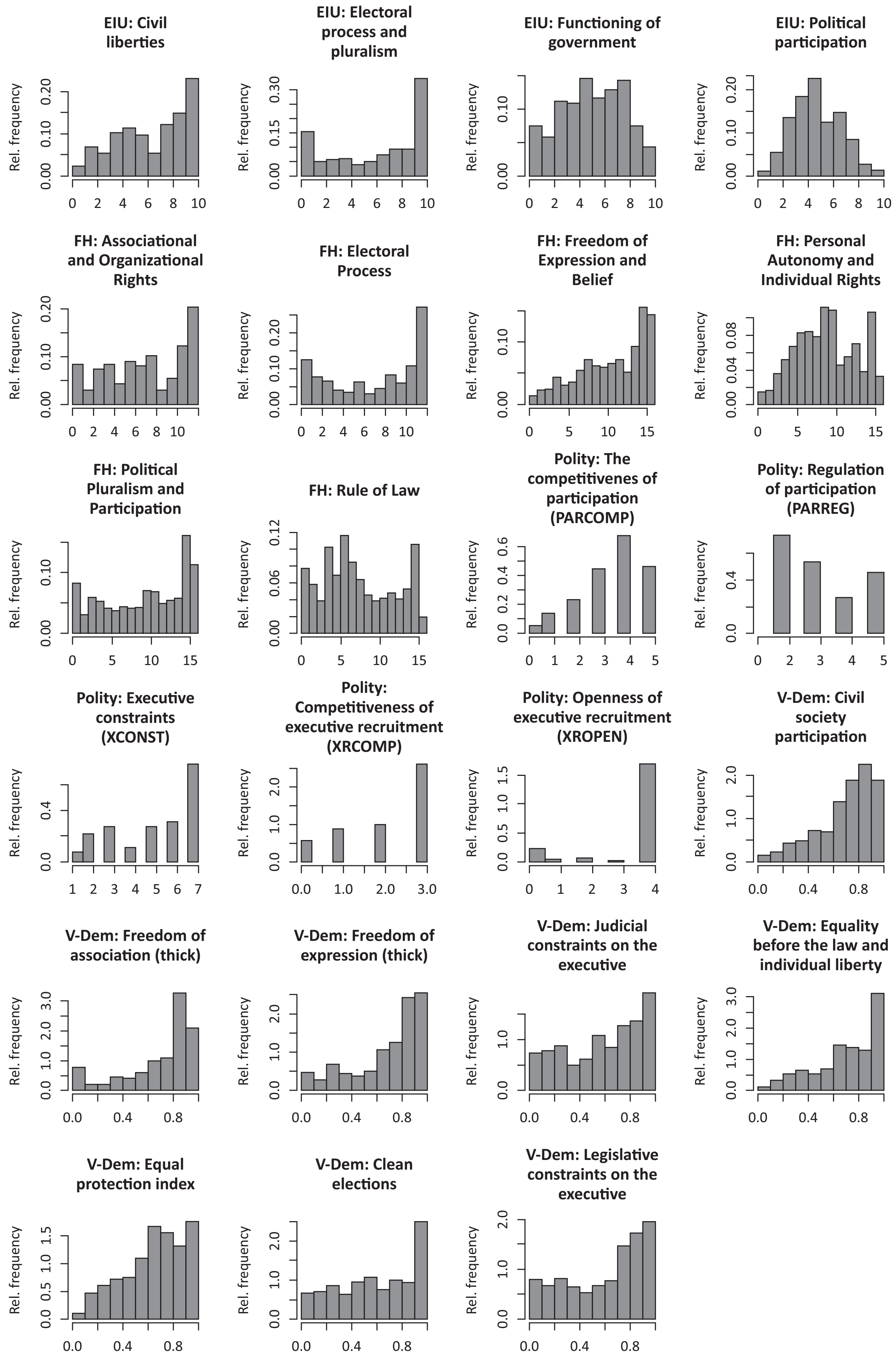


Figure 2. Histograms of the variables contained in *D2* (N = 1,070).

characteristics, with a model that additionally contains method factors corresponding to the various institutes that produced the indicators. If the likelihood ratio test indicates that inclusion of method factors leads to a statistically significant improvement of model fit, then we conclude that democracy measures indeed are affected by method factors. In order to make sure that our results are robust, we repeat the likelihood ratio test with different conceptualisations of regime dimensions as a baseline. Furthermore, we repeat the likelihood ratio tests with respect to the complete set of countries as well as only those countries categorized as authoritarian by Cheibub et al. (2010) and Bjørnskov and Rode (2017). To take into account the panel structure of the data, we conduct the tests first based on the between-country cross-section (i.e. the country averages) of the regime indicator values and second based on the within-country first differences of the indicator values. Table 2 shows the results of the hypothesis tests based on between-country cross-section while Table 3 shows the hypothesis tests results based on within-country first-differences. In addition to the values of the likelihood-ratio statistics, Tables 2 and 3 show three goodness-of-fit indices: the comparative fit index (CFI) which varies between 0 and 1, where 1 indicates perfect fit; the root mean square error of approximation (RMSEA), which also varies between 0 and 1, but where a value below 0,05 indicates an acceptable fit; and the standardised root mean squared residual (SRMR), which again usually varies between 0 and 1, with lower values indicating a better fit between the model and the data.

The results in Tables 2 and 3 are clear: no matter whether one assumes one, two, three, or four conceptual dimensions of regime properties, no matter whether one considers all countries or only non-democratic ones; the improvement of model fit by including method factors into the factor model is statistically significant at any conventional level. Furthermore, the goodness-of-fit indices also show a substantial improvement, except for SRMR in section (b) of Table 2. In summary, we find strong and robust evidence that method factors matter.

Having established the existence of method factors, we should discuss their relevance. How strong is their influence on measures of regime properties? This question can be answered by comparing the sizes of the loadings of the regime measures on the conceptual factors with their loadings on the method factors. In order to take into account the different scale lengths of the regime measures, such a comparison is best made using standardised estimates that are rescaled so that common factors and indicators all have unit variance. If the standardised loading of a regime measure on a method factor is as large as, or larger than, its loading on any conceptual factor then its validity should be considered questionable.

Figure 3 illustrates the factor loadings of the cross-section of regime indicators in the model that fit the data best, the factor model with four conceptual factors, and all four method factors. Diamonds represent conceptual factors, circles represent method factors, and rectangles represent regime measures. Each factor loading in the model is represented by an arrow, where the width indicates the absolute size of the standardised loading.[3] Overall, the

Table 2. Model comparison tests for the presence of method factors in data set *D2* (between-country cross-section).

(a) All countries								
	Deviance	Mod.Df	Chi-squared	Diff. Df	p-value	CFI	RMSEA	SRMR
1 Dimension	1419.8	253				0.813	0.171	0.055
+ Method factors	891.9	230	527.9	23	0.000	0.894	0.135	0.045
2 Dimensions	1368.4	252				0.821	0.168	0.056
+ Method factors	844.0	229	524.5	23	0.000	0.901	0.131	0.045
3 Dimensions	1395.8	250				0.816	0.171	0.055
+ Method factors	859.4	227	536.5	23	0.000	0.898	0.133	0.045
4 Dimensions	1314.3	247				0.828	0.166	0.055
+ Method factors	822.0	224	492.3	23	0.000	0.904	0.130	0.045
(b) Non-democratic countries only								
	Deviance	Mod.Df	Chi-squared	Diff. Df	p-value	CFI	RMSEA	SRMR
1 Dimension	695.1	253				0.697	0.177	0.102
+ Method factors	495.7	230	199.5	23	0.000	0.818	0.144	0.120
2 Dimensions	677.7	252				0.708	0.174	0.102
+ Method factors	460.6	229	217.1	23	0.000	0.841	0.134	0.114
3 Dimensions	693.7	250				0.696	0.178	0.101
+ Method factors	499.6	227	194.2	23	0.000	0.813	0.146	0.088
4 Dimensions	624.5	247				0.741	0.165	0.099
+ Method factors	438.9	224	185.6	23	0.000	0.853	0.131	0.104

[3] Path diagram of the confirmatory factor analysis model with four conceptual and for method factors—within-country first differences.

Table 3. Model comparison tests for the presence of method factors in data set *D2* (within-country first differences).

(a) All countries	Deviance	Mod.Df	Chi-squared	Diff. Df	p-value	CFI	RMSEA	SRMR
1 Dimension	4256.9	253				0.503	0.132	0.105
+ Method factors	1478.5	230	2778.4	23	0.000	0.845	0.077	0.055
2 Dimensions	4236.5	252				0.505	0.132	0.106
+ Method factors	1478.2	229	2758.3	23	0.000	0.845	0.077	0.055
3 Dimensions	4151.3	250				0.515	0.131	0.104
+ Method factors	1403.1	227	2748.1	23	0.000	0.854	0.075	0.054
4 Dimensions	4111.2	247				0.520	0.131	0.104
+ Method factors	1340.0	224	2771.2	23	0.000	0.861	0.074	0.061

(b) Non-democratic countries only	Deviance	Mod.Df	Chi-squared	Diff. Df	p-value	CFI	RMSEA	SRMR
1 Dimension	3123.6	253				0.424	0.113	0.100
+ Method factors	1140.7	230	1982.9	23	0.000	0.817	0.067	0.056
2 Dimensions	3106.7	252				0.427	0.113	0.102
+ Method factors	1140.5	229	1966.2	23	0.000	0.817	0.067	0.057
3 Dimensions	3065.7	250				0.435	0.112	0.100
+ Method factors	1126.4	227	1939.2	23	0.000	0.820	0.067	0.055
4 Dimensions	2977.3	247				0.452	0.111	0.102
+ Method factors	1016.8	224	1960.5	23	0.000	0.841	0.063	0.073

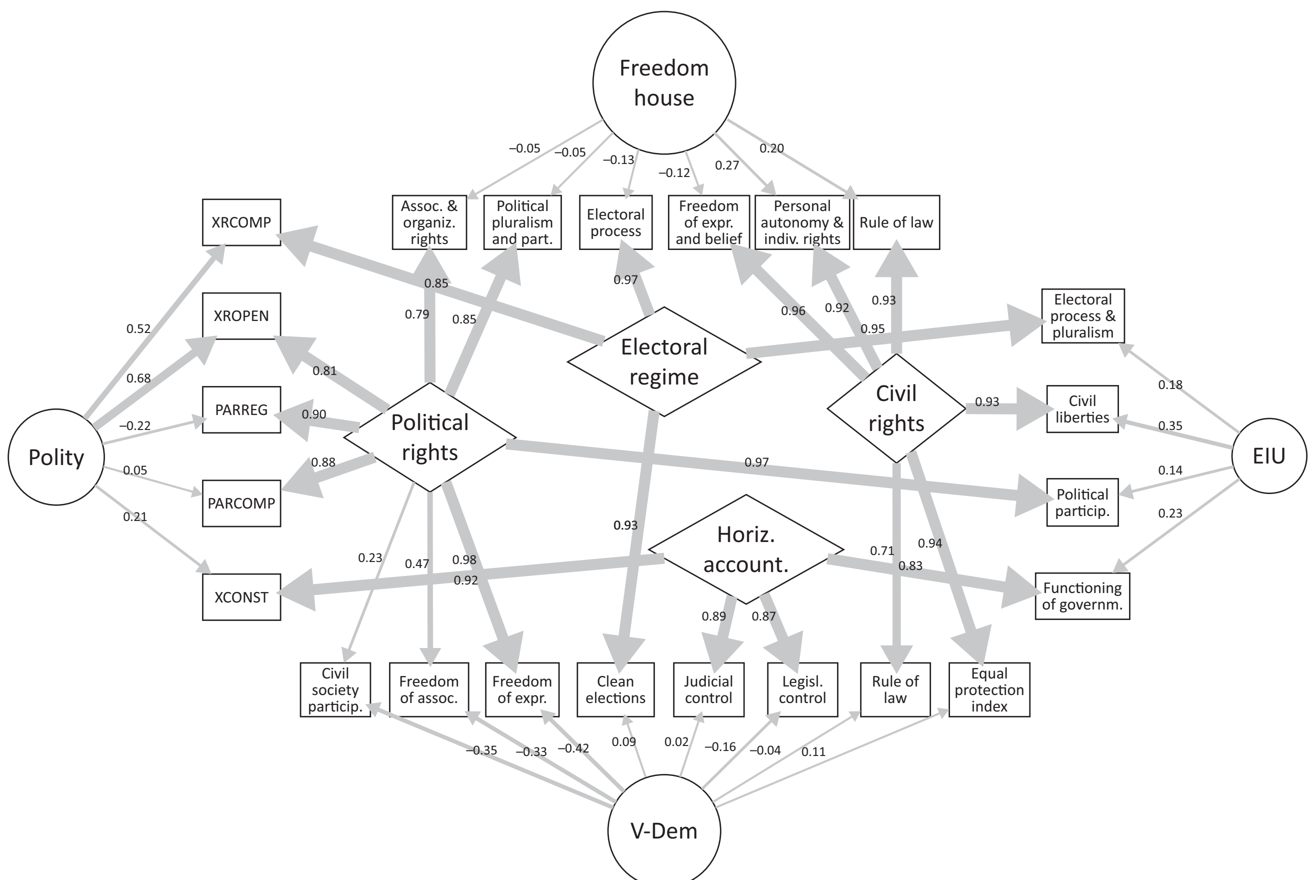


Figure 3. Path diagram of the confirmatory factor analysis model with four conceptual and for method factors—between country cross-section.

loadings of the regime measures on the conceptual factors are larger in absolute value than the loadings on the method factors. This is good news in so far as most regime measures indeed mostly represent substantial regime aspects. Yet, some of the loadings on the method factors are quite large. This affects, in particular, the regime measures from the Polity project: the indicators XRCOMP and XROPEN have strong loadings on the method factor. Even the relatively novel V-Dem measures do not seem to be without problems. All political rights indicators have relatively strong method factor loadings, and both *Civil society participation* and *Freedom of association* have small loadings on the conceptual factor. It appears that V-Dem contradicts more traditional sources on the relative position of countries in terms of political rights.

Figure 4, which illustrates loadings from a factor model fitted to within-country first differences, delivers an even less comforting message: the loadings on the method factors appear at least as large as the loadings on conceptual factors, thus raising doubts regarding the validity of many of the regime measures—at least when it comes to adequately representing change of regime properties within a country.[4] In particular, the Freedom House measures appear to be much more affected by the corresponding method factor than by any of the conceptual factors. Also, some of the Polity measures show very strong loadings on the method factor. In general, these strong loadings indicate that there is much less consensus between the various sources in terms of change than in terms of the average character of a country. This appears to affect the *Political rights* and *Civil rights* factors in particular, and much less so the *Electoral regime* and *Horizontal accountability* measures. A potential explanation for this pattern is that the latter refer to institutional characteristics that are rather easily observable in the form of laws and regulations, while changes of (effective) civil and political rights are unobservable latent properties and therefore more prone to follow a source's bias.

How can we make sense of the divergent assessments of method factors in democracy indicators that the two Figures suggest? One interpretation is that the different producers of democracy indicators vary in their sensitivity to change within countries—some adjust indicators earlier, others later (cf. Lueders & Lust, 2017). Method factors are less salient when it comes to country averages because the producers eventually converge to similar assessments once the dust raised by changing regime properties has settled. An alternative interpreta-

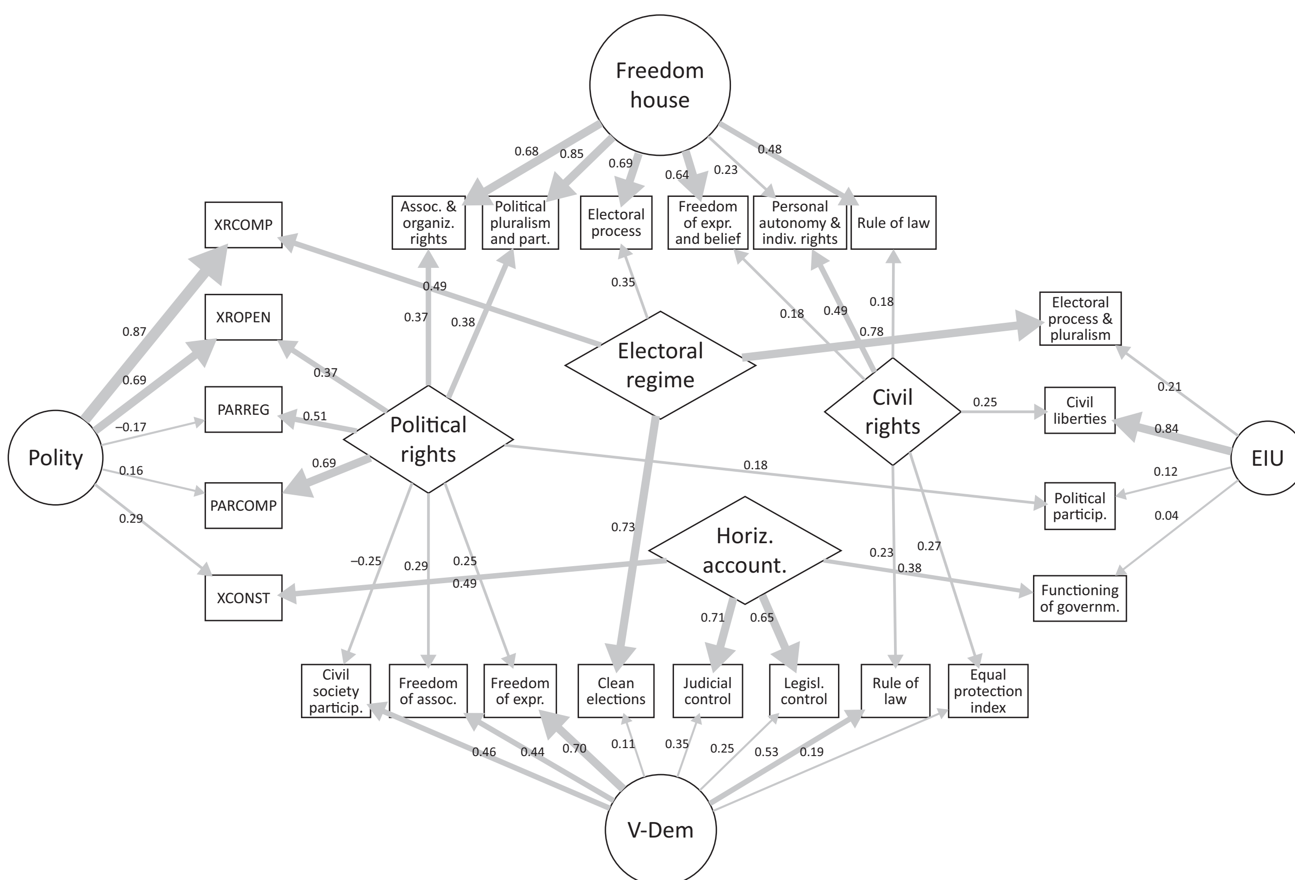


Figure 4. Path diagram of the confirmatory factor analysis model with four conceptual and for method factors—within-country first differences.

[4] One should not be misled by the all negative loadings on the Civil rights factor. This is empirically equivalent with a model fit where all loadings on this factor (and also the covariances with the other factors) have their signs reversed.

tion of the small loadings on the dimension factors in the within-country models is that within-country changes (in particular in civil and political rights) occur not all at the same time but in waves. As a consequence, the covariances between first differences would understate the actual patterns of changes. Alas, the large sizes of method factors make such a benevolent interpretation less likely.

Moreover, the temporal dependence of indicators from the same source has a plausible explanation: if a data producer comes to the conclusion that a larger regime change is occurring, they may also form the expectation that several indicators portraying different aspect of the regime would change at once. A desire to maintain consistency in regime measures may thus increase the correlation between indicators created by the same producer and decrease the correlation with regime measures created by other data producers.

6. Conclusion and Recommendations

Having analysed democracy indicators from four different sources, we find strong evidence for systematic deviations, i.e., method factors. The question of whether these method factors constitute biases in all cases, or whether one source is simply closer to measuring 'true' democracy cannot be answered here. We can state that while most sources converge on cross-sectional variation, they diverge on temporal variation within countries. Much of this uncertainty is not random error affecting individual indicators, but systematic error driven by the source.

As we analyse cross-sectional and within-country variation separately, we can render our speculations on the origins of this systematic error more precisely. Sources agree much more on the cross-sectional data than on the within-country differences.[5] This is a picture that could be explained by a practice of 'guessing until convergence'. Measuring change in democracy is difficult—in particular in the 'softer' dimensions of political and civil rights. When one data producer, at a certain point in time, perceives a change in a country, but others do not, agreement on the affected dimension declines. If those perceived changed stand the test of time, other data producers will follow and also adjust their scores. If the democratic practice observed in the country does not seem to warrant the change in scores, the first mover will revert their scores. As a result, we see substantial agreement in average cross-sectional assessments over the entire time period that we investigate. But we do not see much agreement on the timing of changes. It would be interesting to investigate whether a particular source is better at predicting change—a potential 'superforecaster' among democracy index producers.

