

## Seminar Paper

# Measurement Approaches of Direct Democracy

A Cross-National Theoretical and Empirical Examination

Author: Fabio Votta, B.A.  
Email: fabio.votta@gmail.com  
Student ID: 2891518

Author: Rosa Seitz, B.A.  
Email: rosa.marie.seitz@gmail.com  
Student ID: 2876533

Date of Submission: 04.04.2018

### Abstract

This seminar paper aims to systematically compare selected approaches to the measurement of direct democracy on the national level in a theoretical and empirical way, as well as to examine data sources on direct democratic institutions in regard to their usability for quantitative research. It has to be noted, that this paper does not aim at giving a complete overview of the literature in terms of operationalization and measurement. Instead, only a few contemporary approaches are selected and addressed more thoroughly. In many studies, direct democracy is operationalized with a single dummy variable that specifies the existence of one specific institution. Furthermore, while the study of subnational direct democratic institutions is a worthwhile task, we primarily focus on national level measurements of direct democracy for the sake of a common thread across this work.

Seminar: Representative, direct and cooperative participation in comparison

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Challenges in Measuring Direct Democracy</b>	<b>5</b>
<b>3</b>	<b>Direct Democracy Data Sources</b>	<b>8</b>
3.1	Varieties of Democracy - Direct Democracy variables . . . . .	9
3.2	Direct Democracy Database (IDEA) . . . . .	10
3.3	Database and Search Engine for Direct Democracy (sudd) . . . . .	11
3.4	The Navigator of Direct Democracy . . . . .	11
<b>4</b>	<b>Direct Democracy Measurement Approaches</b>	<b>13</b>
4.1	Approaches assessing Rules in Form . . . . .	13
4.1.1	Gherghina 2016 . . . . .	13
4.1.2	Peters 2016 . . . . .	14
4.1.3	Democracy Barometer . . . . .	17
4.2	Approaches assessing Rules in Use . . . . .	18
4.3	Approaches assessing Rules in Form and Rules in Use . . . . .	19
4.4	Summary of Measurement Approaches . . . . .	21
<b>5</b>	<b>Data Wrangling &amp; Empirical Comparison</b>	<b>23</b>
5.1	Data Wrangling . . . . .	23
5.1.1	Construction of Rules in Form Measures . . . . .	23
5.1.2	Construction of Rules in Use Measures . . . . .	26
5.1.3	Mixed Measures and Additional Variables . . . . .	30
5.2	Comparing Rules in Form Measures . . . . .	31
5.3	Comparing Rules in Use Measures . . . . .	34
5.4	Comparing Mixed Measures . . . . .	37
<b>6</b>	<b>Conclusion and Future Research</b>	<b>42</b>
	<b>References</b>	<b>44</b>
	<b>Appendix</b>	<b>46</b>

## List of Figures

1	Gherghina RiF compared to other Rules in Form Measures . . . . .	31
2	Peters RiF compared to other Rules in Form Measures . . . . .	32
3	Rules in Form Measures by Freedom House Index . . . . .	32
4	Bottom-Up and Top-Down Rules in Form Measures . . . . .	33
5	Rules in Form Measures - World Map . . . . .	34
6	Gherghina - RiU compared to other Rules in Use Measures . . . . .	35
7	Rules in Use Measures by Freedom House Index . . . . .	35
8	Bottom-Up and Top-Down Rules in Use Measures . . . . .	36
9	Rules in Use Measures - World Map . . . . .	37
10	Scatterplot - DDI vs. DPVI . . . . .	38
11	Mixed Measures by Freedom House Index . . . . .	39
12	Direct Democracy Index - World Map . . . . .	40
13	Direct Popular Vote Index - World Map . . . . .	40
14	Bottom-Up and Top-Down Components of DPVI . . . . .	41
A1	Heatmap of all Variable Used . . . . .	46

## List of Tables

1	Definition of Direct Democratic Mechanisms . . . . .	6
2	Overview of the Selected Datasets . . . . .	8
3	Overview of Used Direct Democracy Measures . . . . .	15
4	Rules in Form . . . . .	23
5	Rules in Use . . . . .	26
6	V-Dem Variables used for Gherghina - RiU Indicator Construction . . . . .	27
7	Mixed Indices . . . . .	30
A1	Description of Direct Democracy Index - Kaufmann . . . . .	47