Looking at individual sources, the largest method factors can be observed for some of the Polity IV indicators. This may in part be explained by the divergent conceptual setup of the Polity indicators: in the Appendix, we discuss the assignment of the indicators to the conceptual dimensions, and these decisions are more ambiguous for the Polity indicators. For example, Polity focuses more on the logic of ruler selection and less on participation, as exemplified by the neglect of suffrage issues (Munck & Verkuilen, 2002, p. 11). Freedom House exhibits less bias on the cross-section, but it on average it fares worst of all sources when assessing changes. In this light, Freedom House's self-declared mission 'to defend human rights and promote democratic change',[6] could inspire speculation that temporal distortions in Freedom House data are indeed intentional, with the aim to spur regime change. Previous studies have confirmed an ideological bias of this source (Giannone, 2010; Steiner, 2016). There is less critical literature on V-Dem yet, as most published work using the dataset comes from the large project team itself. The V-Dem team has also shown a large effort to assess and improve the quality of their data. For example, it has presented a comparison of its aggregate *polyarchy* score (a summary measure of electoral democracy) with other high-level indices (Teorell, Coppedge, Skaaning, & Lindberg, 2016, pp. 28–31). Nonetheless, some V-Dem indicators exhibit sizable method factors and should be investigated. We can say little about the EIU's democracy index beyond a diagnosis of moderate method factors, as hardly any complementary research on this source is available.

For applied researchers, our results shall serve as a reminder to adhere to some well-known but not always heeded rules of good practice. In a nutshell, these are (1) use the best source available, (2) use several sources, and (3) use meta indices. Determining the best sources available always depends on the research question at hand. An indicator with high conceptual validity for a particular application should certainly not be replaced with a more reliable measure that is far less valid. Contextual specificity matters (Adcock & Collier, 2001, p. 534). Among our set of indicators, however, we have many close matches that claim to measure very similar issues. In that case, given our results, our best guess for the indicator least affected by method bias will usually be the indicator provided by the V-Dem project. This is based not only on our model estimates but also on what we know about how the data is generated.

This scenario leads us to the second recommendation: should several indicators be available, use them! There is little additional effort in using multiple indicators to assess the robustness of results. Data collections such as those published by the *Quality of Government Institute* provide a large variety of indicators merged ready for the end user.

The third recommendation requires more preparation: the use of meta indices. For democracy as a unidimensional concept, various estimates exist. A prominent example based on a Bayesian measurement model is the

[5] Also note that the failure to agree on changes is all the more disappointing since we are employing yearly data, not monthly or weekly assessments.
[6] Available at https://freedomhouse.org/our-work

Unified Democracy Scores (Pemstein et al., 2010). For sub-dimensions of democracy, there are fewer meta indices available. Examples beyond Bollen's (1990) original approach are the contestation and participation scores provided by Coppedge, Alvarez and Maldonado (2008). In order to provide more choice on meta indices for sub-dimensions of democracy, one could employ the very factor scores that our models produce. This would provide quantitative measures for the dimensions of the Bollen concept, the V-Dem concept, and the Merkel concept. However, before using these in applied research, comprehensive additional vetting will be required, as validly measuring a substantial concept is more demanding than validly detecting method bias.

Our advice to producers of democracy indicators who pursue the goal of unbiased measures of democracy is to further address issues of method bias along all stages of the measurement process with various methodological approaches (see McMann, Pemstein, Seim, Teorell, & Lindberg, 2016 for an example) and with reference to alternative sources. Coppedge et al. (2017a), for example, have taken first steps to compare V-Dem to its main competitors. Additional efforts to more precisely assess temporal change in unobservable traits of democracy are advised.

Acknowledgments

We thank the editors of this thematic issue, two anonymous referees, participants at the Annual Meeting of the Methods Section of the German Political Science Association in May 2017, and participants at the V-Dem Lunch Seminar in September 2017 for comments and suggestions. We acknowledge financial support by Deutsche Forschungsgemeinschaft within the funding programme Open Access Publishing by the Baden-Württemberg Ministry of Science, Research and the Arts and by Ruprecht-Karls-Universität Heidelberg.

Conflict of Interests

The authors declare no conflict of interests.

References

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

Baltagi, B. H. (2013). *Econometric analysis of panel data* (5th ed.). Chichester: Wiley.

Beck, T., Clarke, G., Groff, A., Keefer, P., & Walsh, P. (2001). New tools in comparative political economy: The database of political institutions. *The World Bank Economic Review*, *15*(1), 165–176.

Bertelsmann Transformation Index. (2016). *Transformation index BTI 2016: Political management in international comparison*. Gütersloh: Bertelsmann Stiftung.

Bjørnskov, C., & Rode, M. (2017). *Regime types and regime change: A new dataset*. Aarhus: Aarhus University.

Bollen, K. A. (1990). Political democracy: Conceptual and measurement traps. *Studies in Comparative International Development*, *25*(1), 7–24.

Bollen, K. A. (1993). Liberal democracy: Validity and method factors in cross-national measures. *American Journal of Political Science*, *37*(4), 1207–1230.

Bollen, K. A., & Stine, R. A. (1992). Bootstrapping goodness-of-fit measures in structural equation models. *Sociological Methods & Research*, *21*(2), 205–229.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, *37*(1), 62–83.

Cheibub, J. A., Gandhi, J., & Vreeland, J. R. (2010). Democracy and dictatorship revisited. *Public Choice*, *143*(1/2), 67–101.

Coppedge, M., Alvarez, A., & Maldonado, C. (2008). Two persistent dimensions of democracy: Contestation and inclusiveness. *The Journal of Politics*, *70*(3), 632–647.

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S.-E., & Teorell, J. (2017a). *V-Dem comparisons and contrasts with other measurement projects* (Varieties of Democracy Working Paper Series, 2017(45)). Gothenburg: V-Dem Institute.

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S.-E., Teorell, J., Altman, D., . . . Wilson, S. (2017b). *V-Dem country-year dataset v7*. Gothenburg: V-Dem Institute.

Dahl, R. A. (1971). *Polyarchy: Participation and opposition*. New Haven, CT: Yale University Press.

Economist Intelligence Unit. (2014). *Democracy index 2014: Democracy and its discontents*. Retrieved from http://graphics.eiu.com/PDF/Democracy_Index_2010_web.pdf

Freedom House. (2016). *Freedom in the world 2016*. Retrieved from https://freedomhouse.org/report/freedom-world/freedom-world-2016

Giannone, D. (2010). Political and ideological aspects in the measurement of democracy: The freedom house case. *Democratization*, *17*(1), 68–97.

Giebler, H. (2012). Bringing methodology (back) in: Some remarks on contemporary democracy measurements. *European Political Science*, *11*(4), 509–518.

Lijphart, A. (2012). *Patterns of democracy: Government forms and performance in thirty-six countries* (2nd ed.). New Haven, CT: Yale University Press.

Lindberg, S. I., Coppedge, M., Gerring, J., & Teorell, J. (2014). V-Dem: A new way to measure democracy. *Journal of Democracy*, *25*(3), 159–169.

Lueders, H., & Lust, E. (2017). *Multiple measurements, elusive agreement, and unstable outcomes in the study of regime change* (Varieties of Democ-

racy Working Paper Series, 2017(52)). Gothenburg: V-Dem Institute.

Marshall, M. G., Gurr, T. R., & Jaggers, K. (2016). *Polity IV project: Political regime characteristics and transitions, 1800–2015: Dataset users' manual*. Vienna, VA: Center for Systemic Peace. Retrieved from http://www.systemicpeace.org/inscr/p4manualv2015.pdf

McMann, K. M., Pemstein, D., Seim, B., Teorell, J., & Lindberg, S. I. (2016). *Strategies of validation: Assessing the Varieties of Democracy corruption data* (Varieties of Democracy Working Paper Series, 2016(23)). Gothenburg: V-Dem Institute.

Merkel, W. (2004). Embedded and defective democracies. *Democratization*, *11*(5), 33–58.

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

O'Donnell, G. A. (1994). Delegative democracy. *Journal of Democracy*, *5*(1), 55–69.

Pemstein, D., Meserve, S. A., & Melton, J. (2010). Democratic compromise: A latent variable analysis of ten measures of regime type. *Political Analysis*, *18*(4), 426–449.

R Core Team. (2017). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Retrieved from https://www.R-project.org

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36.

Steiner, N. D. (2016). Comparing Freedom House democracy scores to alternative indices and testing for political bias: Are US allies rated as more democratic by Freedom House? *Journal of Comparative Policy Analysis: Research and Practice*, *18*(4), 329–349.

Teorell, J., Coppedge, M., Skaaning, S.-E., & Lindberg, S. I. (2016). *Measuring electoral democracy with V-Dem data: Introducing a new polyarchy index* (Varieties of Democracy Working Paper Series, 2016(25)). Gothenburg: V-Dem Institute.

Teorell, J., Dahlberg, S., Holmberg, S., Rothstein, B., Khomenko, A., & Svensson, R. (2017). *The quality of government standard dataset, version jan17*. University of Gothenburg: The Quality of Government Institute. Retrieved from http://www.qog.pol.gu.se doi:10.18157/QoGStdJan17

Treier, S., & Jackman, S. (2008). Democracy as a latent variable. *American Journal of Political Science*, *52*(1), 201–217.

About the Authors

Martin Elff is a professor of political sociology at Zeppelin University, Friedrichshafen. He holds a doctorate in social science (Dr. rer. soc.) from the University of Mannheim. His research interests are in political attitudes and political behaviour, party competition, and quantitative methods of political science.

Sebastian Ziaja is a postdoctoral researcher at the Research Center for Distributional Conflict and Globalization at Heidelberg University. He holds a PhD from the Department of Government at the University of Essex. His research focusses on political regimes, democracy aid, state fragility, and the measurement of social science concepts.

COGITATIO

Politics and Governance (ISSN: 2183–2463)
2018, Volume 6, Issue 1, Pages 105–116
DOI: 10.17645/pag.v6i1.1183

Article

Different Types of Data and the Validity of Democracy Measures

Svend-Erik Skaaning

Department of Political Science, Aarhus University, 8000 Aarhus, Denmark, E-Mail: skaaning@ps.au.dk

Submitted: 27 September 2017 | Accepted: 20 November 2017 | Published: 19 March 2018

Abstract
Different measures of democracy rely on different types of data. Some exclusively rely on observational data, others rely on judgement-based data in the form of in-house coded indicators or expert surveys. A third set of democracy measures combines information from indicators based on different types of data, some of them also data from representative surveys of the mass public. This article discusses the advantages and disadvantages of these different types of data for the measurement of electoral and liberal democracy. The discussion is based on the premise that the main priorities must be to establish a high degree of concept-measure consistency, i.e. indicators capture relevant aspects of the core concept of interest in a precise and unbiased manner, and to provide high coverage. The basic argument of the article is that no type of data is superior to others in all respects. The article draws on examples from extant datasets to illustrate the tradeoffs and it offers suggestions about how to reduce some of the potential drawbacks.

Keywords
democracy; measuring democracy; reliability; types of data; validity

Issue
This article is part of the issue "Why Choice Matters: Revisiting and Comparing Measures of Democracy", edited by Heiko Giebler (WZB Berlin Social Science Center, Germany), Saskia P. Ruth (German Institute of Global and Area Studies, Germany), and Dag Tanneberg (University of Potsdam, Germany).

© 2018 by the author; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Not everything that can be counted counts,
and not everything that counts can be counted.
(Cameron, 1963, p. 13)

1. Introduction

The construction and use of measures of democracy in social scientific research has increased considerably in recent decades. This makes good sense; without them, the identification of trends in political rights and liberties must be based on rough impressions not allowing for systematic temporal and cross-country comparisons (Bollen, 1992, p. 189). However, such efforts are only valuable if the quality of the data is high in terms of reliability and validity.[1]

When we attempt to measure democracy, the identification of empirical indicators that tap into the different aspects of the overarching concept is one of the most important tasks. One can either use extant indicators, collect new data, or combine new indicators with old ones. The main priority must be to establish a high degree of concept-measure consistency, i.e. the extent to which the indicators capture all of the components of the core concept of interest (and only those), and the extent to which they do so in a precise and unbiased manner (Adcock & Collier, 2001; Goertz, 2006; Munck, 2009).[2] In the

[1] *Reliability* concerns whether a measurement procedure produces similar results under consistent conditions. *Validity* concerns the extent to which a measure plausibly captures the concept it is supposed to measure. *Reliability* is a necessary but insufficient condition for measurement validity. See Seawright and Collier (2014) for an overview and critical assessment of different validation strategies applied to measures of democracy. The strategies they discuss mainly apply to extant measures, while they neither discuss the data generating procedures nor address the question of different data types in the same level of detail as the present article.

[2] Note that concept-measure consistency, besides the use of adequate indicators, also concerns the aggregation procedures used to combine the information provided by different indicators. However, the question of whether the aggregation of information provided by the indicators is based on theoretically justified, empirically sound procedures is not part of this article's agenda as it constitutes a rather independent issue (see Bollen & Lennox, 1991; Goertz, 2006; Møller & Skaaning, 2011, Appendix; Munck, 2009).

words of the Office of the High Commissioner for Human Rights (OHCHR, 2012, p. 50):

> An important statistical consideration in identifying and developing human rights indicators, or any set of indicators for that matter, is to ensure their relevance and effectiveness in measuring what they are supposed to measure. This relates to the notion of indicator validity. It refers to the truthfulness of information provided by the estimate or the value of an indicator in capturing the state or condition of an object, event, activity or an outcome for which it is an indicator. Most other statistical and methodological considerations follow from this requirement.

Among the supplementary—and related—criteria that scholars take into consideration are: Whether indicators are produced through transparent and replicable data-generating processes, whether they are made publicly available, and whether they have extensive coverage in terms of units (typically countries) and time (typically years). Researchers face numerous tradeoffs when trying to fulfill these criteria.

One of the most important considerations is what type of data the ever-growing industry of measuring democracy, governance, and human rights should rely on (see Arndt & Oman, 2006; Landman & Carvalho, 2009, Chapter 3; OHCHR, 2012; Schedler, 2012; United Nations Development Programme, 2012).

Different measures of democracy are based on different types of data. Four main data types have been used to construct the major democracy measures: observational data (OD), i.e. data on directly observable facts, such as turnout rates or the presence or absence of formal political institutions; 'in-house' coding (IC) by researchers and/or their assistants based on an assessment of country-specific information found in reports, academic works, newspapers, archival material, etc.; expert surveys (ES), where selected country experts provide an evaluation based on their case-specific knowledge; and representative surveys (RS), where a sample of ordinary citizens provide judgements about particular issues.[3]

All of these types of sources have different strengths and shortcomings. Even though this is well-known, contrasting views about what kind of data is better still exist. To illustrate, the Office of the United Nations High Commissioner for Human Rights (OHCHR, 2012), which represents the global commitment to universal ideals of human dignity, takes a clear stand in favor of observable data in its widely cited report on human rights measurement. This preference for fact-based quantitative indicators over judgement-based indicators[4] is motivated by an interest in making assessments less subjective and thus more broadly acceptable. According to Cheibub, Gandhi and Vreeland (2010, p. 77), the data required by judgement-based democracy measures 'are hard, if not impossible, to obtain. Consequently, we suspect that these measures entail coding created on the basis of inferences, extensions, and perhaps even guesses' (see also Merkel et al., 2016; Przeworski, Alvarez, Cheibub, & Limongi, 2000; Vanhanen, 2000).

In contrast, the people behind the Worldwide Governance Indicators state that fact-based indicators are insufficient for capturing the realities of governance outcomes on the ground (Kaufman & Kraay, 2008). They therefore consider judgement-based data as a valuable tool. This position is motivated by the assumption that it is virtually impossible to capture the relevant aspects of governance, including democracy, without relying on the judgement of experts, in-house coders, and/or citizens (see also Bowman, Lehoucq, & Mahoney, 2005; Coppedge, Gerring, Lindberg, Skaaning, & Teorell, 2017a, 2017b; Munck, 2009; Schedler, 2012).

To increase the awareness among producers and users of democracy data, it seems pertinent to critically review and supplement the arguments and suggestions in a single article. More particularly, this article discusses the pros and cons of different data types and suggests how to counter some of the potential problems related to the measurement of electoral democracy (i.e. access to government power is determined by competitive and inclusive elections) and liberal democracy (i.e. electoral democracy combined with respect for civil liberties and the rule of law) (see Møller & Skaaning, 2011). The discussion draws on extant as well as suggested indicators to illustrate the tradeoffs. After presenting an overview of what kind of data extant democracy measures are based on, I discuss—for each of the four types of data in turn—the potential advantages and disadvantages regarding reliability and validity together with suggestions to reduce some of the problems. The basic argument of the article is that no type of data is superior to the others in *all* respects. Researchers should generally pay more attention to different ways of increasing valid measurement, including the combination of different types of data and data from different sources, whenever they construct their measures. It is not reasonable simply to stick to conformist practices and dogmatic doctrines about the general superiority of one type of data.

2. Extant Democracy Measures: What Kinds of Data Are Used?

Table 1 makes clear that there is considerable variation regarding how many kinds and which kind of data

[3] In this article, I exclusively focus on different types of standards-based data and thus disregard different types of events-based data.

[4] The distinction between fact-based and judgement-based indicators 'refers to information content of the indicators in question. Accordingly, objects, facts or events that can, in principle, be directly observed or verified, such as formal political institutions, are categorized as objective [fact-based] indicators. Indicators based on perceptions, opinions, assessment or judgements expressed by individuals are categorized as subjective [judgement-based] indicators' (OHCHR, 2012, p. 17). However, regarding the measurement of democracy and other governance-related concepts, it is often difficult to make a clear-cut distinction.

sources they build on. This plethora of approaches indicates that it not obvious what kind of data—or mix of data—one should prefer when trying to measure democracy. For some indicators, it is not easy to say if they are fact-based or judgement-based (more on this below). But if we take the statements of the different data providers as given, the Democracy–Dictatorship dataset (Cheibub et al., 2010) and Vanhanen's (2000) polyarchy measure only use observational data. The first of these measures uses indicators of legislative and executive elections, status of the legislature, opposition parties, and government turnovers to create a dichotomous distinction between democracies and autocracies. The second only uses share of votes cast for the largest party and electoral turnout rates in national elections to capture the level of democracy.

The underlying data of the Bertelsmann Transformation Index (Bertelsmann Stiftung, 2017), the Freedom in the World survey (Freedom House, 2017), and the Perception of Electoral Integrity index (Norris, Frank, & Martínez i Coma, 2014) are all based on expert assessments. The Polity Measure (Marshall, Gurr, & Jaggers, 2016) and the CIRI Human Rights Dataset (Cingranelli & Richards, 2010) solely rely on in-house coded data. The remaining measures included in the overview presented in Table 1 build on more than one kind of data source. The Democracy Barometer dataset (Merkel et al., 2016), the Unified Democracy Scores (Pemstein, Meserve, & Melton, 2010), and the Worldwide Governance Indicators (Kaufman & Kray, 2017) do not provide original data collection but use extant indicators based on all four kinds of data sources. The Varieties of Democracy (Coppedge et al., 2017b) dataset relies on all types of data apart from representative surveys, and the Democracy Index by the Economist Intelligence Unit (2007) only excludes in-house coded data. Finally, the Lexical Index of Electoral Democracy (Skaaning, Gerring, & Bartusevičius, 2015) combines two kinds of data sources: in-house coded data and observational data. It varies quite a bit from measure to measure whether the

Table 1. Selected characteristics of 13 large-scale democracy datasets. Source: Coppedge et al. (2017a, p. 6) and own assessment.

Names of Data Provider and Dataset	**Years covered**	**Types of sources**				**Based on various datasets**	**Uncertainty estimates**
		IC	OD	ES	RS		
Bertelsmann Stiftung (2017): Bertelsmann Transformation Index (BTI)	2003–2015 (biennial)			X		No	No
Cheibub et al. (2010): Democracy–Dictatorship (DD)	1946–2008		X			No	No
Cingranelli & Richards (2010): CIRI Human Rights Database (CIRI)	1981–2011	X				No	No
Coppedge et al. (2017b): Varieties of Democracy dataset (V-Dem)	1900–2016	X	X	X		No	Yes
Economist Intelligence Unit: Democracy Index (EIU)	2006, 2008, 2010–2016		X	X	X	Yes	No
Freedom House: Freedom in the World (FH)	1972–2016			X		No	No
Kaufmann & Kray (2017): Worldwide Governance Indicators (WGI)	1996, 1998, 2000–2015	X	X	X	X	Yes	Yes
Marshall, Gurr, & Jaggers (2016): Polity IV (Polity)	1800–2015	X				No	No
Merkel et al. (2016): Democracy Barometer (DB)	1990–2014	X	X	X	X	Yes	No
Norris et al. (2014): Perceptions of Electoral Integrity (PEI)	2012–2016			X		No	Yes
Pemstein et al. (2010): Unified Democracy Scores (UDS)	1946–2012	X	X	X	X	Yes	Yes
Skaaning et al. (2015): Lexical Index of Electoral Democracy (LIED)	1800–2016	X	X			No	No
Vanhanen (2000): Polyarchy Dataset (Vanhanen)	1810–2014		X			No	Yes

different types of data are used to measure the same subcomponents and components of the overall democracy measures.

Disagreements about best practices regarding what kind of data to employ continue to flourish. The great variation not only reflects differences in resources; it also indicates different weighting of the potential problems related to data types. But what are the more specific pros and cons of different data types? How can the data type choice matter for reliability and validity?

3. Potential Advantages and Disadvantages of Different Kinds of Data

3.1. Observational Data

Observational data have a high, often preferred standing among users of social science data. In the words of Cheibub et al. (2010, p. 74), 'The reliability of a measure depends on whether knowledge of the rules and the relevant facts is sufficient to unambiguously lead different people to produce identical readings on specific cases'. On this basis, they prefer democracy measures based on directly observable and verifiable indicators rather than subjective and fuzzy indicators. Among the main assets of a fact-based approach to measurement are transparency and a replicable data-generation process, which is generally less susceptible to biases than judgement-based data types (see below). Moreover, observational data often provides scales of the phenomena in question that are both relatively easy to interpret and comparable across countries and over time (OHCHR, 2012).

However, the assumptions underlying this preference are criticized for being unrealistic. According to Schedler (2012, p. 28), the collection and use of non-judgmental data in the social sciences rests on two conditions: '(1) transparent empirical phenomena whose observation do not depend on our judgmental faculties and (2) complete public records on those phenomena'. When we want to measure democracy, none of these criteria are met. Not all aspects of democracy are easily observable and, relatedly, official statistics do not capture many relevant features in meaningful ways. Readily observable empirical information is often incomplete, inconsistent, or insufficient. 'Some empirical phenomena we cannot observe in principle, others we cannot observe in practice' (Schedler, 2012, p. 28).

A particular problem emerges when measuring democracy by examining the official (formal) laws of the land, first and foremost the constitution:[5] There is often a large discrepancy between what appears on the books and what is practiced on the ground. Informal rules and traditions are often more important than formal regulations. To illustrate this point with an extreme example, the Soviet 1936 constitution (aka. the Stalin Constitution) promised free and fair elections and respect for civil liberties on top social and economic rights. In practice, however, the political regime was totalitarian (see Linz, 2000), including a level of state repression that has hardly been matched by any other political regime in world history.

This problem refers to more than discrepancies between *de jure* and *de facto* regulations. For example, OHCHR (2012, p. 97) has suggested using reported cases of killing, disappearances, detention, and torture against journalists to measure freedom of opinion. This could be a relevant indicator but has two significant shortcomings. On the one hand, there is likely to be a reporting bias because reliable information is often not readily available (see Fariss, 2014; McNitt, 1988, pp. 94–99; Weidmann, 2016). The perpetrators normally have a clear interest in keeping the correct number secret and it is often difficult to know why a particular journalist has disappeared or died, or whether they were imprisoned due to a legitimate use of their freedom of expression or some other reasons. On the other hand, anticipated sanctions often lead to self-censorship. Journalists are rarely killed in North Korea (as far as we know), because they know that criticizing the government would have dire consequences. These problems would apply to similar attempts at capturing respect for liberal rights and adherence to the rule of law by the (exclusive) use of observational indicators.

Among the attempts to measure democracy using observational data, we find the democracy–dictatorship dataset (Cheibub et al., 2010; Przeworski et al., 2000). Its reliance on the rule of electoral government turnover to determine whether elections have been free suffers from two problems. First, the so-called Botswana problem, i.e. a government seems to be continuously reelected through free elections, meaning that Botswana (and other such cases) does not fulfill the turnover-criterion, saying that an alternation in government power has taken place under electoral rules identical to those bringing the incumbent into power. Second, the turnover criterion is implemented in a way that could introduce further problems. The coding rule says that a government turnover implies that a particular regime is coded as democratic all the way back to when the previous government took power given that the case also fulfilled the other criteria for democracy in the period, if the electoral rules are identical. However, a judgement call is sometimes needed to determine what counts as electoral rules and what counts as a relevant change to these rules (Knutsen & Wig, 2015, p. 909).[6]

The freeness and fairness of elections could also improve significantly (from no uncertainty to significant uncertainty about the outcome) under the same (formal) electoral rules. This applies, among other cases, to the

[5] This is a prominent feature of the Democracy Barometer and a number of governance measures, such as the Rule of Law Index by the World Justice Project (2016).

[6] This point applies more generally to seemingly fact-based indicators used to measure democracy, where elements of judgement cannot be fully excluded.

Dominican Republic between 1966 and 2002. In this period, election outcomes varied greatly and, according to comprehensive case studies, did not meet the minimum threshold for electoral democracy before 1978 (Marsteintredet, 2009, Chapter 4). In other cases, government turnover merely signifies that the ruling coalition is split and no longer controls the sufficient means to stay in power—a situation that the opposition exploits to gain power through manipulated elections. This problem applies to, for example, the change from conservative to liberal hegemonic rule in Columbia (1930–1931) and President Kurmanbek Bakiyev's rise to power in connection with the Tulip Revolution in Kyrgyzstan. Hence, government turnover often provides strong and relevant indication of free electoral competition, but it is not unproblematic and undisputable evidence (see Bogaards, 2007; Boix, Miller, & Rosato, 2013; Skaaning et al., 2015).

Another well-known example is Vanhanen's (2000) use of voter turnout and the vote share of the largest party in order to capture different degrees of democracy. These indicators tend to fail to tap all of the relevant aspects of democracy, however, such as the degree of freedom of expression and the power of the parliament, while capturing things that do not directly reflect the level of democracy, such as mandatory voting, dissatisfaction with the government, and the weather on voting day (Bollen, 1990, pp. 8, 15). Both the official statistics on turnout rates and vote share could also be unreliable—either because the data has been manipulated or because the data providers have been unable to collect all of the relevant information and aggregate them correctly. Some governments simply do not have the capacity to collect and handle the relevant information, which leads to missingness or flawed estimates. Other governments and/or their agents have strong incentives (and few constraints) to manipulate official data in order to misinform their own citizens and foreign governments and organizations (Herrera & Kapur, 2007).

Both of these circumstances can seriously reduce the availability and quality of data that could be relevant for measuring democracy, since they tend to be politically sensitive. Even in the case of so-called hard economic data (e.g. GDP per capita and trade), where governments and international organizations invest extensive resources in the collection of information and calculate the figures used by countless social scientists, there are remarkable problems regarding reliability and validity (Jerven, 2013; Kerner, 2014). There is therefore good reason to refrain from buying into the claim that fact-based data are always more informative and less biased than judgement-based data (see, e.g., Bollen, 1992; Coppedge et al., 2017a, 2017b; Kaufman & Kraay, 2008; Schedler, 2012). Public statistical information and other types of observable data can be useful for measuring democracy, but directly observable indicators do not capture all aspects of democracy well.

3.2. In-House Coded Data

One way of overcoming problems related to the lack of good observable indicators is to base scores on different kinds of relevant information found in diverse sources providing country-specific information, such as newspapers, election observation reports, human rights reports, and academic works. The construction of in-house coded data normally follows a particular procedure: Relevant information is gathered, after which a coder evaluates the evidence on one or more particular issues and translates the evaluation into a score based on more or less explicit and precise standards. Note, furthermore, that in the case of in-house coding, the coders are not experts on all of the (many) countries (and maybe also not the substantive areas) they assign scores to.

In-house coded data has three major advantages: It can be used to capture important traits that are largely undetectable by observational data (Bollen, 1993, p. 1210; Hadenius & Teorell, 2005, pp. 14–15; Mainwaring, Brinks, & Pérez-Liñán, 2001, p. 61; Munck & Verkuilen, 2002, p. 18). In many cases, bits and pieces of evidence can be put together to create a more general understanding of the actual respect for different democratic rights. On this basis, raters can make an informed estimate of the extent of, say, electoral contestation, freedom of expression, and fair trails, which would otherwise be very difficult to capture in a nuanced manner.

Another positive feature of in-house coded data is that the centralized assignment of scores by one or a few selected coders, ceteris paribus, generally makes for a higher degree of consistency when applying coding criteria. The understandings of concepts and scales will simply be more uniform compared to (more 'decentralized') expert surveys and public opinion surveys. In other words, in-house coding facilitates similar applications of standards across countries, especially if the number of coders is low and they are carefully trained and supervised. The use of multiple coders and inter-coder reliability tests are valuable tools to assess whether the assumptions about consensus among coders are met, i.e. there is consistency in the estimate if the data-generating procedure is repeated by the same or different coders (see Gwet, 2014).[7]

The third potential advantage is that in-house coding facilitates standardized and detailed documentation of why particular observations are assigned certain values. Detailed documentation of the motivation behind the particular scores can obviously be very time-consuming, which is probably why it is not provided in connection to any of the democracy measures based on in-house coding.

There are other reasons for hesitating before accepting values derived from in-house coding. The use of in-house coded data (and judgement-based data more generally) is sometimes rejected with reference to its sub-

[7] Such as those made public in connection to BTI, CIRI, and Polity for a single year and, more appropriately, for a random selection of 10% of the country-years in connection to LIED.

jective nature. In contrast to genuine subjective measures, however, such as data on public attitudes, 'they are not supposed to be subjective, but intersubjective: grounded in public facts and public reasons, defensible in the face of critique' (Schedler, 2012, p. 24). Despite this well-taken qualification, coder-specific biases can still influence the scores in different stages of the coding process (Bollen & Paxton, 1998, 2000):

> First, differential use of sources of information, combined with the filtering of information across the world, could lead to specific judge-centered method factors. Second, judges can process the information available to them in such a way as to differentially weight relevant events or to include irrelevant factors. Finally, the methods of constructing a measure might introduce method effects. (Bollen & Paxton, 2000, p. 64)

In-house coders do not have expert knowledge of all of the countries they code. They must therefore rely on secondary sources, which obviously differ with respect to availability and relevance. Systematic distortion of information is likely as it makes its way from the actual practices and events to the sources of information used by the coders. Accessible data can be ordered according to its informative value. The best situation would be for all relevant information to be available, but this is unrealistic. The following ordering of information therefore applies: recorded, accessible, locally reported, and internationally/foreign reported (Bollen, 1992, pp. 198–199). Movement from the former to the latter resembles a filtering process where some information passes through and some does not.

This process is likely to introduce biases. Filters often tend to be selective in non-random fashions, meaning that the information is neither complete nor representative (Foweraker & Krznaric, 2000, p. 766; Milner, Poe, & Leblang, 1999, p. 420). This is due to differences in the openness of countries, how much international attention they receive (influences by size, language, etc.), ideological preferences of the media, specific agendas of scholarly works and reports, and so forth. While most of the providers of original in-house coding (LIED, Polity, V-Dem) use multiple sources (which are generally unspecified), only CIRI makes use of the Country Reports on Human Rights Practices issued by the US State Department.[8] This fact means that the validity to a very high degree depends on the representativeness and impartiality of a single source, which has been accused of being biased—especially in the early releases (see Innes, 1992; Poe, Carey, & Vazquez, 2001; Qian & Yanagizawa, 2009).[9]

In the next step, raters can introduce random and systematic measurement errors by interpreting the sources differently, either because they based their evaluation of different pieces of relevant or irrelevant information, because they weight the same evidence differently, or because they have different understandings of concepts and scales guiding the coding process.[10] According to Raworth (2001, p. 114), 'The identity of the individuals giving the ratings is inevitably open to questioning'.

Differences in the specific coding processes can also influence the scores. Raters can assign scores to many or few countries (and different groups of countries); they can finalize scores immediately or go back and revise some of them; they can code everything between one year and hundreds of consecutive years at the time; and they can work on the coding in a relatively short but intensive period or carry out the task over a longer, less-intensive period. All of these factors will tend to influence the implicit reference points in the minds of coders and thus have an impact on the scores. The ability of in-house coded data to capture latent regime features in a consistent way is promising, while biases introduced in the coding process and the lack of comprehensive case knowledge are among the potential downsides of this kind of data.

3.3. Expert Survey Data

Expert survey data is generated through assessments of the fulfilment of democratic rights with the help of informed experts, often scholars or other persons working in related fields and intimately acquainted with the subject matter, such as journalists or leading members of NGOs. The main advantage of expert surveys compared to in-house coded data is exactly the case knowledge. The experts presumably know the relevant context and details about the issues in question (Marquardt et al., 2017). If their knowledge is insufficient, they have a superior background for finding relevant information. Experts may even have sufficient contextual knowledge to provide a plausible estimate if there is limited available evidence in terms of written sources directly tapping into a particular phenomenon. Original expert surveys are part of BTI, EIU, FH, PEI, and V-Dem; the three former only use one expert per country, while PEI and V-Dem use multiple experts per country (Coppedge et al., 2017a, p. 8). V-Dem even divides its survey into different categories, and to some degree enlists different experts to fill out different parts of the overall survey for each country (Coppedge et al., 2017b).

The potential problems identified in relation to in-house coded data also apply to expert surveys. The filtering of information might not be as big a problem due to

[8] To code physical integrity rights, CIRI also employs the Annual Reports from Amnesty International.
[9] For detailed discussions of potential biases in the Freedom House scores, see Bollen and Paxton (2000), Giannone (2010), and Steiner (2016).
[10] As stated by Bollen (1990, p. 18), 'A variety of personal factors could unconsciously affect a judge's ratings. These include the relation of the country being rated to the judge's home country, the political orientation of the judge, or any personal stakes in the rating'. Actually, one kind of personal stake, namely academic credibility, will tend to increase the quality of the data, while disinterest in the quality of the product (as is probably the case, at least in relative terms, for many research assistants) can produce low reliability.

the case expertise. However, the selection and weighting of evidence and the coding process will differ somewhat from expert to expert, partly depending on personal factors, such as updated and relevant familiarity with the cases, political leaning, job situation, and work effort (Bollen & Paxton, 1998). Expert knowledge varies and is sometimes inadequate, and the experts often lack strong incentives to enlist and spend much time doing a serious coding job, including searches for additional information. Furthermore, limited and differentiated knowledge leaves room for the so-called 'halo effect,' which is the tendency for a good (or bad) impression of performance in one area to influence opinion regarding other areas (Sequeira, 2012). These circumstances draw attention to the three-fold challenge related to the recruitment of experts. The experts should preferably be the most knowledgeable, unbiased, and be ready to do a careful job. However, the enrolled experts are rarely the best possible according to these criteria.

Experts are also more prone to apply different coding criteria than in-house coders because expert surveys are mostly carried out as decentralized coding without prior training, meaning that the basic understanding of concepts and scales can vary greatly (see Martinez i Coma & Van Ham, 2015; Steenbergen & Marks, 2007). BTI, EUI, and FH combine their expert assessments with review and deliberation across a team of in-house analysts. For good reason, this approach is assumed to increase cross-country consistency. The procedures are not transparent, however, since it is not made public which changes are introduced to the original expert-based values and why for any of the cases.

V-Dem has a different approach to increase the comparability and reduce the influence of potential biases. A complex Bayesian IRT measurement model uses information about agreement across coders, self-assigned uncertainty estimates by the experts about their own ratings, personal coder characteristics (extracted from a post-questionnaire survey), links between countries based on experts assessing more than one country (either for all years or one year), and vignettes related to the survey questions in order to align the experts' thresholds (see Pemstein et al., 2017). This procedure also supplements the scores with a systematic assessment of measurement uncertainty. This is also done for PEI but only based on the degree of expert agreement.

The documentation of the justifications for the scores is desirable, just as in the case of in-house coding. Even though it is usually impossible for experts 'to relate the numerical conclusions they reach to the precise pieces and bits of information that have gone into them…they should be able to document the big picture [and] describe the range of uncertainty and controversy regarding their judgmental decisions with reference to concrete documentary evidence (or the lack of such evidence)' Schedler (2012, p. 32). The extra workload for the experts and coordinators to provide and standardize the information makes this procedure very resource-demanding. Nonetheless, BTI and FH complement their scores with relatively detailed country reports, meaning that one can get an impression of what events and circumstances have influenced the scores for different aspects of democracy (but they do not provide adequate references to the material on which the reports are based).

In sum, the comparative advantage of expert surveys comes to the fore in situations of incomplete or inconsistent information, where contextual knowledge can be used to bridge informational gaps (Schedler, 2012, p. 28). However, the reliance on the personal judgements of a few experts means that the data might lack comparability and might be affected by different kinds of biases.

3.4. Representative Survey Data

The final type of data, representative surveys of the general population, brings the knowledge and opinions of ordinary citizens into play. Mayne and Geissel (2016) argue in favor of including a citizen component in the measurement of democratic quality. It should capture the citizens' democratic commitments, political capacities, and political participation. This perspective, however, seems more relevant for the measurement of deliberative and participatory democracy than electoral and liberal democracy. In connection to these more limited understandings of democracy, the suggested additions are better understood as possible causes or consequences of democracy. Pickel, Breustedt and Smolka (2016) also advocate for the inclusion of representative survey data in the measurement of democratic quality. They propose that citizen evaluations of democratic performance should complement other types of data.

For some purposes, representative surveys can provide valuable information. Respondents can function as 'everyday experts' on issues that are otherwise hard to get firm knowledge about. A case in point is petty corruption, where the experiences of citizens with having to pay bribes could be a superior source of information (see Naval, Walter, & de Miguel, 2008; Razafindrakoto & Roubaud, 2010). Another would be information about whether citizens experience or participate in political violence (see Bhavnani & Backer, 2007).

However, there are also noteworthy problems associated with the use of data from representative surveys to measure democracy. Most citizens lack nuanced knowledge about the general dynamics and performance of particular political institutions. Gut feelings and personal opinions are thus likely to influence the scores. Most drawbacks of judgement-based data apply more strongly to representative survey data than in-house coded data and data based on expert surveys (cf. Marquardt et al., 2017). Experts and in-house coders generally have better backgrounds for carrying out such assessments. They generally possess a broader knowledge regarding the political history of other countries and data collection procedures, a higher degree of shared understanding about

the meaning of particular concepts, and a strong scientific ethos (or least an interest in maintaining their academic credibility). This implies that individual biases and dissimilar standards (both within and across countries) in the interpretation of questions and scales are more pronounced. Ordinary citizens also tend to be more susceptible to collective cultural biases (nation-wide inclinations), and the respondents in representative surveys are very unlikely to provide any form of systematic reasoning for their entries. Ordinary citizens might also be afraid to share their experiences or express their honest opinion, especially in the case of an oppressive regime (Tannenberg, 2017).

Does this mean that we should generally refrain from using representative surveys to measure at all? Ordinary citizens might possibly possess valuable knowledge based on their real-life experiences that could supplement that of experts. Here, it seems pertinent to distinguish between experience-based questions and perception-based questions. The former ask citizens about their own experiences regarding particular situations (e.g. how often they have been asked to pay a bribe or been subjected to violent assaults in the previous year). The latter is typically based on more abstract questions, asking about the lay of the land regarding democracy, civil liberties, corruption, etc.

The experience-based questions have greater potential for providing relevant information than the second type, which are likely to produce unreliable and biased democracy indicators. Combined with the relatively low coverage in terms of years and countries,[11] it is therefore unadvisable to use perception-based data from representative surveys for democracy measurement. None of the evaluated measures are based on original data collection using this approach, but DB, EIU, UDS, and WGI rely on such data—either directly or indirectly (by including composite measures that use them).

4. Discussion

There are several ways of countering the disadvantages identified above. In relation to in-house coded and expert survey data, the documentation should ideally provide answers to the following questions: What evidence has been used and why? And how has the evidence been weighed and processed and why? That is, the criteria for identifying and selecting relevant sources and the criteria for extracting and using relevant information must be pinned down. This work can be done to different degrees of perfection to the point where every score is supplemented with nuanced description of the evidence (using active citation; see Moravcsik, 2014), how and why it has been weighted in certain ways (with relevance as the main criteria; see Bowman et al., 2005; Lustik, 1996; Møller & Skaaning, 2017), and who has been involved and how in the data collection and processing (Schedler, 2012, p. 33).

Inconsistency and personal biases can be reduced by the construction and application of specific and justified definitions of what one attempts to measure and the scales used to distinguish between different levels of fulfilment. The clarification should preferably be presented as precisely as possible and linked to concrete (maybe even paradigmatic) examples. This would support the establishment of shared anchors for the assignment of values. Another useful tool is to reduce conceptual complexity through disaggregation. This would imply the coding of more concrete issues than just freedom of expression, including media censorship, freedom of private discussion, harassment of journalists, and monopoly of news media.

Other factors, such as the exposure of coders to extensive relevant variation, can also improve the consistency. As a rule of thumb, they are more likely to employ similar standards across and within cases when the following conditions are fulfilled: The coders assign scores to a diverse set of many countries; they are willing and allowed to revise scores; they score long time series; and they score the cases within a relatively short period.