# 1 Introduction

Direct democracy has been a topic of interest to political theory and empirical research for a long time. It is often asserted that direct democracy is closest to the ideal form of democracy, which is rule of the people and by the people themselves. In this light, direct democratic institutions are an often discussed remedy of the contemporary “crisis” of representative democracy which is postulated by some scholars (cf. Pogrebinschi 2015). Therefore, the causes and effects of direct democracy are of great interest to empirical research. For example, studies indicate positive effects on political efficacy (cf. Bernard & Bühlmann 2014) and satisfaction with democracy (cf. Stadelmann-Steffen & Vatter 2012). However, other authors have pointed out that direct democracy could exclude minorities and might foster inequalities in political participation (cf. Merkel & Ritzi 2017: 241). An examination of the literature shows that effects of direct democracy are sometimes heavily disputed, empirical research does not always provide conclusive evidence and many questions remain unanswered. In this context, quantitative comparative research has assigned great importance to the question of how to measure direct democracy, as empirical results could rely heavily on the way direct democracy is operationalized. In the literature, different approaches to this question can be identified, having some elements in common but diverging in many ways.

The aim of this work is to systematically compare selected approaches to the measurement of direct democracy on the national level in a theoretical and empirical way, as well as to examine data sources on direct democratic institutions in regard to their usability for quantitative research. It has to be noted, that this paper does not aim at giving a complete overview of the literature in terms of operationalization and measurement. Instead, only a few contemporary approaches are selected and addressed more thoroughly. In many studies, direct democracy is operationalized with a single dummy variable that specifies the existence of one specific institution.<sup>1</sup> Such studies were not examined in this paper, as the focus here lies on measurements that account for the multidimensionality of direct democracy, which cannot be captured with a simple dummy variable. Furthermore, while the study of subnational direct democratic institutions is a worthwhile task, we primarily focus on national level measurements of direct democracy for the sake of a common thread across this work. As most of the indicators are derived from studies restricted to democracies, the *democracy* dimension in direct democracy is neglected in most of them. This can be seen as a shortcoming, as some autocratic or semi-democratic regimes have adopted direct popular vote instruments, using them to seemingly legitimate their rule. For a first exploratory application of the constructed measures, we also assess the indices in regard to their Freedom House classification. The approaches examined theoretically and empirically are derived from different studies or datasets, namely: Gherghina (2016), a subcomponent of the Democracy Barometer Barometer (2016), Peters (2016), Fiorino et al. (2017), and Altman (2017) / M. Coppedge et al. (2017).

<sup>1</sup>For example a citizens’ initiative dummy as in Bernard & Bühlmann (2014).

The following Chapter 2 discusses some of the general challenges in measuring direct democracy as well as criteria relevant to the measurement of direct democracy (like the decisiveness or ease of initiation/approval). In Chapter 3, some of the available data sources on direct democracy are briefly examined in regard to their content and scope, as well as their usability for quantitative research (the *IDEA-Database*, the “*Varieties of Democracy*”-Dataset, the *Database and Search Engine for Direct Democracy* and the *Direct Democracy Navigator*). Afterwards, Chapter 4 introduces some of the contemporary approaches for the measurement of direct democracy and compares them to the previously established criteria. After reconstructing some indicators from the literature, Chapter 5 examines some of the discussed approaches empirically. Finally, the conclusion (Chapter 6) gives an overview of the examined measurements and their empirical comparison and derives implications for further research.

## 2 Challenges in Measuring Direct Democracy

Before discussing operationalization and measurement of direct democracy, it is necessary to determine a working definition of direct democracy first. Generally speaking, direct democracy is a form of government in which collective decisions are made by the people themselves and not by their respective delegates, as would be the case in a representative democracy (cf. Clark et al. 2017: 499). Despite this rather straightforward definition, it is not an easy task to assess direct democracy empirically. One challenge arises from the fact that all democracies rely to some degree on delegation, as the government processes in modern states are too complex to organize them in a completely direct democratic way. Therefore, countries can be more or less direct democratic, depending on their institutional designs. Moreover, there are different institutional arrangements that can be classified as direct democratic mechanisms, tools or institutions, for example mandatory and optional referendums or citizens’ initiatives. Especially challenging in this context is the rather unclear terminology of direct democratic institutions. As Altman puts it, “what we understand as direct democracy has different meanings in different places, and the different institutional components of this concept [...] have diverse normative undertones. For instance, a referendum in one country is called a plebiscite or even a popular initiative in another” (Altman 2017: 1208). For the sake of clarity, we define six direct democratic mechanisms, which can be mapped onto the datasets and measurement approaches that will be discussed later on. The definitions can be found in Table 1, and are primarily derived from the V-Dem Codebook (Coppedge Michael et al. 2017: 138), and additionally Beramendi et al. (2008) pp. 9-14.<sup>2</sup>

---

<sup>2</sup>Another notable typology, which allows for a more refined differentiation and is shortly introduced in Chapter 2, is provided by the Democracy Navigator (cf. Navigator 2018).