If in-house coders or experts score the same cases, formal measurement models can produce replicable point estimates and estimates of uncertainty.[12] One should note, however, that whereas it will almost always be good to increase the number of in-house coders (although there will be a diminishing return), more is not always better in the recruitment of expert coders because there will be a rather limited number of people with high levels of relevant expertise. Moreover, an increase in the number of coders will increase the costs attached to the data collection, thereby emphasizing the latent tradeoff between high quality data and coverage.[13]

Formal measurement models can also be used to combine data from different datasets based on different data-generating approaches (i.e. observational data, in-house coded data, expert survey data, and/or representative survey data) (Bollen & Paxton, 2000, p. 79). The advantage of such composite measures is their utilization of information from several variables. The combination of information from different data types can increase the ability to capture related, but distinct, aspects of the variable in question. In addition, it can reduce the impact of idiosyncratic measurement errors associated

[11] In most cases, it is overly demanding to request respondents to answer questions for several years, coding back in time. Moreover, for different reasons (e.g. regime type, geography, level of socio-economic development), it is extremely difficult to carry out high quality representative surveys in some countries.

[12] Besides the original scores, formal measurement models can utilize other types of information, such as data on the personal characteristics of the experts or in-house coders and their responses to vignettes linked to the variables (see Pemstein et al., 2017). It is also possible to use a measurement model approach to calculate point estimates and uncertainty in the case of representative surveys.

[13] This caveat about a higher demand for resources applies to several of the suggestions, including circumstances where data providers do not themselves possess the relevant skills for implementing them.

with individual indicators. The use of multiple indicators for the same phenomenon also facilitates an assessment of how precise the point estimates are through the construction of confidence levels (see Fariss, 2014; Pemstein et al., 2010). This integrative approach is used (in full or in part) to construct several of the democracy measures (see Table 1). By reducing some problems, however, it risks introducing or increasing others. The integration can lead to an accumulation of the problems associated with the individual indicators rather than resolving them. Moreover, the products tend to be more complex. This means that the relationship between measures and the concepts they should capture becomes more blurred.

Extant democracy measures build on different kinds of data; some only employ in-house coded data, expert survey data, or observable indicators, while others use different combinations of two or more of these types and representative survey data. The identification of the pros and cons typically associated with the respective data types has demonstrated that the different methodological choices about this issue matter for the reliability and validity of democracy measures. Table 2 summarizes some of the most important strengths and weaknesses typically associated with the different data types. Some of the similarities and dissimilarities follow the overall fact-based and judgement-based distinction, while others do not. The overview reveals that the pros and cons associated with the respective kinds of data are not simply mirror images of each other.

The discussion reveals the simplicity of the bullet points in Table 2. They neglect the many nuances of coding rules and processes that can influence the quality of the data. The comparative advantages and disadvantages of the different data types vary in both kind and degree. The reliability and validity depend on the particular procedures used in the data generating process and the aspects of democracy one attempts to capture.

The discussion has also revealed that no type of data is superior to all of the others in all respects when it comes to measuring the fulfilment of democratic rights. Hence, the arguments presented have challenged what many consider conventional wisdom, namely the general superiority on one kind of data—directly observable (fact-based) data. Actually, this belief tends to be a dogmatic doctrine resting on invalid assumptions. As neatly summarized by Schedler (2012, p. 21), 'Banning judgment from measurement is neither a feasible methodological imperative nor a desirable one', and:

> If we were to renounce our judgmental faculties in the measurement of regime properties and regime dynamics, we would have to renounce the measurement of most of the most interesting regime properties and regime dynamics. If we truly had expelled judgment from data development, quantitative research on political regimes could not have blossomed as it has over the past decades. (Schedler, 2012, p. 33)

This point applies more to the measurement of thicker understandings of democracy, such as liberal democracy. Respect for civil liberties and adherence to the rule of law tend to be even harder to capture without judgement-based indicators than narrow electoral criteria (regular, inclusive, and competitive elections). Considerable effort has already been invested in improving democracy measures. More can still be done to increase the reliability and validity, however, and greater awareness about these issues among data users is required.

Acknowledgments

I would like to thank the editors, the reviewers, Carl Henrik Knutsen, Jørgen Møller, and participants in the NOPSA workshop on the Progress and Decline in Democracy Outside the West for valuable comments.

Conflict of Interests

The author declares no conflict of interests.

Table 2. General advantages and disadvantages associated with different data types.

	Advantages	Disadvantages
Observational data	• Avoid personal biases • Fixed and comparable scales • Transparent documentation of scores	• Relevant information often not directly observable • Biases and limitations in available information
In-house coded data	• Consistency in the application of coding criteria • Capture latent traits	• Personal biases • Biases and limitations in the available information • Limited, case-specific knowledge
Expert survey data	• Case-specific knowledge • Capture latent traits	• Personal biases • Inconsistently applied coding criteria
Representative survey data	• Experience-based knowledge • Capture latent traits	• Personal biases • Unfeasible in particular settings • Inconsistently applied coding criteria

References

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

Arndt, C., & Oman, C. (2006). *Uses and abuses of governance indicators*. Paris: OECD Development Centre.

Bertelsmann Stiftung. (2017). *Bertelsmann transformation index*. Retrieved from http://www.bti-project.org/en/home

Bhavnani, R., & Backer, D. (2007). *Social capital and political violence in Sub-Saharan Africa* (Afrobarometer Working Paper 90). Michigan, MI: Michigan State University.

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

Boix, C., Miller, M., & Rosato, S. (2013). A complete dataset of political regimes, 1800–2007. *Comparative Political Studies*, *46*(12), 1523–1554.

Bollen, K. (1990). Political democracy: Conceptual and measurement traps. *Studies in Comparative International Development*, *25*(1), 7–24.

Bollen, K. (1992). Political rights and political liberties. In T. Jabine & R. Claude (Eds.), *Human rights and statistics: Getting the record straight* (pp. 188–213). Philadelphia, PA: University of Pennsylvania Press.

Bollen, K. (1993). Liberal democracy: Validity and method factors in cross-national measures. *American Journal of Political Science*, *37*(4), 1207–1230.

Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, *100*(2), 305–314.

Bollen, K., & Paxton, P. (1998). Detection and determinants of bias in subjective measures. *American Sociological Review*, *63*(3), 465–478.

Bollen, K., & Paxton, P. (2000). Subjective measures of political democracy. *Comparative Political Studies*, *33*(1), 58–86.

Bowman, K., Lehoucq, F., & Mahoney, J. (2005). Measuring political democracy: Case expertise, data adequacy, and Central America. *Comparative Political Studies*, *38*(8), 939–970.

Cameron, W. (1963). *Informal sociology: A casual introduction to sociological thinking*. New York, NY: Random House.

Cheibub, J. A., Gandhi, J., & Vreeland, J. R. (2010). Democracy and dictatorship revisited. *Public Choice*, *143*(1/2), 67–101.

Cingranelli, D., & Richards, D. (2010). The Cingranelli and Richards (CIRI) human rights data project. *Human Rights Quarterly*, *32*(2), 401–424.

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S.-E., & Teorell, J. (2017a). *V-dem comparisons and contrasts with other measurement projects* (The Varieties of Democracy Institute Working Paper Series 2017: 45). Gothenburg: University of Gothenburg.

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S.-E., Teorell, J., Krusell, J., . . . Wilson, S. (2017b). *V-dem methodology v7.1*. Gothenburg: Varieties of Democracy (V-Dem) Project.

Economist Intelligence Unit. (2007). *The Economist intelligence unit's index of democracy*. Retrieved from https://www.economist.com/media/pdf/DEMOCRACY_INDEX_2007_v3.pdf

Fariss, C. (2014). Respect for human rights has improved over time: Modeling the changing standard of accountability. *American Political Science Review*, *108*(2), 297–318.

Foweraker, J., & Krznaric, R. (2000). Measuring liberal democratic performance: An empirical and conceptual critique. *Political Studies*, *48*(4), 759–787.

Freedom House. (2017). About freedom in the world: An annual study of political rights and civil liberties. *Freedom House*. Retrieved from https://freedomhouse.org/report-types/freedom-world

Giannone, D. (2010). Political and ideological aspects in the measurement of democracy: The Freedom House case. *Democratization*, *17*(1), 68–97.

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

Gwet, K. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among multiple raters*. Gaithersburg: Advanced Analytics.

Hadenius, A., & Teorell, J. (2005). *Assessing alternative indices of democracy* (Committee on Concepts and Methods Working Paper Series, WP 6). Mexico City: The Committee on Concepts and Methods.

Herrera, Y., & Kapur, D. (2007). Improving data quality: Actors, incentives, and capabilities. *Political Analysis*, *15*(4), 365–86.

Innes, J. (1992). Human rights reporting as a policy tool: An examination of the State Department country reports. In T. Jabine & R. Claude (Eds.), *Human rights and statistics: Getting the record straight* (pp. 235–257). Philadelphia, PA: University of Pennsylvania Press.

Jerven, M. (2013). *Poor numbers: How we are misled by African development statistics and what to do about it*. Ithaca, NY: Cornell University Press.

Kaufman, D., & Kray, A. (2017). *Worldwide governance indicators*. Retrieved from http://info.worldbank.org/governance/wgi/index.aspx#home

Kaufmann, D., & Kraay, A. (2008). Governance indicators: Where are we and where should we be going? *World Bank Research Observer*, *23*(1), 31–36.

Kerner, A. (2014). What we talk about when we talk about foreign direct investment. *International Studies Quarterly*, *58*(4), 805–881.

Knutsen, C. H., & Wig, T. (2015). Government turnover and the effects of regime type: How requiring alternation in power biases against the estimated economic benefits of democracy. *Comparative Political Studies*, *48*(7), 882–914.

Landman, T., & Carvalho, E. (2009). *Measuring human rights*. London: Routledge.
Linz, J. (2000). *Totalitarian and authoritarian regimes*. Boulder, CO: Lynne Rienner Publishers.
Lustik, I. (1996). History, historiography, and political science: Multiple historical records and the problem of selection bias. *American Political Science Review*, *90*(3), 605–618.
Mainwaring, S., Brinks, D., & Pérez-Liñán, A. (2001). Classifying political regimes in Latin America, 1945–1999. *Studies in Comparative International Development*, *36*(1), 37–65.
Marquardt, K., Pemstein, D., Petrarca, C. S., Seim, B., Wilson, S. L., Bernhard, M., . . . Lindberg, S. I. (2017). *Experts, coders, and crowds: An analysis of substitutability* (The Varieties of Democracy Institute Working Paper Series 2017: 53). Gothenburg: University of Gothenburg.
Marshall, M. G., Gurr, T., & Jaggers, K. (2016). *Polity IV Project: Political regime characteristics and transitions, 1800–2016*. Vienna, VA: Center for Systemic Peace. Retrieved from http://www.systemicpeace.org/inscr/p4manualv2016.pdf
Marsteintredet, L. (2009). *Political institutions and democracy in the Dominican Republic: A comparative case-study*. Saarbrücken: VDM Verlag.
Martinez i Coma, F., & Van Ham, C. (2015). Can experts judge elections? Testing the validity of expert judgments for measuring election integrity. *European Journal of Political Research*, *54*(2), 305–325.
Mayne, Q., & Geissel, B (2016). Putting the demos back into the concept of democratic quality. *International Political Science Review*, *37*(5), 634–644.
McNitt, A. (1988). Some thoughts on the systematic measurement of the abuse of human rights. In D. Cingranelli (Ed.), *Human rights: Theory and measurement* (pp. 89–103). Basingstoke: Macmillan.
Merkel, W., & Bochsler, D. Bousbah, K., Bühlmann, M., Giebler, H., Hänni, M., . . . Wessels, B. (2016). *Democracy barometer: Methodology. Version 5*. Retrieved from http://www.democracybarometer.org/Data/Methodological_Explanatory_1990-2014.pdf
Milner, W. T., Poe, S. C., & Leblang, D. (1999). Security rights, subsistence rights, and liberties: A theoretical survey of the empirical landscape. *Human Rights Quarterly*, *21*(2), 403–443.
Møller, J., & Skaaning, S.-E. (2011). *Requisites of democracy*. London: Routledge.
Møller, J., & Skaaning, S.-E. (2017). *Reducing bias when enlisting the work of historians: A criterial framework*. Manuscript in preparation.
Moravcsik, A. (2014). Transparency: The revolution in qualitative research. *PS: Political Science & Politics*, *47*(1), 48–53.
Munck, G. (2009). *Measuring democracy: A bridge between scholarship and politics*. Baltimore, MD: John Hopkins University Press.
Munck, G., & Verkuilen, J. (2002). Conceptualizing and measuring democracy: Alternative indices. *Comparative Political Studies*, *35*(1), 5–34.
Naval, C., Walter, S., & de Miguel, R. S. (Eds.). (2008). Measuring human rights and democratic governance: Experiences and lessons from metagora [Special issue]. *OECD Journal on Development*, *9*(2).
Norris, P., Frank, R., & Martínez i Coma, F. (2014). Measuring electoral integrity around the world: A new dataset. *PS: Political Science & Politics*, *47*(4), 1–10.
Office of the High Commissioner for Human Rights. (2012). *Human rights indicators: A guide to measurement and implementation*. Geneva: OHCHR.
Pemstein, D., Meserve, S., & Melton, J. (2010). Democratic compromise: A latent variable analysis of ten measures of regime type. *Political Analysis*, *18*(4), 426–449.
Pemstein, D., Marquardt, K. L., Tzelgov, E., Wang, Y., Krusell, J., & Miriet, F. (2017). *The varieties of democracy measurement model: Latent variable analysis for cross-national and cross-temporal expert-coded data* (The Varieties of Democracy Institute Working Paper Series 2017: 21). Gothenburg: University of Gothenburg.
Pickel, S., Breustedt, W., & Smolka, T. (2016). Measuring the quality of democracy: Why include the citizens' perspective? *International Political Science Review*, *37*(5), 645–655.
Poe, S., Carey, S., & Vazquez, T. (2001). How are these pictures different? A quantitative comparison of the US State Department and Amnesty International human rights reports, 1976–1995. *Human Rights Quarterly*, *23*(3), 650–677.
Przeworski, A., Alvarez, M. E., Cheibub, J. A., & Limongi, F. (2000). *Democracy and development*. New York, NY: Cambridge University Press.
Qian, N., & Yanagizawa, D. (2009). The strategic determinants of US human rights reporting: Evidence from the Cold War. *Journal of the European Economic Association*, *7*(2/3), 446–457.
Raworth, K. (2001). Measuring human rights. *Ethics and International Affairs*, *15*(1), 111–131.
Razafindrakoto, M., & Roubaud, F. (2010). Are international databases on corruption reliable? A comparison of expert opinion surveys and household surveys in sub-Saharan Africa. *World Development*, *38*(8), 1057–1069.
Schedler, A. (2012). Judgment and measurement in political science. *Perspectives on Politics*, *10*(1), 21–36.
Seawright, J., & Collier, D. (2014). Rival strategies of validation: Tools for evaluating measures of democracy. *Comparative Political Studies*, *47*(1), 111–138.
Sequeira, S. (2012). Advances in measuring corruption in the field. *Research in Experimental Economics*, *15*(1), 145–175.
Skaaning, S.-E., Gerring, J., & Bartusevičius, H. (2015). A lexical index of electoral democracy. *Comparative Political Studies*, *48*(12), 1491–1525.
Steenbergen, M., & Marks, G. (2007). Evaluating expert

judgments. *European Journal of Political Research*, *46*(3), 347–366.

Steiner, N. (2016). Comparing Freedom House democracy scores to alternative indices and testing for political bias: Are US allies rated as more democratic by Freedom House? *Journal of Comparative Policy Analysis: Research and Practice*, *18*(4), 329–349.

Tannenberg, M. (2017). *The autocratic trust bias: Politically sensitive survey items and self-censorship* (The Varieties of Democracy Institute Working Paper Series 2017: 49). Gothenburg: University of Gothenburg.

United Nations Development Programme. (2012). *Governance indicators: A user's guide*. Oslo: UNDP.

Vanhanen, T. (2000). A new dataset for measuring democracy, 1810–1998. *Journal of Peace Research*, *37*(2), 251–265.

Weidmann, N. (2016). A closer look at reporting bias in conflict event data. *American Journal of Political Science*, *60*(1), 206–218.

World Justice Project. (2016). *WJP rule of law index 2016*. Retrieved from https://worldjusticeproject.org/our-work/wjp-rule-law-index/wjp-rule-law-index-2016

About the Author

Svend-Erik Skaaning is professor of political science at Aarhus University and co-principal investigator of Varieties of Democracy (V-Dem). His research interests include comparative methodology and the conceptualization, measurement, and explanation of democracy and human rights. He has published numerous articles and books on these issues, including *Requisites of Democracy, Democracy and Democratization in Comparative Perspective*, and *The Rule of Law*.

COGITATIO

Politics and Governance (ISSN: 2183–2463)
2018, Volume 6, Issue 1, Pages 117–144
DOI: 10.17645/pag.v6i1.1187

Article

Does the Conceptualization and Measurement of Democracy Quality Matter in Comparative Climate Policy Research?

Romy Escher * and Melanie Walter-Rogg

Department for Political Science Research Methods, Institute of Political Science, University of Regensburg, 93040 Regensburg, Germany; E-Mails: romy.escher@ur.de (R.E.), melanie.walter-rogg@ur.de (M.W.-R.)

* Corresponding author

Submitted: 29 September 2017 | Accepted: 7 December 2017 | Published: 19 March 2018

Abstract
Previous empirical research on democracy and global warming has mainly questioned whether democracy contributes to climate protection. However, there is no consensus in the *theoretical* literature on what institutional traits of democracy are crucial for climate policy. Thus, results based on indices that summarize multiple democracy quality dimensions could be misleading, as their effects could balance each other out or hide the relative importance of each institutional trait. This article examines whether the analysis of the effects of democracy quality dimensions, measured by separate indicators, contributes to a better understanding of cross-national variance in climate policy compared to the focus on the regime type difference, measured by democracy quality measures. Compared to earlier research, the results indicate that the positive effect of democracy on commitment to climate cooperation depends on the realization of political rights. We find little to support the claim that democracy quality dimensions matter for climate policy outcomes. The main implication of our findings is that it could be fruitful to use more disaggregated democracy measures for the analysis of substantive research questions.

Keywords
climate change policy; democracy; democracy quality; environmental policy; measures of democracy

Issue
This article is part of the issue "Why Choice Matters: Revisiting and Comparing Measures of Democracy", edited by Heiko Giebler (WZB Berlin Social Science Center, Germany), Saskia P. Ruth (German Institute of Global and Area Studies, Germany), and Dag Tanneberg (University of Potsdam, Germany).

© 2018 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

This article examines whether the analysis of the effects of specific dimensions of democracy quality, as opposed to the focus on the regime type difference, improves our understanding of cross-national variation in commitment to climate cooperation and climate change mitigation performance. Referring to the so-called Churchill hypothesis, which regards democracy as the best form of government, political scientists study the policy performance of democracies and autocracies. In answer to the political recognition of global environmental change, a growing number of studies have focused on the relationship between the regime type and climate change. Empirical research finds that democracies are more likely to join international environmental agreements (e.g., Bernauer, Kalbhenn, Koubi, & Spilker, 2010; Neumayer, 2002a) and perform better in solving local and regional environmental problems that do not demand considerable behavioral changes (e.g., Bernauer & Koubi, 2009; Li & Reuveny, 2006; Ward, 2008; Wurster, 2013) than their autocratic counterparts (see also Fiorino, 2011, pp. 375ff.). More recently, this conclusion has been questioned regarding global warming (e.g., Beeson, 2010; Shearman & Smith, 2007). While quantitative research supports the relationship between democracy and the

ratification of international climate agreements (e.g., Bättig & Bernauer, 2009; Fredriksson & Gaston, 2000; Gallagher & Thacker, 2008; Neumayer, 2002a, 2002b; von Stein, 2008), the empirical literature is unclear about the relevance of regime type for climate policy outcomes (e.g., Bättig & Bernauer, 2009; Fredriksson & Neumayer, 2013, 2016; Gleditsch & Sverdrup, 2002; Kneuer, 2012; Li & Reuveny, 2006; Midlarsky, 1998; Spilker, 2012, 2013; Wurster, 2013). Wurster (2013, p. 89) argues that "it has become clear that a dichotomous distinction between democracy and autocracy is not sufficient to explain the performance results" (see also Christoff & Eckersley, 2011, p. 439).