Table 1: Definition of Direct Democratic Mechanisms

<b>Mechanism</b>	<b>Concerning</b>	<b>Initiator</b>	<b>Level</b>
<i>Recall</i>	Recall of elected officials	Citizens	Bottom-Up
<i>Agenda Initiatives</i>	Proposals to legislative body	Citizens	Bottom-Up
<i>Citizen Initiatives</i>	Ballot Proposals	Citizens	Bottom-Up
<i>Facultative Referendum</i>	Rejection of law	Citizens	Bottom-Up
<i>Plebiscite</i>	Unspecified	Authorities	Top-Down
<i>Obligatory Referendum</i>	Matters specified in constitution	Constitution	Top-Down

Adapted from: Altman (2017) p. 138; Beramendi et al. (2008) pp. 9-14

We categorize the direct democratic institutions by two criteria:

- 1.) *who or what* initiated the direct democratic institution
- 2.) *which topic/issue* they concern (e. g. proposals or rejection of law)

Nevertheless, it is sometimes difficult to follow our defined terminology, as for example when referring to optional referendums, some authors do not specify which actor can initiate them, therefore it is unclear which specific institution the respective source refers to. Such issues will also be addressed in the data wrangling part, as data sources differ in their taxonomy as well.

In empirical research, there are different approaches in assessing direct democracy on the national level. The examination of the selected measurement approaches is structured according to criteria which are explicitly or implicitly included. The approaches differ in regard to which democratic mechanisms they assess, whether they capture those mechanisms only in their form or in their use (or both), differentiate between bottom-up and top-down direct democracy, include hurdles/easiness as well as the decisiveness of the mechanisms and whether they account for subnational levels. Not all indices include or consider all the criteria, which can be explained by different data sources as well as conceptions of direct democracy. First, the examined measurements differ in regards to which *mechanisms of direct democracy* they cover. Most of the examined indices are in principle based on whether a certain set of direct democratic institutions exists or not. The variety of institutions referred to mostly consists of the ones defined in Table 1 (*obligatory referendums*, *plebiscites*, *facultative referendums*, *citizen' initiatives*, *agenda initiatives* and *recall*).

An important differentiation between the institutional availability of direct democratic tools (*rules in form*) and the actual use of these institutions (*rules in use*) is emphasised for example by Bauer & Fatke (2014), Gherghina (2016), Blume et al. (2007) and Peters (2016). The relevance of this distinction lies in the possibly different causes and effects of the formal existence of direct democratic

institutions and their actual use. For example, Bauer and Fatke argue that the institutional availability of direct democracy should be positively related to trust in authorities, whereas the actual use of such mechanisms increases distrust (cf. Bauer & Fatke 2014: 53f).

Another crucial distinction is the one between *bottom-up* and *top-down* mechanisms of direct democracy (cf. Peters 2016; Altman 2017). Here, the core criteria is whether the direct democratic mechanisms are initiated by the citizens (bottom-up) or by constitutional or state organs, for example government or parliament (top-down). The theoretical importance of the difference between top-down and citizen initiated mechanisms is once again rooted in possibly different assumptions about the effects of direct democracy (see for example Peters 2016). Besides the differentiation between citizen-initiated and top-down mechanisms, some measurement approaches take into account whether the initiator is a veto-player and also consider the writer of the ballot proposal (cf. Merkel et al. 2016: 47; see also Navigator 2018).

Another criterion which is captured in some approaches is the “*easiness*” of direct democratic institutions, once they are available. This criterion captures the relevance of thresholds or quora for initiation and/or approval of direct democratic mechanisms. The term is adopted from Altman, who explicitly distinguishes between ease of approval and ease of initiation, but other sources account for the easiness as well (cf. Altman 2017; Fiorino et al. 2017; Gherghina 2016).

In addition to or sometimes without consideration of easiness, the “*decisiveness*” of direct democratic mechanisms plays an important role. Decisiveness refers to whether the outcomes of direct democratic mechanisms are binding or merely consultative. This is important for some aggregation methods, as they apply a weighting that is based on the assumption that actually binding institutions are more direct democratic than merely consultative ones (cf. Kaufmann 2004; Fiorino et al. 2017; Peters 2016).

Lastly, two important dimensions that are neglected in most of the measurement approaches discussed in this paper will be briefly mentioned. A rarely considered dimension is *subnational direct democracy* (for example Gherghina 2016), although research that also considers variations in direct democracy on subnational levels could be of importance. Moreover, one could argue that a country which provides direct democratic institutions at least at the subnational levels is more direct democratic than a country that does not allow for any direct democratic processes at all, and thus neglecting this dimension might lead to suboptimal results. In this paper, we only focus on the national level for the sake of brevity, although still emphasizing that this is an important dimension that should be studied more intensely. A different dimension rarely covered, but nevertheless important is the scope and content of *issues* for which direct democratic mechanisms are allowed or restricted. After discussing the dimensions which are often captured within measurements of direct democracy, we examine selected data sources that researchers can rely on in order to gather information on provisions as well as actual use of direct democratic instruments.



### 3 Direct Democracy Data Sources

Whenever the effects of direct democracy are meant to be evaluated by researchers, they first need to gather (systematic) data that expresses the degree of direct democracy. Before the measurement approaches and their indicator construction is discussed, we therefore describe data sources on direct democratic institutions that allow for cross-national comparison. Historically, data on direct democracy was mostly available for two regions: Switzerland and the individual states of the United States of America. It wasn't until Butler and Ranney's attempt at collecting information on referendums from all over the world that measuring direct democracy became comparable across a wide variety of countries (cf. Butler & Ranney 1978). Since then, established worldwide democracy indices (for example Polity IV or Freedom House) have been fairly neglectant on the topic of direct democracy and do not include such variables in their measurements.

Table 2: Overview of the Selected Datasets

Dataset	Level	Types	Available Information	Subnationality	Scope
IDEA	Country-Level	Mandatory and Optional referendum, Citizens' Initiative, Agenda Initiative, Recall	Primarily Rules in Form		
			Top-down/bottom-up not separated	yes, but limited information	198 Countries, up-to-date
			Information on different quora/ thresholds and bindingness		
			Data preparation necessary		
V-Dem	Country-Year	Obligatory Referendum, Plebiscite, (Facultative) Referendum, Citizens' Initiative	Rules in Form and Use		
			Top-down/bottom-up can be differentiated	yes, but limited information	178 Countries, 1900 - 2016
			Information on different quora/ thresholds, and bindingness		
			Data can be used for quantitative research without much preparation		
sudd	Referendum-Level	Inofficial Vote, Plebiscite, Obligatory Referendum, Facultative Referendum, Initiative	Primarily Rules in Use		
			Top-down/bottom-up can be differentiated	yes	140 Countries on national level, 1791 - present
			Information on bindingness, quora and thresholds depend on each popular vote		
			Data preparation necessary		
Direct Democracy Navigator	Country-Level	Agenda Initiative, Initiative, Referendum, Plebiscite  (with subtypes for the latter three)	Top-down/bottom-up can be differentiated		
			Information initiating authority and on the author(s) of the ballot proposal(s)	yes	112 Countries, with their regions and municipalities
			Information on bindingness of plebiscites, quora and thresholds not available		
			Data preparation necessary, incomplete data		up-to-date, but incomplete

Other individual researchers have been more focused on specific regions, for example Latin America (cf. Lissidini 2011: 60–7; Zovatto 2015; Madroñal 2005), Asia (cf. Hwang 2005) and Central and Eastern Europe (cf. Auer & Bützer 2001).

The following section introduces three datasets that include institutional provisions and/or usage of direct democracy for countries around the world: *Varieties of Democracy* (short: V-Dem), the *Direct Democracy Database* by the Institute for Democracy and Electoral Assistance (short: IDEA), and the *Democracy Navigator* operated by the Research Center of Citizen Participation/Institute for Democracy And Participation Research of the University of Wuppertal.<sup>3</sup> For an extensive rules in use database, we rely on the *Database and Search Engine for Direct Democracy* (short: sudd) compiled by Beat Müller, which was translated from German into English by Research and Documentation Centre on Direct Democracy. The *Democracy Barometer* is not discussed at this point, as it only includes superficial data for provision and use of direct institutions (compared to the other data sources), namely a total of three already aggregated indicators.