Previous empirical research has mainly questioned whether democracies contribute to climate protection. However, there is no consensus in the theoretical literature on how democracy influences climate policy. Different aspects of democracy are emphasized as crucial for the environment and there is no agreement on a uniform effect of the democracy quality dimensions (Burnell, 2012; Held & Hervey, 2011; Payne, 1995). Therefore, the question is which institutional traits of democracy affect the global atmosphere (Burnell, 2012, p. 823). This issue becomes important for the measurement of democracy quality in the statistical analysis as well. In accordance with their research question, empirical studies test the effect of summary measures of democracy (e.g., from Freedom House and Polity IV) on climate policy. While several studies test the robustness of their results using multiple indicators, democracy indices vary, not only in their validity and reliability but also in their underlying democracy concept (e.g., Munck & Verkuilen, 2002; Pickel, Stark, & Breustedt, 2015). Moreover, results based on indices that summarize multiple democracy quality dimensions could be misleading, as their effects could balance each other out or hide the relative importance of each trait.

Our analysis contributes to the academic literature on the measurement of democracy quality. More recently, disaggregated democracy quality data for large-N studies has been made available by the Democracy Barometer and the Varieties of Democracy Project (V-Dem). These indicators enable us to analyze the specific mechanisms that link democracy quality to policy outputs and outcomes. Thus, we will be able to go beyond the question of whether democracy contributes or undermines policy performance, and we will be able to state which institutional traits of democracy are responsible for a better or worse policy performance. Second, we contribute to comparative climate policy research. Global warming is a worldwide environmental problem. Hence, it is fruitful to analyze the willingness and ability of governments to tackle climate change in different institutional contexts. Only a limited amount of work has been done regarding the effect of specific democracy quality dimensions on climate policy (Böhmelt, Böker, & Ward, 2016, p. 1273). Fredriksson and Neumayer (2013, p. 18), in separate regression models, find that only the historical experience of executive constraints and not the cumulative effect of competition leads to stricter climate policies. On the contrary, Wurster (2013, pp. 86f.) ascertains that there is no significant effect of checks and balances on carbon dioxide (CO_2) emissions. To our knowledge, the effects of multiple democracy quality dimensions on climate policy have not been studied simultaneously.

Our main argument is that, if democracy quality has an impact at all, it is political rights that contribute to climate protection. Since solving global warming implies considerable changes in our daily lives and economy, it depends foremost on a demand for such measures by the citizenry. Political rights together with an independent civil society enable citizens to inform themselves about global environmental change, and they enable supporters of climate change mitigation to pressure the government via the media and public opinion to address global warming. In comparison, there is little reason to believe that competitive elections alone make governments implement climate policies. The diffuse character of the climate change problem makes it unlikely that emission reductions are relevant for most citizens' decisions in elections or their organization and participation in political parties. In addition, democratic governments are presumably reluctant to adopt stricter climate policies due to the considerable short-term socio-economic costs which could affect their re-election. While checks and balances imply that more interests are considered, veto players with divergent interests are likely to hinder the adoption of stricter climate policies. Civil rights enable individuals to focus on their self-interest even if it is against the common interest of environmental protection. However, they might also contribute, via the rule of law, to the acceptance of international agreements and the implementation of climate policies.

To answer our research question, we test the effects of four democracy quality dimensions—electoral and horizontal accountability, political and civil rights—on climate policy commitment and performance using data from the V-Dem-Project (Coppedge et al., 2017; Pemstein et al., 2017). We agree with Bättig and Bernauer (2009) that it is important to distinguish between commitment and performance. Compared to climate change mitigation, participation in climate cooperation comes with little cost. With the democracy measures from Polity IV and Freedom House (Freedom House, 2017; Marshall, Gurr, & Jaggers, 2016), we show the differences of our analytic strategy compared to former publications on the relationship between democracy and global warming. The focus on the effect of democracy quality limits the analysis to countries that have been classified as "free" or "partly-free" by Freedom House in the majority of years of our research period. Our findings shed new light on the causal mechanisms that link democracy to global warming. While results based on the Freedom House and Polity IV measures indicate that democracy quality, in general, contributes to a commitment

to international climate cooperation, the disaggregated measurement approach shows that the positive effect of democracy on commitment to climate cooperation depends on the realization of political rights. We find limited support for the claim that the other democracy quality dimensions matter for climate policy outcomes.

In the next section (2), we summarize the theoretical discussion regarding our research question and formulate hypotheses for the empirical analysis. The dependent and independent variables are operationalized in section 3. The following section (4) conducts cross-sectional analyses to explore our research question. The final section summarizes the findings and presents our conclusions (5).

2. Institutional Traits of Democracy and Climate Policy

In our discussion of the literature on the environmental consequences of democracy quality dimensions, we adopt the "embedded democracy" concept from Merkel (2004). Following Dahl, narrow definitions of democracy focus on competitive elections and enable the analysis of the relative performance of democracies and autocracies (e.g., Wurster, 2013, p. 82). They are less suited for the analysis of the relative importance of institutional traits of democracies for climate policy. Scholars disagree about what constitutes democracy besides competitive elections (Geissel, Kneuer, & Lauth, 2016, p. 574). Merkel's (2004) distinction between internal and external embeddedness enables us to focus on the effects of political democracy and not on the hypotheses that link democracy to climate policy via economic development or international cooperation (Burnell, 2012; Payne, 1995). The "embedded democracy" concept distinguishes five partial regimes, namely electoral regime, political rights, civil rights, horizontal accountability, and the effective power to govern. In comparison to democracy concepts that focus on political rights and civil liberties (e.g., Collier & Levitsky, 1997, p. 434; Freedom House, 2017), it identifies with the former four partial regimes the democracy-quality dimensions that are regarded as the most important in the democracy-environment literature. Our robustness analyses also control for three indicators regarding the effective power to govern (see Annex). With regard to the number of democracy quality dimensions the embedded democracy concept is relatively parsimonious (e.g. Diamond & Morlino, 2005, p. xii).

Several scholars expect that electoral accountability, i.e. the right to participate in the free and fair election of political authorities (Merkel, 2004, p. 42), contributes to environmental commitment and protection. First, democratically elected governments are responsive to their citizens' policy preferences (Barrett & Graddy, 2000; Congleton, 1992). The median voter should be more willing to accept stricter environmental regulations since they imply lower costs for citizens, compared to political and economic elites that possess a larger part of the national income (Bernauer & Koubi, 2009, p. 1356; Congleton, 1992, pp. 416f., 421; Winslow, 2007, p. 772). Additionally, non-elected governments might not adopt long-term environmental policies since their power is uncertain (Congleton, 1992, p. 417). Second, democracies presumably provide more environmental public goods to stay in power (Bueno De Mesquita, Smith, Siverson, & Morrow, 2003) since the price of public good provision relative to private goods falls with the size of the winning coalition (Cao & Ward, 2015, p. 265). Finally, Fredriksson, Neumayer, Damania and Gates (2005, p. 350), List and Sturm (2006, p. 1259) and Wilson and Damania (2005) emphasize the importance of competitive elections. Governments would only consider citizens' policy preferences if their participation made a difference.

The underlying assumption that citizens are climate-friendly is questionable (Spilker, 2013; Ward, 2008, p. 389). Empirical research finds that climate concern varies among countries (e.g., Kim & Wolinsky-Nahmias, 2014). Moreover, democratic governments are accountable to citizens within the nation-state and therefore might not be willing to deal with global environmental pollution (Held & Hervey, 2011, p. 90). Additionally, global warming mainly affects future generations and climate policies will only impact emissions in the long-term (Bernauer & Koubi, 2009, p. 1357; Cao & Ward, 2015, p. 271; Wurster, 2011, pp. 546f., 2013, p. 90). The diffuse character of the climate problem makes it unlikely that emissions are relevant for most citizens' decisions in elections. Democratic governments might also face citizens who are unwilling to accept the socio-economic costs of climate change mitigation (Holden, 2002, p. 10) and, therefore, prioritize economic development (Shearman & Smith, 2007, pp. xivf., 83). Non-elected governments also might not pursue climate-friendly policies; their legitimacy rests on their socio-economic performance. In sum, we expect that competitive elections cannot explain cross-national variation in commitment to climate cooperation (Hypothesis 1a) and climate change mitigation (Hypothesis 1b). It depends on whether supporters or opponents of climate protection are elected (see also Wurster, Auber, Metzler, & Rohm, 2015, pp. 183f.).

In democracies, horizontal accountability (i.e. checks and balances) makes it more likely that alternative policy choices are discussed and that the public is informed about environmental policies and their implementation (Burnell, 2012, p. 823; Held & Hervey, 2011; Wurster, 2013, p. 83). In comparison, the environmentalist authoritarian literature emphasizes that the democratic decision-making process hinders fast action to tackle climate change. Democratic governments must find an agreement with veto players with divergent economic interests (Beeson, 2010, p. 289; Fliegauf & Sanga, 2010, p. 2; Giley, 2012, p. 289; Wurster, 2011, p. 547, 2013, p. 79). Empirical research finds no clear support that institutional constraints contribute to or impede climate protection (e.g., Fredriksson & Neumayer, 2013; Garman, 2014; Madden, 2014; Wurster, 2013). Overall, we believe

that the positive and negative effects of horizontal accountability balance each other out and assume no effect of horizontal accountability on cross-national variation on climate policy commitment and performance (Hypotheses 2a and 2b).

In the 1970s, green political theorists argued that civil rights, i.e. constitutional rights that protect the individual against the state (Merkel, 2004, p. 39), contribute to individuals following their self-interest versus the common interest of environmental protection (e.g., Hardin, 1968; Ophuls, 1977, pp. 145ff.). However, this argument depends on the climate policy preferences of citizens. In addition, the effectiveness of using repression to enforce environmental policies is limited (Stehr, 2015, p. 450; Wurster, 2013, p. 80). Civil rights enable citizens to demand the implementation of climate policies via the courts (Spilker, 2013, pp. 55, 59; Winslow, 2007, p. 772). However, this possibility depends on existing environmental regulations. In sum, we expect that civil rights do not explain country-differences in climate policy commitment and performance (Hypotheses 3a and 3b).

Many social scientists link democracy to environmental protection via political rights, i.e. freedoms of expression, association, and the media, and the autonomy of the civil society (Merkel, 2004, p. 39). These institutional traits enable citizens to inform themselves regarding pollution (Barrett & Graddy, 2000; Bernauer, Böhmelt, & Koubi, 2013; p. 93f.; Payne, 1995, p. 43), to express their environmental policy-preferences (Bernauer et al., 2013, p. 93), to form environmental interest groups (ENGOs), to mobilize public support (Fredriksson & Neumayer, 2013, p. 12; Gleditsch & Sverdrup, 2002, p. 48), and to influence the government's decisions (Burnell, 2009, p. 6; Payne, 1995, p. 43). An independent civil society makes it more likely that citizens express their policy-preferences (Böhmelt et al., 2016, p. 1277). Free media enables citizens, journalists, and scientists to monitor government policy (Payne, 1995, p. 45) and support technological innovation as well as the spread of scientific knowledge (Gleditsch & Sverdrup, 2002, p. 47). However, powerful special interest groups may block environmental policy reforms (e.g., Bernauer & Koubi, 2009, p. 1357; Never & Betz, 2014, p. 12; Shearman & Smith, 2007, pp. 89, 91) or undermine their implementation (Midlarsky, 1998, p. 344). We expect that countries with higher levels of political rights are more committed to climate cooperation (Hypothesis 4a). Since the reduction of greenhouse gas emissions is associated with considerable short-term costs for society and economy, climate change mitigation is dependent on public awareness of global warming and support for climate protection. Democratic freedoms ensure that diffuse interests such as climate protection are at least *considered* as part of the political process. It is presumably more difficult for the public to control climate change mitigation compared to the ratification of climate treaties (Cao & Prakash, 2012). Thus, we expect no effect of political rights on climate policy outcomes (Hypotheses 4b).

To conclude, if democracy quality has an impact at all, it is political rights that contribute to commitment to climate cooperation. A uniform effect of the democracy quality dimensions on climate policy outputs and outcomes cannot be expected. Hence, they should be tested separately.

3. Research Design

To answer our research question, we examine the relevance of democratic institutional traits using cross-sectional OLS regression based on country averages from 1990–2005/2010. This is because, firstly, cross-country variations are of primary interest in this analysis. Secondly, institutional traits of democracy are relatively stable over time. Finally, we assume that political institutions affect climate policy only in the long-term. The selected period is relevant as climate change has only been recognized at the end of the last century on the international level as a global environmental problem. As our focus lies on democracy quality, we examine 99 countries that have been classified by Freedom House as "free" or "partly free" during most years of our research period. Before we present the results, this section describes the measurement of the dependent and independent variables.

3.1. Measurement of the Dependent Variables

Following previous research, we examine climate cooperation commitment and state efforts to mitigate global warming separately. Our indicator for commitment is the climate policy output component from the climate cooperation index created by Bättig, Brander and Imboden (2008); taken from Bättig and Bernauer (2009). It summarizes information on state behavior within the climate change regime (ratification and ratification delay of the UNFCCC and the Kyoto Protocol, reporting and reporting delay, and timely financial contributions to the UNFCCC core budget) from 1990–2005. It varies from 0–1, where higher values imply more cooperative state behavior. Because CO_2 is the most important greenhouse gas, CO_2 emissions per capita can be used to measure climate policy outcomes. The data is taken from the online database of the World Bank World Development Indicators (WDI) (data access in 2017). Following previous studies, we examine pollution levels. Our appendix also studies long-term and short-term changes in CO_2 emissions per capita using cross-sectional and time-series cross-sectional regression analysis.

3.2. Measurement of the Independent Variables

In comparison to earlier research on the relationship between democracy and climate policy, we measure the effect of the democracy quality dimensions side-by-side in one regression model. Following our theoretical discussion, it is important to assess the influence of the

democracy attributes separately. Several studies apply multiple democracy measures with different underlying democracy conceptions in separate regression models (e.g., Midlarsky, 1998; Neumayer, 2002a). However, the democracy quality dimensions should be tested simultaneously. Specific democracy quality indicators might only show a significant effect because they are highly correlated with the other institutional traits.

The effects of the democracy quality dimensions are captured by indicators developed by Lührmann, Marquardt and Mechkova (2017) and the V-Dem project (Coppedge et al., 2017; Pemstein et al., 2017). All indicators are based on expert evaluations. Our indicator of electoral accountability is the "Vertical Accountability Index" of Lührmann et al. (2017, pp. 11ff.). This index focuses on the mechanisms of formal political participation via elections and political parties in the exercise of accountability. It summarizes indicators of the quality of free and fair elections, the percentage of the population that is enfranchised, whether the chief executive is elected, and whether there is the right to organize and participate in political parties. The latter aspect enables us to consider the assumption that competitive elections are crucial. Checks and balances are captured by the "Horizontal Accountability Index" (Lührmann et al., 2017, p. 13), which represents the extent to which state institutions are able to hold the executive branch of the government to account. Three institutions are considered in this regard: the legislature, the judiciary, and special bodies designed for this purpose (e.g., ombudsman). We use the "Equality before the law and individual liberty Index" from V-Dem to measure the democratic subdimension of civil rights. This index captures the extent to which laws are transparent and rigorously enforced, and whether the public administration is impartial, the extent to which citizens enjoy access to justice, the ability to secure property rights, freedom from forced labor, freedom of movement, physical integrity rights, as well as freedom of religion. Finally, for the operationalization of political rights, we apply the "Diagonal Accountability Index" (Lührmann et al., 2017, p. 15), which captures the extent to which citizens are able to hold a government accountable outside of formal political participation. It summarizes information on media freedom, civil society characteristics, freedom of expression, and the degree to which citizens are engaged in politics.

These variables enable us to measure the dimensions of democracy quality separately and test their effects on climate policy simultaneously. In comparison to the disaggregated data from Freedom House and Polity IV, they are more valid and reliable. Both Freedom House measures and the polity2 index have validity and reliability problems with regard to conceptualization, measurement, and aggregation (e.g., Munck & Verkuilen, 2002). In contrast to the indicators from the Democracy Barometer, our variables are not based on policy output and outcome indicators. Finally, in comparison to data from the Bertelsmann Stiftung's Transformation Index, the Democracy Barometer, and Freedom House, the V-Dem indicators cover our country sample and the whole research period we are interested in. We compare the disaggregated measurement approach to the Freedom House (2017) political rights and civil liberties indices and the polity2 indicator from Polity IV (Marshall et al., 2016). They have been used in most studies of democracy and climate policy. The Freedom House indicators are rescaled so that higher values on the measures from Freedom House and Polity IV indicate higher levels of democracy quality.

Our statistical analyses control for additional variables that have been applied in similar studies. Population density is included as it is associated with natural resource use (Spilker, 2012). Since emissions result mainly from economic activities, we consider the level of economic development (GDP per capita) and economic growth (GDP growth). Countries that export fossil fuels should be less likely to participate in climate cooperation and mitigate global warming. Thus, we control for the percentage of merchandise exports that are fossil fuel exports. The effect of international trade is theoretically ambiguous. Our indicator is the percentage of the sum of exports and imports of a country's GDP. Data on our socio-economic variables come from the WDI online database (data accessed in 2017). Recent research suggests that state involvement in international governmental organizations (IGOs) increases a country's willingness and ability to reduce pollution (e.g., Spilker, 2012). Data on country memberships in IGOs comes from Pevehouse, Nordstrom and Warnke (2010). On the domestic level, ENGOs pressure governments to consider environmental issues. ENGO strength is captured by data from Bernauer et al. (2013) on the number of ENGOs registered in a country with the International Union for Conservation of Nature. We consider a country's vulnerability to the consequences of global warming with the climate change index from Bättig, Wild and Imboden (2007) in Bättig and Bernauer (2009). It covers climate variability due to global warming in comparison to natural developments on a scale from 0–1. Higher values indicate higher climate variability. More vulnerable countries should be more active in this policy area (Sprinz & Vaahtoranta, 1994).

The climate policy outcome models test, following EKC-theory (Grossman & Krueger, 1995), a curvilinear effect of GDP per capita. We used the mean-centered variable GDP per capita to avoid problems with non-essential multicollinearity. Our commitment indicator is also included as an independent variable.[1] Countries that are more committed to climate cooperation might be more willing to reduce emissions (Bättig & Bernauer, 2009). Table 1 summarizes the descriptive statistics of all variables included in the statistical analyses.

[1] To acknowledge the endogeneity of commitment to climate cooperation with regard to the explanatory variables, we also examined the climate policy outcome models with the residuals of the commitment model. The results of our climate policy outcome models stay stable.

Table 1. Descriptive statistics.

Variable	Mean	Std. Dev.	Min	Max
Commitment to climate cooperation	.680	.156	.226	.978
CO_2 emissions per capita	4.712	5.271	.063	28.284
Electoral accountability	.975	.479	.031	1.854
Horizontal accountability	.799	.752	–.655	2.254
Political rights	1.106	.573	–.253	2.106
Civil rights	.786	.177	.358	.995
Freedom House Political rights	5.340	1.428	2.315	7.000
Freedom House Civil liberties	5.032	1.240	2.375	7.000
Polity2	5.830	4.409	–7.133	10.000
Population density	172.413	595.542	1.557	5, 861.425
GDP per capita	13143.732	16979.596	270.110	78793.39
GDP growth	3.453	1.619	–1.786	7.553
Fuel exports	11.567	20.972	.002	95.484
Trade openness	75.843	42.91	21.977	359.634
Memberships in IGOs	67.909	19.113	33.933	123.313
ENGO strength	5.818	7.790	.000	51.444
Climate change vulnerability	.489	.136	.261	.897

Notes: N = 99. The units of analysis are country averages from 1990–2010, except commitment to climate cooperation (1990–2005), IGO membership (1990–2005) and vulnerability (1990–2005).

4. Results

4.1. Commitment to Climate Cooperation

Table 2 presents the results of our first dependent variable commitment to climate cooperation from 1990–2005. Positive skewness of variables is reduced using the natural logarithm (ln). Tolerance values of our regression models indicate no problems with multicollinearity. Our results remain stable if we apply robust standard errors. Regarding the control variables, our models show that trade openness contributes to commitment. R^2 values increase considerably when we exclude outliers (see Annex). While nearly all democracy aspects are associated with higher levels of climate policy commitment, only the effect of political rights is significant. Electoral accountability is negatively but insignificantly associated with commitment. This supports our hypotheses 1a, 2a and 3a that electoral and horizontal, as well as civil rights cannot explain country-differences in commitment to climate cooperation. In accordance with hypothesis 4a, we find that countries with high levels of political rights are also more committed.

Models 2–4 in Table 2 show that both Freedom House indicators and the polity2 indicator have a positive and significant effect on commitment. If the democracy quality dimensions are examined in separate models, each contributes significantly to commitment as well (see Annex). Thus, the summary measures exhibit significant positive effects as they either encompass political rights indicators or are highly correlated with political rights.

We performed several robustness analyses (see Annex). Testing cumulative effects of the democracy quality dimensions (Fredriksson & Neumayer, 2013; Gallagher & Thacker, 2008) from 1950–2005/2010, we yield the same conclusions. Our results also remain stable if we control for the Annex-I status and the level of CO_2 emissions per capita. If we add countries classified as "not free" for most years from 1990–2010, the positive effect of political rights on commitment is also significant. The results of the commitment model remain stable if we estimate our models without outliers and if we exclude developed countries or countries from a particular region.