Before delving deeper into the separate datasets, it would be wise to identify certain common dimensions along which they can be classified. For example, the examined datasets differ in their scope (temporal/ geographical), in terms of which direct democratic mechanisms they provide data for and whether it can be used to account for dimensions like top-down/bottom-up, easiness or decisiveness. Also, the datasets vary in whether they are composed of cross-sectional or time-series country-level data or are based on individual referendums, which is especially relevant for the distinction of rules in form and rules in use. All of the discussed features imply that the data sources might be more or less suitable for researchers, depending on the desired operationalization of direct democracy. Table 2 gives a short overview of the selected datasets.

### 3.1 Varieties of Democracy - Direct Democracy variables

The most comprehensive and systematic approach might be included in the Varieties of Democracy dataset (v7). It includes data in the range from 1900 to 2016 for a total of 178 countries (Coppedge Michael et al. 2017: 137–53). The codebook specifically addresses the top-down and bottom-up differentiation in their typology of direct democratic institutions, which makes comparison between these dimensions convenient. The covered top-down institutions are obligatory referendums and plebiscites, initiatives and (facultative) referendums are featured as bottom-up mechanisms. The V-Dem data is in a country-year format so that it provides the legal provisions for each institution in a given year. Concerning rules in use, V-Dem includes a variable for each institution, capturing the number of occurrences in a specific year. Moreover, a variable is included which indicates if there

<sup>3</sup>A different database that would also have been worthy of including in our comparison is compiled by the Research and Documentation Centre on Direct Democracy (*c2d*), however at the time of writing this paper, the website containing the data has been offline and thus inaccessible (as of February and March 2018).

was any “credible” use of a direct popular vote institution in a given year (by credible it is referred to whether the official results reflect the actual vote). The V-Dem data refers almost completely to the national level and does not include any subnational forms of direct democracy, besides one item that asks whether the respective institution exists at national, subnational, or at both levels (except for obligatory referendums). Besides information about the existence of institutions, it is assessed whether they are binding and whether certain quora apply (approval and turnout quorum, administration threshold and supermajority requirements for every institution as well as signature requirements and gathering periods for the bottom-up institutions).

### 3.2 Direct Democracy Database (IDEA)

The direct democracy database by *IDEA* can be seen as collection of information rather than a dataset that is meant to be used in quantitative analysis. It does not have a specified temporal range, instead it simply features the most recent available information. It includes several columns assessing provisions of direct democracy in a total of 198 countries and mostly focuses on rules in form.<sup>4</sup> The IDEA database covers legal provisions for five institutions: *mandatory/obligatory referendums*, *optional referendums*, *citizens’ initiatives*, *agenda initiatives* and *recall*. This information is available for 198 countries also includes regional and local levels, though only general information about the existence of such provisions is provided. A greater range of information on the respective institutions is only available on the national level. For referendums, the information consists of restricted/allowed issues, possible initiators, who drafts the referendum question, who decides the final form of the ballot text, approval quorum, majority requirements and bindingness. A disadvantage in the structure of the IDEA data is that its underlying typology of institutions does not differentiate between citizen initiated optional referendums and top-down optional referendums/plebiscites (though extraction is possible and is performed in Chapter 5). Moreover optional and mandatory referendums are mixed up in the same column, resulting in strings like “Mandatory referendum - always/ Optional referendum - sometimes”. This makes it challenging to use the data in quantitative analysis, especially if one is interested in data for a large number of countries. Regarding initiatives, the same mixed-up data structure applies for agenda and citizen’ initiatives. For both initiatives information is available for issues voted on, required materials, disqualification, legality checks, the author of the initiatives title and ballot text. For recall institutions, informations are provided in similar fashion. Regarding signature collection, detailed information is available, for example on time gathering periods and requirements, although different institutions are again captured in the same variable (this time optional referendums, agenda initiatives, citizen initiatives and recall), making it difficult to extract the information for quantitative analysis. In general, this coding induces uncertainty whether some entries that lack further specification refer to all of the provided institutions or not.

<sup>4</sup>IDEA only addresses rules in use by including information on the first referendum or initiative held in a specific country and whether a national referendum was held since 1980.