Following our theoretical expectations, political rights might contribute to commitment since they enable citizens and interest groups to pressure the government to consider climate change. However, we find no significant interaction effect between political rights and ENGO influence (see Annex). This finding might result from our indicator of the political influence of ENGOs which just counts the number of domestic ENGOs. In addition, many domestic ENGOs focus on local environmental problems (Never & Betz, 2014, p. 12).

4.2. Climate Change Mitigation

Table 3 summarizes the results for average CO_2 emissions per capita levels (ln) from 1990–2010. In our models, population density, fuel exports and, contrary to our expectations, ENGO strength contribute to emissions. In accordance with EKC-research, we find an inverse U-shaped relationship between GDP per capita and CO_2 emissions per capita. The findings support that growth,

Table 2. Democracy quality dimensions and climate cooperation commitment (OLS regression analysis).

	(1)	(2)	(3)	(4)
Population density (ln)	.008	.004	.013	.000
	(.010)	(.009)	(.009)	(.010)
GDP per capita (ln)	–.002	–.005	–.017	.015
	(.016)	(.014)	(.014)	(.014)
GDP growth	.006	.006	.006	.010
	(0.007)	(.007)	(.007)	(.007)
Fuel exports (ln)	–.008	–.005	.002	–.012
	(.012)	(.012)	(.011)	(.012)
Trade openness (ln)	.091**	.078**	.058*	.072**
	(.035)	(.033)	(.032)	(.035)
Memberships in IGOs	.000	.001	.000	.001
	(.001)	(0.001)	(.001)	(0.001)
ENGO strength (ln)	.036*	.021	.027	.020
	(.021)	(.020)	(.019)	(.022)
Climate change vulnerability	–.028	–.073	–.027	–.116
	(.099)	(.095)	(.091)	(.100)
Electoral accountability	–.009			
	(.060)			
Horizontal accountability	.010			
	(.031)			
Political rights	.106**			
	(.049)			
Civil rights	.138			
	(.158)			
Freedom House Political rights		.061***		
Freedom House Civil liberties		(.013)	.093***	
			(.016)	
Polity2				.013***
				(.004)
Countries	99	99	99	99
R^2	.452	.451	.503	.391

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** p < .01, ** p < .05, * p < .10. The units of analysis are country averages from 1990–2005.

country memberships in IGOs as well as vulnerability are associated with lower emissions levels.

Moreover, the analysis indicates that cross-national variation in the democracy quality dimensions does not explain country-differences in climate policy outcomes. In accordance with our theoretical expectations (hypotheses 1b, 2b, 3b and 4b), we find no effect of vertical and horizontal accountability as well as civil and political rights on cross-national variation in CO_2 emissions. However, while political rights and horizontal accountability are negatively associated with emission levels, electoral accountability and civil rights show positive effects. Both Freedom House indicators and the polity2 indicator are not associated with emission levels.

Our results remain stable if we exclude developed countries (see Annex). We also examined the effects of the democracy quality dimensions on average CO_2 emissions in two sub-periods (1990–1999, 2000–2010). In contrast to our theoretical expectations, we find a significant positive effect of electoral accountability on average CO_2 emissions per capita from 1990–1999, if we exclude countries from the Middle East and North Africa. However, many countries might not have adopted mitigation measures in the early years of the international climate change regime. Additionally, the time it takes for climate policies to affect greenhouse gas emissions has to be considered. The robustness analyses suggest no effect of the democracy quality dimensions on short-term and long-term changes in CO_2 emissions per capita.

It appears that the cross-national variance in political rights explains commitment to climate cooperation better than country-differences in vertical accountability, horizontal accountability and civil rights. Our results are based on correlations. Case studies could give us more insight into the causal processes that link political rights to commitment to international efforts to tackle climate change. In the following, we examine Ecuador in more detail, a country that performs well with re-

Table 3. Democracy quality dimensions and CO_2 emissions per capita, ln (OLS regression analysis).

	(1)	(2)	(3)	(4)
Population density (ln)	.087**	.093**	.089**	.092**
	(.039)	(.038)	(.040)	(.038)
GDP per capita (ln)	.856***	.897***	.911***	.900***
	(.069)	(.061)	(.063)	(.054)
GDP per capita (ln) squared	–.128***	–.135***	–.135***	–.133***
	(.028)	(.026)	(.026)	(.027)
GDP growth	–.092**	–.080**	–.080**	–.077**
	(.035)	(.034)	(.034)	(.035)
Fuel exports (ln)	.249***	.235***	.229***	.235***
	(.051)	(.148)	(.051)	(.050)
Trade openness (ln)	.077	.104	.096	.102
	(0.157)	(.004)	(.146)	(.147)
Memberships in IGOs	–.014***	–.014***	–.013***	–.014***
	(.004)	(.004)	(.004)	(.004)
ENGO strength (ln)	.142	.152*	.153*	.149*
	(.086)	(.085)	(.085)	(.086)
Climate change vulnerability	–1.222***	–1.218***	–1.237***	–1.221**
	(.419)	(.570)	(.417)	(.412)
Commitment to climate cooperation	.535	.467	.555	.465
	(.438)	(.437)	(.456)	(.418)
Electoral accountability	.347			
	(.268)			
Horizontal accountability	–.034			
	(.136)			
Political rights	–.248			
	(.212)			
Civil rights	.391			
	(0.675)			
Freedom House Political rights		.017		
		(.060)		
Freedom House Civil liberties			–.011	
			(.081)	
Polity2				.007
				(0.017)
Countries	99	99	99	99
R^2	.917	.915	.915	.915

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** p < .01, ** p < .05, * p < .10. The units of analysis are country averages of from 1990–2010. IGO membership, ENGO strength and commitment refer to country averages from 1990 to 2005.

gard to political rights and commitment to climate cooperation but has deficits in another democracy quality dimension (horizontal accountability).[2] Simultaneously, Ecuador performs above-average on our measure of climate cooperation commitment. The political system of Ecuador is characterized by free and fair elections and the respect of civil rights. Ecuador shows deficits with regard to horizontal accountability during our research period. The independence of the judiciary in Ecuador is restricted. The executive and the legislative branches of government have repeatedly influenced court decisions for their benefit. In comparison, there are no restrictions on the freedoms of association, expression, or the press. Ecuador has above average values on our political rights indicator. Since the 1990s, civil society organizations representing indigenous people have gained in influence in the Ecuadorian political system. They have also been organized around environmental issues such as the ecological consequences of petroleum extraction in the Amazon lowlands. Case studies have shown that the Ecuadorian government's international climate policy has been influenced by civil society organizations (e.g. Espinosa, 2013; Martin, 2011). In the mid-1990s domestic NGOs, scientists and indigenous groups de-

[2] We use information from the country reports of the Bertelsmann Stiftung's Transformation Index from 2003 and 2006.

manded a halt to oil drilling in the Ecuadorian Amazon. Espinosa (2013) argues that the environmental interests groups were able to change the public discourse on petroleum extraction in the oil exporting country. The government of President Correa was responsive to the influence of civil society organizations (Martin, 2011, pp. 26f.). It worked together with domestic NGOs to formulate a proposal to the international community (Martin, 2011, pp. 31f., 39). The Ecuadorian government finally proposed that it would commit itself to not extracting the country's largest oil reserves in the Yasuní National Park in the Amazon and thus avoid considerable greenhouse gas emissions under the condition that it would receive international financial compensation (Martin, 2011, p. 22). Martin (2011, p. 31) concludes that the Ecuadorian government represented the climate policy position of the civil society organizations on the international level. The example of Ecuador shows that political rights enabled civil society to influence the country's climate policy (Martin, 2011, pp. 27f.). This case also illustrates the difference between climate policy commitment and climate change mitigation. The Ecuadorian government introduced the condition of financial compensation. It was, in the end, unwilling to stop oil drilling in the Yasuní National Park since there was little financial support from the international community. With regard to commitment to climate policy goals, states appear to be responsive to the climate policy demands of citizens and civil society organizations facilitated by political rights although they are more reluctant to implement climate change mitigation policies.

5. Discussion and Conclusion

Our results show that the conceptualization and measurement of democracy matters in comparative climate policy research. With the measures from Freedom House and Polity IV, we observe positive effects on climate policy commitment. In comparison, the disaggregated measurement approach indicates that only the realization of political rights is crucial. The results suggest that previous research might have only found significant effects of democracy quality measures on commitment because they contain or are highly correlated with the dimension of political rights. This finding sheds new light on the causal mechanisms that link democracy to commitment. In accordance with our theoretical expectations, we find that the positive effect of democracy quality depends on political rights. It appears that electoral and horizontal accountability, as well as civil rights, are not decisive for country-differences in commitment to climate cooperation. The effect of competitive elections depends on the electoral success of supporters and opponents of climate protection. Horizontal accountability provides incentives *and* constraints to participation in climate cooperation. The effect of civil rights depends on a country's existing climate policy regulations. We do not believe that political rights alone contribute to climate cooperation either. Climate change mitigation is dependent on public awareness of global warming and support for climate protection. For instance, Harrison and Sundstrom (2010) conclude from their case studies of Kyoto Protocol ratification that the EU member states and Japan were able to ratify the agreement since public and business interests supported it. In comparison, Australia withdrew its ratification and the United States never ratified it since public climate concern was low and industry interest groups opposed ratification. Our argument is that political rights enable supporters of climate change mitigation to raise awareness of climate change, articulate their climate policy preferences, and mobilize public support in the first place. These democratic freedoms make it, therefore, more likely that diffuse interests such as climate protection are considered in the political process (see also Martin, 2011). For instance, Never and Betz (2014, p. 12) demonstrate that civil society organizations had no influence on climate policy in India and South Africa since they had no access to the domestic political decision-making process. In accordance with our theoretical expectations and earlier research, we find little support for the claim that democracy quality or the democracy quality dimensions—electoral and horizontal accountability, political and civil rights—matter for climate policy outcomes. If at all, electoral accountability may have been associated with higher emission levels in the 1990s. While political rights contribute to commitment to climate cooperation, they are not associated with lower emission levels. An explanation for this finding is that it is, presumably, more difficult for the public to control the implementation of environmental policies (Cao & Prakash, 2012).

The main implication of our results for the analysis of substantive research questions in empirical democracy research is that it could be fruitful to study the implications of democracy on a disaggregated basis to gain more analytical clarity and, therefore, to conceptualize and measure democracy quality dimensions separately. Further climate policy research could investigate the relationship between political rights and global warming policy in more detail. Our study has focused on institutional traits. To gain more knowledge on the effect of political rights on climate cooperation, more attention should also be given to the policy-preferences of political decision-makers and citizens and their interaction with political rights. It would also be interesting to examine the relative importance of political rights such as freedom of expression, freedom of association, and civil society autonomy for climate policy. We also have not considered elements of direct democracy.

Acknowledgments

The authors would like to thank the editors of the special issue, Heiko Giebler, Saskia Ruth and Dag Tanneberg, and the two anonymous reviewers for their helpful comments. This article has benefited also from the valuable

comments and suggestions of Dieter Fuchs, Brigitte Geissel, Hans-Joachim Lauth, Quinton Mayne, Elena Rinklef, Gary S. Schaal, Oliver Schlenkrich, Sebastian Stier and Stefan Wurster on previous versions of this manuscript. This work was supported by the German Research Foundation (DFG) within the funding programme Open Access Publishing.

Conflict of Interests

The authors declare no conflict of interests.

References

Barrett, S., & Graddy, K. (2000). Freedom, growth, and the environment. *Environment and Development Economics*, *5*, 433–456.

Bättig, M. B., & Bernauer, T. (2009). National institutions and global public goods. Are democracies more cooperative in climate change policy? *International Organization*, *63*(2), 281–308.

Bättig, M. B., Brander, S., & Imboden, D. (2008). Measuring countries' cooperation within the international climate change regime. *Environmental Science & Policy*, *11*, 478–489.

Bättig, M. B., Wild, M., & Imboden, D. (2007). A climate change index. Where climate change may be most prominent in the 21st century. *Geophysical Research Letters*, *34*(1), 1–6.

Beeson, M. (2010). The coming of environmental authoritarianism. *Environmental Politics*, *19*(2), 276–294.

Bernauer, T., Böhmelt, T., & Koubi, V. (2013). Is there a democracy-civil society paradox in global environmental governance? *Global Environmental Politics*, *13*(1), 88–107.

Bernauer, T., Kalbhenn, A., Koubi, V., & Spilker, G. (2010). A comparison of international and domestic sources of global governance dynamics. *British Journal of Political Science*, *40*, 509–538.

Bernauer, T., & Koubi, V. (2009). Effects of political institutions on air quality. *Ecological Economics*, *68*, 1355–1365.

Böhmelt, T., Böker, M., & Ward, H. (2016). Democratic inclusiveness, climate policy outputs, and climate policy outcomes. *Democratization*, *23*(7), 1272–1291.

Bueno De Mesquita, B., Smith, A., Siverson, R. M., & Morrow, J. D. (2003). *The logic of political survival*. Cambridge, MA and London: MIT Press.

Burnell, P. (2009). *Is democratization bad for global warming?* (CSGR Working Paper no. 257). Warwick: University of Warwick.

Burnell, P. (2012). Democracy, democratization and climate change. Complex relationships. *Democratization*, *19*(5), 813–842.

Cao, X., & Prakash, A. (2012). Trade competition and environmental regulations: Domestic political constraints and issue visibility. *The Journal of Politics*, *74*(1), 66–82.

Cao, X., & Ward, H. (2015). Winning coalition size, state capacity, and time horizons. An application of modified selectorate theory to environmental public goods provision. *International Studies Quarterly*, *59*, 264–279.

Christoff, P., & Eckersley, R. (2011). Comparing state responses. In J. S. Dyrzek, R. B. Norgaard, & D. Schlosberg (Eds.), *Oxford handbook of climate change and society* (pp. 431–448). Oxford: Oxford University Press.

Collier, D., & Levitsky, S. (1997). Democracy with adjectives. Conceptual innovation in comparative research. *World Politics*, *49*, 430–451.

Congleton, R. D. (1992). Political institutions and pollution control. *The Review of Economics and Statistics*, *74*(3), 412–421.

Coppedge, M., Gerring, J., Lindberg, S., Skaaning, S.-E., Teorell, J., Altman, D., . . . Wilson, S. (2017). V-Dem [country-year/country-date] Dataset v7.1. *Varieties of Democracy (V-Dem) Project*. Retrieved from https://www.v-dem.net/en/data/data-version-7-1

Diamond, L., & Morlino, L. (2005). Introduction. In L. Diamond & L. Morlino (Eds.), *Assessing the quality of democracy* (pp. xi–xliii). Baltimore, MD: The John Hopkins University Press.

Espinosa, C. (2013). The riddle of leaving the oil in the soil—Ecuador's Yasuní project from a discourse perspective. *Forest and Economics*, *36*, 27–36.

Fiorino, D. (2011). Explaining national environmental performance: Approaches, evidence, and implications. *Policy Sciences*, *44*, 367–389.

Fliegauf, M. T., & Sanga, S. (2010). *Steering toward climate tyranny?* Retrieved from http://www.eisa-net.org/sitecore/content/be-bruga/eisa/publications/feeds/stockholm.aspx

Fredriksson, P. G., & Gaston, N. (2000). Ratification of the 1992 climate change convention. What determines legislative delay? *Public Choice*, *104*, 345–368.

Fredriksson, P. G., & Neumayer, E. (2013). Democracy and climate change policies. Is history important? *Ecological Economics*, *95*, 11–19.

Fredriksson, P. G., & Neumayer, E. (2016). Corruption and climate change policies: Do the bad old days matter? *Environmental and Resource Economics*, *63*(2), 451–469.

Fredriksson, P. G., Neumayer, E., Damania, R., & Gates, S. (2005). Environmentalism, democracy, and pollution control. *Journal of Environmental Economics and Management*, *49*, 343–365.

Freedom House. (2017). Freedom in the world 2017. *Freedom House*. Retrieved from https://freedomhouse.org/report/freedom-world/freedom-world-2017

Gallagher, K. P., & Thacker, S. C. (2008). *Democracy, income, and environmental quality* (Working Paper Series no. 164). Amherst, MA: Political Economy Research Institute (PERI).

Garman, S. (2014). Do government ideology and frag-

mentation matter for reducing CO_2-emissions? Empirical evidence from OECD countries. *Ecological Economics*, *105*, 1–10.

Geissel, B., Kneuer, M., & Lauth, H.-J. (2016). Measuring the quality of democracy: Introduction. *International Political Science Review*, *37*(5), 571–579.

Giley, B. (2012). Authoritarian environmentalism and China's response to climate change. *Environmental Politics*, *21*(2), 287–307.

Gleditsch, N. P., & Sverdrup, B. (2002). Democracy and the environment. In E. A. Page & M. Redclift (Eds.), *Human security and the environment. International comparisons* (pp. 45–70). Cheltenham and Northampton: Edward Elgar.

Grossman, G. M., & Krueger, A. B. (1995). Economic growth and environment. *Quarterly Journal of Economics*, *110*, 353–377.

Hardin, G. (1968). The tragedy of the commons. *Science*, *162*, 1243–1248.

Harrison, K., & Sundstrom, L. M. (2010). Conclusion: The comparative politics of climate change. In K. Harrison & L. M. Sundstrom (Eds.), *Global commons, domestic decisions. The comparative politics of climate change* (pp. 261–290). Cambridge, MA and London: MIT Press.

Held, D., & Hervey, A. (2011). Democracy, climate change and global governance. Democratic agency and the policy menu ahead. In D. Held, A. Fane-Hervey, & M. Theros (Eds.), *The governance of climate change. Science, economics, politics and ethics* (pp. 89–110). Cambridge and Malden, MA: Polity Press.

Holden, B. (2002). *Democracy and global warming*. London and New York, NY: Continuum.

Kim, S. Y., & Wolinsky-Nahmias, Y. (2014). Cross-national public opinion on climate change: The effects of affluence and vulnerability. *Global Environmental Politics*, *14*(1), 79–106.

Kneuer, M. (2012). Who is greener? Climate action and political regimes: Trade-offs for national and international actors. *Democratization*, *19*(5), 865–888.

Li, Q., & Reuveny, R. (2006). Democracy and environmental degradation. *International Studies Quarterly*, *50*, 935–956.

List, J. A., & Sturm, D. M. (2006). How elections matter: Theory and evidence from environmental policy. *The Quarterly Journal of Economics*, *121*(4), 1249–1281.

Lührmann, A., Marquardt, K. L., & Mechkova, V. (2017). *Constraining governments: New indices of vertical, horizontal and diagonal accountability* (Working Paper Series no. 46). Gothenburg: The Varieties of Democracy Institute.

Madden, N. J. (2014). Green means stop: Veto players and their impact on climate-change policy outputs. *Environmental Politics*, *23*(4), 570–589.

Marshall, M. G., Gurr, T. R., & Jaggers, K. (2016). Polity IV project, political regime characteristics and transitions, 1800–2015. *Center for Systemic Peace*. Retrieved from http://www.systemicpeace.org/inscrdata.html

Martin, P. L. (2011). Global governance from the Amazon: Leaving oil underground in Yasuní National Park, Ecuador. *Global Environmental Politics*, *11*(4), 22–42.

Merkel, W. (2004). Embedded and defective democracies. *Democratization*, *11*(5), 33–58.

Midlarsky, M. I. (1998). Democracy and the environment. An empirical assessment. *Journal of Peace Research*, *35*(3), 341–361.

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

Neumayer, E. (2002a). Do democracies exhibit stronger international environmental commitment? A cross-country analysis. *Journal of Peace Research*, *39*(2), 139–164.

Neumayer, E. (2002b). Does trade openness promote multilateral environmental cooperation? *The World Economy*, *25*(6), 815–832.

Never, B., & Betz, J. (2014). Comparing the climate policy performance of emerging economies. *World Development*, *59*, 1–15.

Ophuls, W. (1977). *Ecology and the politics of scarcity. Prologue to a political theory of the steady state*. San Francisco, CA: W. H. Freeman and Company.

Payne, R. A. (1995). Freedom and the environment. *Journal of Democracy*, *6*(3), 41–55.

Pemstein, D., Marquardt, K. L., Tzelgov, E., Wang, Y., Krusell, J., & Miri, F. (2017). *The V-Dem measurement model: Latent variable analysis for cross-national and cross-temporal expert-coded data* (Working Paper no. 21, 2nd ed.). Gothenburg: Varieties of Democracy Institute.

Pevehouse, J. C., Nordstrom, T., & Warnke, K. (2010). The COW-2 international organisations dataset version 2.3. *Conflict Management and Peace Science*, *21*, 101–119.

Pickel, S., Stark, T., & Breustedt, W. (2015). Assessing the quality of quality measures of democracy. A theoretical framework and its empirical application. *European Political Science*, *14*, 496–520.

Shearman, D., & Smith, J. W. (2007). *The climate change challenge and the failure of democracy*. Westport, CT and London: Praeger Publishers.

Spilker, G. (2012). Helpful organisations. Membership in inter-governmental organisations and environmental quality in developing countries. *British Journal of Political Science*, *42*(02), 345–370.

Spilker, G. (2013). *Globalization, political institutions and the environment in developing countries*. New York, NY and London: Routledge.

Sprinz, D., & Vaahtoranta, T. (1994). The interest-based explanation of international environmental policy. *International Organization*, *48*(1), 77–105.

Stehr, N. (2015). Democracy is not an inconvenience. *Nature*, *525*(24), 449–450.

von Stein, J. (2008). The international law and politics of climate change. Ratification of the United Nations

Framework Convention and the Kyoto Protocol. *Journal on Conflict Resolution*, *52*(2), 243–268.
Ward, H. (2008). Liberal democracy and sustainability. *Environmental Politics*, *17*(3), 386–409.
Wilson, J. K., & Damania, R. (2005). Corruption, political competition and environmental policy. *Journal of Environmental Economics and Management*, *49*, 516–535.
Winslow, M. (2007). Is democracy good for the environment? *Journal of Environmental Planning and Management*, *48*(5), 771–783.
Wurster, S. (2011). Sustainability and regime type: Do democracies perform better in promoting sustainable development than autocracies? *Zeitschrift für Staats- und Europawissenschaften*, *9*(4), 538–559.
Wurster, S. (2013). Comparing ecological sustainability in autocracies and democracies. *Contemporary Politics*, *19*(1), 76–93.
Wurster, S., Auber, B., Metzler, L., & Rohm, C. (2015). Institutionelle Voraussetzungen nachhaltiger Politikgestaltung [Institutional preconditions for sustainable policy-making]. *Zeitschrift für Politik*, *62*(2), 177–196.

About the Authors

Romy Escher is a research assistant and PhD candidate at the Institute of Political Science of the University of Regensburg. Her research interests include Climate Policy, Comparative Policy Analysis, and empirical research methods.

Melanie Walter-Rogg is Professor of Political Science and Methodology at the University of Regensburg. She has co-edited *The Political Ecology of the Metropolis* (ECPR Press, 2013) and authored a number of scholarly articles and book chapters related to Public Policy Analysis, Metropolitan Governance, Urban Democracy as well as Political Culture and Political Behaviour.

Annex

1. Models with Indicators of the Effective Power to Govern

The concept of “embedded democracy” from Merkel (2004) distinguishes five interdependent partial regimes, namely electoral regime, political rights, civil rights, horizontal accountability, and the effective power to govern. Table A1 tests, also the effect of the effective power to govern. The effective power to govern means that only elected authorities participate in political decision-making processes. We considered three indicators: The Domestic Autonomy (v2svdomaut) item from the V-Dem Project (Coppedge et al., 2017; Pemstein et al., 2017) evaluates a state’s autonomy in domestic politics from external actors. The variable International autonomy (v2svinlaut) from the V-Dem Project (Coppedge et al., 2017; Pemstein et al., 2017) captures a state’s independence in foreign policy from external actors. The external constraints (exconst) indicator from Marshall et al. (2016) measures constraints on the government by various accountability groups (e.g., the legislature, the judiciary, the military). The results of our main analysis remain stable. In addition, we find a significant positive effect of a state’s autonomy in foreign policy on commitment. In comparison, the domestic and international autonomy does not matter for CO_2 emissions per capita. The exconst indicator overlaps with our measure of horizontal accountability. However, the results remain the same, when we exclude horizontal accountability.

Table A1. Effective power to govern and commitment to climate cooperation and CO_2 emissions per capita, ln (OLS regression analysis).

	(2) Commitment to climate cooperation	(4) CO_2 emissions per capita, ln
Population density (ln)	.004	–.090**
	(.009)	(.040)
GDP per capita (ln)	–.007	.862***
	(.016)	(.071)
GDP per capita (ln) squared		.130***
		(.031)
GDP growth	.004	–.100***
	(.007)	(.035)
Fuel exports (ln)	–.009	.254***
	(.012)	(.052)
Trade openness (ln)	.097***	.063
	(.035)	(.162)
Memberships in IGOs	.000	–.014***
	(.001)	(.004)
ENGO strength (ln)	.032	.135*
	(.021)	(.072)
Climate change vulnerability	.035	–1.111**
	(.100)	(.435)
Commitment to climate cooperation		.435
		(.447)
Electoral accountability	–.047	.224
	(.063)	(.281)
Horizontal accountability	.012	–.034
	(.031)	(.137)
Political rights	.093*	–.248
	(.049)	(.216)
Civil rights	.178	.447
	(.157)	(.684)
Domestic autonomy	.002	–.162
	(.041)	(.183)
International autonomy	.068*	.118
	(.035)	(.157)
External constraints	.005	.034
	(.004)	(.022)

Table A1. Effective power to govern and commitment to climate cooperation and CO_2 emissions per capita, ln (OLS regression analysis). (Cont.)

	(2) Commitment to climate cooperation	(4) CO_2 emissions per capita, ln
Countries	99	99
R^2	.500	.921

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** p < .01, ** p < .05, * p < .10. The units of analysis are country averages from 1990–2005 (Model 1) and from 1990–2010 (Model 2). IGO membership, ENGO strength and commitment refer to country averages from 1990–2005.

2. Test of the Effect of the Democracy Quality Dimensions in Separate Regression Models

The summary measures of democracy in our main analysis showed significant positive effects on commitment. In comparison, the separate analysis of the democracy quality dimensions found that only political rights matter for commitment. Table A2 tests the effect of the democracy quality dimensions in separate regression models. Table A2 shows that in separate regression models all democracy quality dimensions are significantly associated with commitment to climate cooperation. This result indicates that summary measures of democracy, as well as our indicators of electoral, horizontal accountability, and civil rights in Table A3, are only significantly associated with commitment as they are highly correlated with political rights.

Table A3 examines the effect of the democracy quality dimensions on CO_2 emissions levels in separate regression models. In accordance with our main analysis, it suggests that the democracy quality dimensions do not matter for pollution.

Table A2. Democracy quality dimensions and commitment to climate cooperation (OLS regression analysis).

	(1)	(2)	(3)	(4)
Population density (ln)	.000	.001	.007	.004
	(.010)	(.010)	(.009)	(.010)
GDP per capita (ln)	.003	.007	.004	–.001
	(.016)	(.015)	(.014)	(.015)
GDP growth	.003	.005	.006	.004
	(.008)	(.008)	(.007)	(.007)
Fuel exports (ln)	–.011	–.013	–.010	–.009
	(.013)	(.012)	(.012)	(.012)
Trade openness (ln)	.086**	.089**	.098***	.066*
	(.035)	(.036)	(.034)	(.034)
Memberships in IGOs	.000	.000	.000	.000
	(.001)	(.001)	(.001)	(.001)
ENGO strength (ln)	.028	.038*	.035*	.037*
	(.022)	(.021)	(.020)	(.021)
Climate change vulnerability	.097	–.055	–.046	–.037
	(.101)	(.103)	(.096)	(.100)
Electoral accountability	.134***			
	(.046)			
Horizontal accountability		.076***		
		(.025)		
Political rights			.139***	
			(.030)	
Civil rights				.411***
				(.107)
Countries	99	99	99	99
R^2	.374	.378	.445	.411

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** p < .01, ** p < .05, * p < .10. The units of analysis are country averages from 1990–2005.

Table A3. Democracy quality dimensions and CO_2 emissions per capita, ln (OLS regression analysis).

	(1)	(2)	(3)	(4)
Population density (ln)	.095**	.091**	.088**	.095**
	(.038)	(.038)	(.039)	(.039)
GDP per capita (ln)	.867***	.907***	.914***	.890***
	(.066)	(.060)	(.056)	(.062)
GDP per capita (ln) squared	–.131***	–.135***	–.136***	–.136***
	(.026)	(.026)	(.026)	(.026)
GDP growth	–.085**	–.080**	–.080**	–.081**
	(.034)	(.034)	(.034)	(.034)
Fuel exports (ln)	.247***	.231***	.228***	.237***
	(.051)	(.050)	(.049)	(.050)
Trade openness (ln)	.130	.097	.084	.098
	(.038)	(.149)	(.152)	(.146)
Memberships in IGOs	–.014***	–.014***	–.013***	–.014***
	(.004)	(.004)	(.004)	(.004)
ENGO strength (ln)	.152*	.154*	.152*	.156*
	(.085)	(.085)	(.085)	(.085)
Climate change vulnerability	–1.184***	–1.229***	–1.247***	–1.199***
	(.412)	(.417)	(.416)	(.416)
Commitment to climate cooperation	.396	.525	.583	.444
	(.411)	(.412)	(.434)	(.422)
Electoral accountability	.189			
	(.195)			
Horizontal accountability		.001		
		(.106)		
Political rights			–.045	
			(.016)	
Civil rights				.238
				(.479)
Countries	99	99	99	99
R^2	.916	.915	.915	.915

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** p < .01, ** p < .05, * p < .10. The units of analysis are country averages of from 1990–2010. IGO membership, ENGO strength and commitment refer to country averages from 1990–2005.

3. Main Models without Influential Cases

Table A4 and A5 present our main regression models without influential cases. First, countries with standardized residuals of at least +/−2 were successively excluded. Second, outliers were identified using Cook's D (> 4/n), Leverage (> 3*k/n) and DfBetas (> +/−1). The results of our commitment models stay stable (see Table A4).

With regard to CO_2 emissions per capita, the positive effect of political rights on CO_2 emissions per capita becomes significant (see Table A5). However, this effect turns insignificant again if we estimate further robustness analyses with and without influential cases.

Table A4. Democracy quality dimensions and climate cooperation commitment without influential cases (OLS regression analysis).

	(1)	(2)	(3)	(4)
Population density (ln)	.013*	.003	.002	.002
	(.007)	(.006)	(.007)	(.007)
GDP per capita (ln)	.008	.014	.008	.021
	(.011)	(.010)	(.012)	(.011)
GDP growth	.002	.000	–.003	–.005
	(.005)	(.005)	(.006)	(.006)
Fuel exports (ln)	.003	.014*	.013	.003
	(.008)	(.008)	(.009)	(.009)
Trade openness (ln)	.129	.088***	.069***	.108***
	(.025)	(.023)	(.026)	(.028)
Memberships in IGOs	.000	.001**	.000	.001
	(.000)	(.001)	(.001)	(.001)
ENGO strength (ln)	.031**	.009	.024	.006
	(.015)	(.013)	(.015)	(.017)
Climate change vulnerability	.026	–.086	–.095	–.137
	(.065)	(.063)	(.069)	(.072)
Electoral accountability	–.067			
	(.046)			
Horizontal accountability	.014			
	(.022)			
Political rights	.169***			
	(.040)			
Civil rights	.024			
	(.119)			
Freedom House Political rights		.039***		
		(.010)		
Freedom House Civil liberties			.065***	
			(.014)	
Polity2				.009**
				(.004)
Countries	78	76	82	84
R^2	.725	.724	.678	.569

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** $p < .01$, ** $p < .05$, * $p < .10$. The units of analysis are country averages from 1990–2005.

Table A5. Democracy quality dimensions and CO_2 emissions per capita, ln without influential cases (OLS regression analysis).

	(1)	(2)	(3)	(4)
Population density (ln)	.042	.048	.040	.053*
	(.034)	(.035)	(.036)	(.034)
GDP per capita (ln)	.832***	.896***	.896***	.893***
	(.055)	(.057)	(.060)	(.054)
GDP per capita (ln) squared	–.135***	–.145***	–.140***	–.143***
	(.023)	(.022)	(.023)	(.022)
GDP growth	–.127***	–.101***	–.087***	.092***
	(.027)	(.031)	(.031)	(.030)
Fuel exports (ln)	.193***	.171***	.176***	.208***
	(.044)	(.045)	(.044)	(.042)
Trade openness (ln)	.337***	.316***	.316***	.278**
	(.121)	(.127)	(.127)	(.129)
Memberships in IGOs	–.007*	–.009**	–.008**	–.007**
	(.003)	(.003)	(.004)	(.003)
ENGO strength (ln)	.161**	.227***	.222***	.165
	(.067)	(.070)	(.069)	(.072)
Climate change vulnerability	–1.805***	–1.562***	–1.562***	–1.220***
	(.357)	(.374)	(.380)	(.369)
Commitment to climate cooperation	.275	.183	.278	–.222
	(.342)	(.366)	(.388)	(.369)
Electoral accountability	.355*			
	(.207)			
Horizontal accountability	–.054			
	(.104)			
Political rights	–.413**			
	(.181)			
Civil rights	.225			
	(.538)			
Freedom House Political rights		–.062		
		(.058)		
Freedom House Civil liberties			–.079	
			(.072)	
Polity2				–.016
				(.017)
Countries	86	88	90	86
R^2	.958	.950	.9476.952	

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** p < .01, ** p < .05, * p < .10. The units of analysis are country averages of from 1990–2010. IGO membership, ENGO strength and commitment refer to country averages from 1990–2005.

4. Jackknife Analysis

Using jackknife analysis, we tested whether our results depend on certain country groups. Table A6 tests the effect of our main models without developed countries. We identified developed countries by OECD-membership. The effect of political rights on commitment becomes insignificant if we exclude developed countries. However, it becomes significant if we exclude outliers. The following countries are identified as outliers in model 1 in Table A6: Central African Republic, Republic Congo, Costa Rica, Dominican Republic, Gabon, Gambia, Kuwait, New Zealand, Singapore, Ukraine, and Zambia. The results of CO_2 emissions per capita model remain stable.

Table A6. Democracy quality dimensions, commitment to climate cooperation and CO_2 emissions per capita, ln in non-developed countries (OLS regression analysis).

	(1) Commitment to climate cooperation	(2) Commitment to climate cooperation without influential cases	(3) CO_2 emissions per capita, ln	(4) CO_2 emissions per capita, ln without influential cases
Population density (ln)	.006	.002	.107*	.016
	(.011)	(.009)	(.053)	(.042)
GDP per capita (ln)	.005	.024	.798***	.723***
	(.045)	(.017)	(.108)	(.076)
GDP per capita (ln) squared			–.121**	–.128***
			(.057)	(.041)
GDP growth	.009	.010	–.122**	–.190***
	(.009)	(.007)	(.049)	(.030)
Fuel exports (ln)	–.013	.000***	.290***	.189***
	(.015)	(.011)	(.071)	(.047)
Trade openness (ln)	.039	.066*	.203	.679***
	(.039)	(.034)	(.258)	(.165)
Memberships in IGOs	–.001	–.001	–.012	.004
	(.002)	(.001)	(.008)	(.005)
ENGO strength (ln)	.012	.012*	.165	.287***
	(.026)	(.020)	(.125)	(.079)
Climate change vulnerability	.034	.088	–1.157**	–2.061***
	(.111)	(.080)	(.523)	(.370)
Commitment to climate cooperation			1.148	1.000**
			(.652)	(.437)
Electoral accountability	.056	–.042	.199	–.021
	(.071)	(.053)	(.353)	(.222)
Horizontal accountability	–.002	–.002	–.033	–.034
	(.035)	(.027)	(.171)	(.108)
Political rights	.065	.084*	–.201	.217
	(.056)	(.048)	(.307)	(.200)
Civil rights	.084	.144	.357	–.122
	(.173)	(.139)	(.918)	(.589)
Countries	68	57	68	56
R^2	.321	.488	.891	.962

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** $p < .01$, ** $p < .05$, * $p < .10$. The units of analysis are country averages of from 1990–2005 (Model 1–2) and from 1990–2010 (Model 2–4). IGO membership, ENGO strength and commitment refer to country averages from 1990–2005.

The results of our analysis of commitment to climate cooperation also remain stable if we exclude countries from particular regions. Tables A7, A8 and A9 summarize the results of our jackknife-analysis. We applied the World Bank country classification by region (https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups). The positive effect of political rights on commitment to climate cooperation remains significant in nearly all regression models (see Table A6). It becomes marginally insignificant if countries from East European or from Southeastern European countries are removed from the analysis. However, if we estimate the models without influential cases the effect of political rights becomes significant again (see Table A7). If we exclude countries from Sub-Sahara Africa, we find a significant negative effect of electoral accountability on commitment. In addition, several models without outliers exhibit

a significant negative effect of electoral accountability and a significant positive effect of civil rights on commitment. However, these effects are not as robust as the effect of political rights. In our analysis without countries from the Middle East and North Africa, electoral accountability is significantly associated with CO_2 emissions per capita (see Table A9). Model 2 and 4 in table A14 show that the significant positive effect of electoral accountability, in the analysis of our sample without countries from the Middle East and Northern Africa, is only stable for the period from 1990–1999. Political rights are negatively associated with CO_2 emissions per capita if we exclude countries from Central Asia or Latin America (see Model 1 and 4 in Table A9). However, this effect becomes insignificant if we estimate further robustness analyses.

Table A7. Jackknife analysis: Democracy quality dimensions and commitment to climate cooperation (OLS regression analysis).

	(1) Central Asia	(2) East Asia and Pacific	(3) East Europe	(4) Latin American	(5) Middle East and North Africa	(6) North America	(7) South Asia	(8) South East Europe	(9) Sub-Saharan Africa	(10) Western European
Population density (ln)	.008 (.068)	.016 (.011)	.008 (.010)	.005 (.012)	.010 (.010)	.009 (.010)	.012 (.010)	.006 (.010)	–.010 (.011)	.009 (.010)
GDP per capita (ln)	–.002 (–.015)	–.005 (.018)	–.005 (.017)	–.012 (.019)	.004 (.017)	.002 (.017)	–.008 (.018)	.001 (.017)	.005 (.021)	–.004 (0.18)
GDP growth	.007 (.008)	.004 (.008)	–.001 (.009)	.009 (.008)	.008 (.008)	.007 (.007)	.008 (.008)	.006 (.008)	.007 (.008)	.008 (.008)
Fuel exports (ln)	–.008 (.012)	–.006 (.013)	–.009 (.013)	–.011 (.015)	.000 (.012)	–.009 (.012)	–.007 (.013)	–.012 (.012)	–.005 (.013)	–.003 (.014)
Trade openness (ln)	.091** (.035)	.085** (.018)	.097*** (.036)	.115*** (.041)	.073** (.036)	.079** (.036)	.089** (.037)	.082** (.035)	.094** (.038)	.073* (.042)
Memberships in IGOs	.000 (.001)	.000 (.001)	.000 (.001)	.000 (.001)	–.001 (.001)	.000 (.001)	.000 (.001)	.000 (.001)	.001 (.001)	–.001 (.001)
ENGO strength (ln)	.035* (.021)	.047** (.022)	.042** (.021)	.037 (.024)	.024 (.021)	.038* (.021)	.050** (.023)	.023 (.021)	.031 (.168)	.030 (.024)
Climate change vulnerability	–.024 (.100)	–.005 (.102)	–.018 (.106)	.032 (.188)	–.040 (.099)	–.039 (.099)	–.081 (.109)	–.058 (.100)	.011 (.100)	.003 (.066)
Electoral accountability	–.006 (.061)	–.028 (.062)	–.015 (.062)	.018 (.074)	.021 (.062)	–.024 (.061)	–.021 (.065)	.058 (.062)	–.138* (.072)	.020 (.066)
Horizontal accountability	.011 (.032)	.016 (.032)	.014 (.032)	–.007 (.039)	.015 (.073)	.009 (.031)	.016 (.034)	.014 (.032)	.005 (.035)	–.012 (.035)
Political rights	.107** (.049)	.099* (.058)	.083 (.108) .051)	.131** (.060)	.095* (.052)	.112** (.049)	.124** (.052)	.071 (.157) .050)	.136** (.055)	.102* (.054)
Civil rights	.126 (.161)	.206 (.187)	.162 (.160)	.205 (.183)	.100 (.158)	.142 (.158)	.071 (.171)	.042 (.158)	.267 (.199)	.167 (.169)
Countries	98	87	90	78	95	97	94	93	77	83
R^2	.451	.489	.438	.529	.465	.457	.454	.465	.470	.380

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** $p < .01$, ** $p < .05$, * $p < .10$. The units of analysis are country averages from 1990–2005.

Table A8. Jackknife analysis: Democracy quality dimensions and commitment to climate cooperation without outliers (OLS regression analysis).

	(1) Central Asia	(2) East Asia and Pacific	(3) East Europe	(4) Latin American	(5) Middle East and North Africa	(6) North America	(7) South Asia	(8) South East Europe	(9) Sub-Saharan Africa	(10) Western European
Population density (ln)	.008	.013*	.015**	.019***	.015**	.020***	.019***	.006	.005	.007
	(.007)	(.008)	(.006)	(.007)	(.007)	(.006)	(.007)	(.007)	(.008)	(.010)
GDP per capita (ln)	.017	.010	.007	.000	.008	.008	.004	.003	.015	.015
	(.012)	(.013)	(.010)	(.010)	(.012)	(.010)	(.005)	(.011)	(.013)	(.013)
GDP growth	.003	.001	.005	.010	.000	.012***	.019***	–.001	.012**	.012*
	(.006)	(.006)	.005)	(.005)	(.006)	(.004)	(.007)	(.005)	(.006)	(.006)
Fuel exports (ln)	.005	.001	.005	.021**	.005	.019**	.001	.006	.010	.009
	(.008)	(.009)	(.007)	(.009)	(.008)	(.007)	(.008)	(.007)	(.008)	(.010)
Trade openness (ln)	.125***	.091***	.142***	.146***	.122***	.124***	.132	.157***	.116***	.084***
	(.026)	(.032)	(.022)	(.025)	(.027)	(.021)	(.013)	(.024)	(.025)	(.030)
Memberships in IGOs	.000	–.001	.001	.000	–.001	.000	–.001	.000	.001	–.001
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
ENGO strength (ln)	.039**	.056***	.047***	.047***	.033**	.044***	.058***	.040***	.044**	.021
	(.015)	(.018)	(.013)	(.013)	(.016)	(.012)	(.016)	(.015)	(.016)	(.018)
Climate change vulnerability	–.003	.012	.104*	.278**	.021	.173***	–.055	–.009	.082	.081
	(.068)	(.069)	(.058)	(.107)	(.067)	(.062)	(.073)	(.063)	(.063)	(.076)
Electoral accountability	–.058	–.088*	–.088**	–.028	–.040	–.094***	–.085*	–.043	–.145***	–.033
	(.045)	(.049)	(.038)	(.042)	(.049)	(.034)	(.046)	(.043)	(.022)	(.049)
Horizontal accountability	.00	.045*	.006	–.018	.004	.004	.028	.013	.003	–.016
	(.022)	(.023)	(.020)	(.020)	(.024)	(.018)	(.023)	(.021)	(.022)	(.026)
Political rights	.143***	.146***	.126***	.113***	.201***	.123***	.211***	.151***	.142***	.091*
	(.040)	(.048)	(.034)	(.037)	(.046)	(.032)	(.039)	(.038)	(.038)	(.045)
Civil rights	.073	.097	.156	.269**	–.029	.271***	–.089	.051	.239*	.204
	(.122)	(.135)	(.099)	(.113)	(.123)	(.096)	(.120)	(.109)	(.122)	(.131)
Countries	80	72	67	60	77	74	76	71	60	70
R^2	.726	.736	.797	.843	.739	.815	.730	.789	.784	.577

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** $p < .01$, ** $p < .05$, * $p < .10$. The units of analysis are country averages from 1990–2005.

Table A9. Jackknife analysis: Democracy quality dimensions and CO_2 emissions per capita, ln (OLS regression analysis).

	(1) Central Asia	(2) East Asia and Pacific	(3) East Europe	(4) Latin American	(5) Middle East and North Africa	(6) North America	(7) South Asia	(8) South East Europe	(9) Sub-Saharan Africa	(10) Western European
Population density (ln)	.086**	.142***	.086**	.011	.098**	.098**	.069	.086**	.036	.098**
	(.039)	(.045)	(.040)	(.043)	(.039)	(.040)	(.042)	(.040)	(.042)	(.045)
GDP per capita (ln)	.865***	.866***	.877***	.846***	.810***	.832***	.891***	.887***	.771***	.863***
	(.070)	(.072)	(.072)	(.070)	(.071)	(.070)	(.071)	(.072)	(.100)	(.080)
GDP per capita (ln) squared	–.131***	–.109***	–.128***	–.173***	–.115***	–.136***	–.125***	–.124***	–.096**	–.096**
	(.028)	(.030)	(.029)	(.028)	(.030)	(.028)	(.028)	(.029)	(.041)	(.037)
GDP growth	–.082**	–.109***	–.055	–.093**	–.115***	–.096***	–.105***	–.083**	–.062	–.113***
	(.036)	(.030)	(.040)	(.036)	(.035)	(.035)	(.035)	(.036)	(.042)	(.040)
Fuel exports (ln)	.240***	.260***	.241***	.219***	.274***	.253***	.227***	.254***	.258***	.257***
	(.052)	(.053)	(.054)	(.057)	(.054)	(.051)	(.052)	(.053)	(.060)	(.059)
Trade openness (ln)	.067	.050	.032	–.065	.098	.104	.083	.023	.166	.204
	(.158)	(.178)	(.161)	(.171)	(.156)	(.158)	(.156)	(.162)	(.180)	(.205)
Memberships in IGOs	–.013***	–.016***	–.015***	–.009**	–.015***	–.013***	–.014***	–.015***	–.008	–.010
	(.004)	(.005)	(.005)	(.004)	(.004)	(.004)	(.004)	(.004)	(.005)	(.006)
ENGO strength (ln)	.133	.184**	.100	.193**	.130	.096	.100	.146	.057	.275*
	(.086)	(.091)	(.088)	(088)	(.086)	(.088)	(.093)	(.089)	(.111)	(.101)
Climate change vulnerability	–1.190***	–1.003**	–1.204***	–2.617***	–1.146***	–1.187***	–.886*	–1.152**	–1.169**	–1.218***
	(.420)	(.424)	(.455)	(.699)	(.415)	(.417)	(.447)	(.439)	(.447)	(.453)
Commitment to climate cooperation	.517	.278	.735	.581	.483	.594	.624	.683	.056	.748
	(.438)	(.457)	(.466)	(.452)	(.445)	(.441)	(.435)	(.470)	(.542)	(.504)
Electoral accountability	.360	.364	.337	.450	.514*	.404	.300	.225	.142	.252
	(.268)	(.270)	(.273)	(.283)	(.283)	(.268)	(.279)	(.283)	(.369)	(.293)
Horizontal accountability	–.030	–.057	–.017	.097	–.129	–.013	–.078	–.079	–.027	.006
	(.136)	(.135)	(.140)	(.150)	(.138)	(.135)	(.143)	(.141)	(.161)	(.154)
Political rights	–.233	–.453*	–.211	–.614**	–.060	–.289	–.323	–.189	–.175	–.065
	(.212)	(.237)	(.218)	(.017)	(.260)	(.212)	(.220)	(.219)	(.261)	(.780)
Civil rights	.316	1.168	.355	.265	.260	.373	.843	.469	.672	.065
	(.678)	(.773)	(.685)	(.028)	(.670)	(.669)	(.710)	(.690)	(.917)	(.780)
Countries	98	87	90	78	95	97	94	93	77	83
R^2	.919	.928	.920	.944	.922	.918	.920	.922	.853	.911

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** p < .01, ** p < .05, * p < .10. The units of analysis are country averages from 1990–2010. IGO membership, ENGO strength and commitment refer to country averages from 1990–2005.

5. Cumulative Effects of Democracy Quality Dimensions from 1950–2005/2010

Following Fredriksson and Neumayer (2013), we tested the cumulative effects of the democracy quality dimensions. Table A10 uses the sum of the values of the democracy quality dimensions from 1950 to 2005 (dependent variable: commitment to international climate cooperation)/2010 (dependent variable: CO_2 emissions per capita, ln) as an independent variable. We come to the same conclusions as in our main analysis of the current values of the democracy quality dimensions. The long-term experience with political rights contributes to a commitment to climate cooperation. There are no significant effects of a country's historical experience with electoral and horizontal accountability or political and civil rights on climate policy outcomes.

Table A10. Democracy quality dimensions (cumulative effect from 1950–2010), commitment to climate cooperation and CO_2 emissions per capita, ln (OLS regression analysis).

	(1) Commitment to climate cooperation	(2) Commitment to climate cooperation without influential cases	(3) CO_2 emissions per capita, ln	(4) CO_2 emissions per capita, ln
Population density (ln)	.008	.013*	.080**	.049
	(.010)	(.007)	(.037)	(.035)
GDP per capita (ln)	–.002	.008	.868***	.927***
	(.016)	(.011)	(.063)	(.060)
GDP per capita (ln) squared			–.143***	–.116***
			(.031)	(.027)
GDP growth	.006	.002	–.096***	–.074**
	(.007)	(.005)	(.035)	(.031)
Fuel exports (ln)	–.008	.003	.274***	.210***
	(.012)	(.008)	(.049)	(.046)
Trade openness (ln)	.091**	.129	.168	.256**
	(.035)	(.025)	(.147)	(.060)
Memberships in IGOs	.000	.000**	–.015***	–.009**
	(.001)	(.001)	(.004)	(.004)
ENGO strength (ln)	.036*	.031	.186**	.234***
	(.021)	(.015)	(.083)	(.071)
Climate change vulnerability	–.029	.026	–1.265***	–1.350***
	(.099)	(.065)	(.410)	(.373)
Commitment to climate cooperation			.249	.146
			(.391)	(.349)
Electoral accountability	–.001	–.004	.001	.000
	(.004)	(.003)	(.004)	(.004)
Horizontal accountability	.001	.001	.001	–.002
	(.002)	(.001)	(.003)	(.002)
Political rights	.007**	.010***	–.003	–.003
	(.003)	(.002)	(.003)	(.003)
Civil rights	.009	.002	.009	.004
	(.010)	(.007)	(.008)	(.007)
Countries	99	78	98	87
R^2	.452	.726	.925	.954

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** $p < .01$, ** $p < .05$, * $p < .10$. The units of analysis are country averages from 1990–2005 (Model 1–2) and from 1990–2010 (Model 3–4). IGO membership, ENGO strength and commitment refer to country averages from 1990–2005.

6. Main Model of Commitment to Climate Cooperation with Interaction Effect between ENGO Strength and Political Rights

We tested an interaction effect between ENGO strength and political rights (see table A11). The assumption is that political rights enable citizens to organize in ENGOs and exert influence via public opinion policy outputs. The results indicate no significant interaction effect between ENGO strength and political rights. It has to be considered that the number of ENGOs captures only the number of ENGO's in a country. However, alternative measures were only available for a limited number of countries.

Table A11. Interaction effect between ENGO strength and political liberties (OLS regression analysis).

	(1) Commitment to climate cooperation	(2) CO_2 emissions per capita, ln
Population density (ln)	.007	.088**
	(.010)	(.040)
GDP per capita (ln)	.000	.853***
	(.017)	(.070)
GDP per capita (ln) squared		–.132***
		(.030)
GDP growth	.005	.090**
	(.007)	(.088)
Fuel exports (ln)	–.008	.250***
	(.012)	(.052)
Trade openness (ln)	.086**	.082
	(.036)	(.159)
Memberships in IGOs	.000	–.014***
	(.001)	(.004)
ENGO strength (ln)	.036*	.141
	(.021)	(.086)
Climate change vulnerability	–.035	–1.223***
	(.100)	(.421)
Commitment to climate cooperation		.547
		(.442)
Electoral accountability	–.013	.344
	(.100)	(.269)
Horizontal accountability	.010	–.030
	(.032)	(.030)
Political rights	.100*	–.236
	(.050)	(.216)
Civil rights	–.020	.377
	(.029)	(.680)
ENGOlnXpolitical rights	–.020	.045
	(.029)	(.127)
Countries	99	99
R^2	.455	.918

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** p < .01, ** p < .05, * p < .10. The units of analysis are country averages of from 1990–2005 (Model 1) and from 1990–2010 (Model 2). IGO membership, ENGO strength and commitment refer to country averages from 1990–2005.

7. Control of Annex-I Status

International climate change regime distinguishes between Annex-I and non-Annex-I countries. Annex-I (Developed) countries are regarded as historically responsible for global warming and, therefore, should take the lead in climate change mitigation. Thereby, the Kyoto Protocol specified legally binding greenhouse gas emissions targets for Annex-I member states. Table A12 controls the effect of the Annex-I status on commitment to climate cooperation. Overall, our results remain stable.

Table A12. Annex-I status and commitment to climate cooperation (OLS regression analysis).

	(1) Commitment to climate cooperation
Population density (ln)	.007
	(.010)
GDP per capita (ln)	–.006
	(.018)
GDP per capita (ln) squared	.007
	(.008)
GDP growth	.007
	(.008)
Fuel exports (ln)	–.007
	(.012)
Trade openness (ln)	.092**
	(.035)
Memberships in IGOs	.000
	(.000)
ENGO strength (ln)	.038*
	(.021)
Climate change vulnerability	.003
	(.115)
Electoral accountability	–.007
	(.060)
Horizontal accountability	.009
	(.032)
Political rights	.102**
	(.050)
Civil rights	.12
	(.159)
Annex I country	.026
	(.048)
Countries	99
R^2	.454

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** $p < .01$, ** $p < .05$, * $p < .10$. The units of analysis are country averages of from 1990–2005.

8. Control of CO_2 Emissions Per Capita

Table A13 adds CO_2 emissions per capita, ln, as an additional control variable in our analysis of commitment to climate change cooperation. Countries with low emission levels might be more willing to enter international climate change treaties than countries with high emission levels. The results of our main model in Table 2 remain stable. In contrast with our expectations, we find a positive but insignificant effect of CO_2 emission per capita on commitment to climate cooperation.

Table A13. Democracy quality dimensions and commitment to climate cooperation (OLS regression analysis).

	(1) Commitment to climate cooperation
Population density (ln)	.005
	(.010)
GDP per capita (ln)	–.025
	(.026)
GDP growth	.008
	(.008)
Fuel exports (ln)	–.015
	(.014)
Trade openness (ln)	.084**
	(.036)
Memberships in IGOs	.000
	(.000)
ENGO strength (ln)	.031
	(.021)
Climate change vulnerability	–.004
	(.102)
Electoral accountability	–.026
	(.062)
Horizontal accountability	.017
	(.032)
Political rights	.107**
	(.049)
Civil rights	.135
	(.158)
CO_2 emissions per capita, ln	.026
	(.024)
Countries	99
R^2	.460

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** $p < .01$, ** $p < .05$, * $p < .10$. The units of analysis are country averages of from 1990–2005.

9. Democracy Quality Dimensions and Average CO_2 Emissions in Two Sub-Periods (1990–1999, 2000–2010)

We also estimated our climate policy outcome model for two sub-periods—1990–1999 and 2000–2010. Table A13 shows, as in our main analysis, no significant effects of the democracy quality dimensions on CO_2 emissions per capita. Model 2 and 4 in table A13 show that the significant positive effect of electoral accountability in the analysis of our sample without countries from the Middle East and Northern Africa is only stable for the period from 1990–1999.

Table A14. Democracy quality dimensions and CO_2 emissions per capita, ln (OLS regression analysis).

	(1) 1990–1999	(2) 1990–1999 Without countries from Middle East and North Africa	(3) 2000–2010	(4) 2000–2010 Without countries from Middle East and North Africa
Population density (ln)	.077*	.084*	.093**	.098**
	(.046)	(.046)	(.045)	(.046)
GDP per capita (ln)	.861***	.858***	.942***	.927***
	(.074)	(.074)	(.074)	(.078)
GDP per capita (ln) squared	–.114***	–.098***	–.137***	–.129***
	(.032)	(.033)	(.031)	(.033)
GDP growth	–.094***	–.096***	.008	–.002
	(.023)	(.023)	(.043)	(.049)
Fuel exports (ln)	.185***	.213***	.216***	.233***
	(.056)	(.056)	(.057)	(.060)
Trade openness (ln)	.066	.065	.022	–.017
	(.158)	(.157)	(.169)	(.172)
Memberships in IGOs	–.016***	–.018***	–.011**	–.013**
	(.004)	(.005)	(.005)	(.005)
ENGO strength (ln)	.166*	.144	.106	.099
	(.098)	(.099)	(.083)	(.085)
Climate change vulnerability	–1.105***	–1.050**	–1.052**	–1.046**
	(.463)	(.460)	(.455)	(.466)
Commitment to climate cooperation	.866	.728	.736	.659
	(.527)	(.524)	(.469)	(.480)
Electoral accountability	.361	.486*	.061	.150
	(.137)	(.281)	(.259)	(.278)
Horizontal accountability	–.135	–.218	–.016	–.046
	(.137)	(.141)	(.143)	(.147)
Political rights	–.150	.053	–.333	–.284**
	(.232)	(.246)	(.241)	(.251)
Civil rights	.532	.220	.811	.773
	(.680)	(.688)	(.781)	(.792)
Countries	90	87	95	92
R^2	.914	.919	.907	.907

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** $p < .01$, ** $p < .05$, * $p < .10$. The units of analysis are country averages of from 1990–1999 (Model 1–2) and from 2000–2010 (Model 3–4). IGO membership, ENGO strength and commitment refer to country averages from 1990–2005.

10. Democracy Quality Dimensions and Changes in CO_2 Emissions over Time

10.1. Dependent Variable: Linear Trend in CO_2 Emissions Per Capita from 1990 to 2010

Our main analysis focuses on variation in CO_2 emissions per capita among countries. Table A15 tests our model on the variation of long-term changes in CO_2 emissions among countries. For this purpose, we estimated for each country the slope (unstandardized regression coefficient) of our year variable on CO_2 emissions per capita from 1990–2010. The slopes of each country constitute our measure of long-term changes in CO_2 emissions per capita in the cross-sectional OLS regression analysis (slope regression) in Table A15 (Babones, 2014). We find no significant effect of the democracy quality dimensions on the linear trend of CO_2 emissions per capita.

Table A15. Democracy quality dimensions and CO_2 emissions per capita, ln long-term changes (OLS regression analysis/slope regression).

	(1)
Population density (ln)	–.003
	(.011)
GDP per capita (ln)	.030
	(.020)
GDP per capita (ln) squared	.001
	(.008)
GDP growth	.027***
	(.010)
Fuel exports (ln)	.020
	(.015)
Trade openness (ln)	–.085*
	.045)
Memberships in IGOs	–.001
	(.001)
ENGO strength (ln)	–.049*
	(.025)
Climate change vulnerability	.343***
	(.120)
Commitment to climate cooperation	.033
	(.125)
Electoral accountability	.030
	(.077)
Horizontal accountability	.018
	(.039)
Political rights	.048
	(.061)
Civil rights	–.164
	(.193)
Countries	99
R^2	.266

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** $p < .01$, ** $p < .05$, * $p < .10$. The units of analysis are country averages of from 1990–2010. IGO membership, ENGO strength and commitment refer to country averages from 1990–2005. The dependent variable is the linear trend of CO_2 emissions per capita from 1990–2010.

10.2. Dependent Variable: Annual Changes in CO_2 Emissions Per Capita

We estimated pooled OLS regression models with annual changes of CO_2 emission per capita as a dependent variable (first differences) with country and year dummies. As we have only annual data on state memberships in IGOs and the number of ENGOs in a country until 2005, these models only analyze the CO_2 emission changes from 1990–2005. Table A15 finds no effect of the institutional traits of democracy on annual changes in CO_2 emission per capita.

Table A16. Democracy quality dimensions and CO_2 emissions per capita, ln short-term changes (OLS regression analysis/slope regression).

	(1)
Population density (ln)	–.039
	(.081)
GDP per capita (ln)	–.005
	(.043)
GDP per capita (ln) squared	–.005
	(.010)
GDP growth	.006***
	(.001)
Fuel exports (ln)	.000
	(.008)
Trade openness (ln)	.000
	(.000)
Memberships in IGOs	–.000
	(.001)
ENGO strength (ln)	.002
	(.012)
Electoral accountability	.008
	(.018)
Horizontal accountability	.012
	(.017)
Political rights	–.034
	(.029)
Civil rights	.050
	(.0830)
Observations	1160

Notes: Unstandardized regression coefficients and standard errors in parentheses. *** $p < .01$, ** $p < .05$, * $p < .10$.

References

Babones, S. J. (2014). *Methods for quantitative macro-comparative research*. Los Angeles, CA: Sage.

Coppedge, M., Gerring, J., Lindberg, S., Skaaning, S.-E., Teorell, J., Altman, D., . . . Wilson, S. (2017). V-Dem [country-year/country-date] dataset v7.1. *Varieties of Democracy (V-Dem) Project*. Retrieved from https://www.v-dem.net/en/data/data-version-7-1

Fredriksson, P. G., & Neumayer, E. (2013). Democracy and climate change policies. Is history important? *Ecological Economics*, *95*, 11–19.

Marshall, M. G., Gurr, T. R., & Jaggers, K. (2016). Polity IV project, political regime characteristics and transitions, 1800–2015. *Center for Systemic Peace*. Retrieved from http://www.systemicpeace.org/inscrdata.html

Merkel, W. (2004). Embedded and defective democracies. *Democratization*, *11*(5), 33–58.

Pemstein, D., Marquardt, K. L., Tzelgov, E., Wang, Y., Krusell, J., & Miri, F. (2017). *The V-Dem measurement model: Latent variable analysis for cross-national and cross-temporal expert-coded data* (Working Paper no. 21, 2nd ed.). Gothenburg: Varieties of Democracy Institute.

Politics and Governance (ISSN: 2183-2463)

Politics and Governance is an innovative new offering to the world of online publishing in the Political Sciences. An internationally peer-reviewed open access journal, Politics and Governance publishes significant, cutting-edge and multidisciplinary research drawn from all areas of Political Science.

www.cogitatiopress.com/politicsandgovernance

COGITATIO