### 3.3 Database and Search Engine for Direct Democracy (sudd)

Similarly to the direct democracy database by IDEA, the *Database and Search Engine for Direct Democracy* (sudd) is not inherently designed to be used in quantitative analyses. However, it includes a wide range of data on all kinds of actual direct popular votes from 1791 to the current day, making it quite valuable for researchers seeking to explore direct democratic mechanisms over time. The unit of analysis for the sudd data is any direct popular vote that occurred in any given year in any given country or region (autonomous/independent/unrecognized or colonized regions, does not refer to subnational level, eg. Swiss Cantons). This leads to a total of 2815 referendums in 400 countries and regions. Given the structure of the data, rules in form are not measurable (or only per referendum), making it mostly useful for research assessing rules in use. For each popular vote, different variables are captured, for example question type, majority requirements, number of people entitled to vote and the results of the vote. Also majority requirements and sometimes participation or approval quora are documented. Rather useful information is captured in the column *type of vote*, which categorizes amongst others the type of vote (*unofficial*, *plebiscite*, *obligatory referendum*, *facultative referendum* and *initiative*), who initiated it (for example parliament, president or the people) and the bindingness. Again, if one wishes to use the data for quantitative analysis, the information has to be extracted with some effort. The coding scheme makes it challenging to use without some extensive data cleaning and transforming and, given that the data is not available for download, it has to be web scraped from the website itself.<sup>5</sup>

### 3.4 The Navigator of Direct Democracy

A noteworthy data compilation is the *Direct Democracy Navigator* database, although it is still under development. The aim of the Democracy Navigator is to feature available legal designs, practices and events of direct democracy in all jurisdictions around the world. In their classification typology, it also accounts for the (veto player or minority) status of a top-down initiator as well as the author of the ballot proposal(s) (cf. Navigator 2018), which results in nine different subtypes of direct democratic votes, with three overarching main types:

- 1.) *initiative* (citizens' initiative, citizens' initiative + authorities' counter-proposal, agenda setting initiative)
- 2.) *referendum* (citizen-initiated referendum, citizen-initiated referendum + counter-proposal, obligatory referendum)
- 3.) *plebiscites* (plebiscite, veto-plebiscite, authorities' minority plebiscite, authorities' minority veto-plebiscite)

---

<sup>5</sup>On request, Beat Müller kindly offered to provide the Data in XML format after restructuring, which would possibly have made data preparation somewhat easier, however such data was not accessible before the end of March 2018. Nevertheless there are some issues which will be addressed in the empirical section, when discussing the construction of rules in use indicators.

Unfortunately, the data available for legal provisions seems to be rather incomplete and is not available as a dataset, even on the national level (thus it needs to be web scraped). Moreover, it is not clear at all whether not-listed legal provisions are actually not provided or rather missing values. Furthermore, sometimes types of direct democratic votes are counted twice without obvious reasons (for example the agenda initiative in Liechtenstein) or because they are separately regulated for different issues (for example plebiscites in Taiwan). All of this makes it difficult to use for quantitative comparative analyses, although the database is useful for cross-checking information with other sources.

In general, there is no recommendation as to which of the data sources a quantitative researcher should use in their empirical analysis. This decision depends heavily on the research question to be answered and the underlying concept of direct democracy. For example, many studies exclude the recall mechanism from their measurement, because it can be seen as a accountability function of representative democracy and not an element of direct democracy. If one is interested in examining provisions for institutions in regard to whether they are top-down or bottom up initiated, a researcher is better advised in using the V-Dem data, as it takes much more effort to obtain the information from other sources such as IDEA. Generally speaking, of the discussed data sources V-Dem is the most convenient and user-friendly dataset to use in quantitative analyses, mostly numeric in nature and available in time-series format since 1900, with the drawback of some unavailable information (for example on agenda initiatives or the authors of ballot proposals). On the other hand, the IDEA Database offers more detailed documentation of legal provisions and the corresponding constitutional paragraphs are often provided as background information, which makes it a valuable data source as well. The Democracy Navigator is a rather detailed database as well, which is especially useful for cross-checking information or to gain additional information if a researcher is interested in legal designs in regard to the status of the initiating authority or the author of ballot proposals. Lastly, the sudd database should be considered by researchers who are specifically interested in the occurrence of referendums and other mechanisms in practice (on the national level but also for independent and dependent regions).

## 4 Direct Democracy Measurement Approaches

After examining available data on direct democracy, we discuss how such data can be aggregated into general assessments of direct democracy. We therefore introduce measurement approaches derived from contemporary research and examine them in regard to which dimensions they cover in which way. First, we elaborate on measurement approaches assessing direct democratic rules in form and rules in use separately, starting with the former. Then, two indices that cover both rules in use and rules in form in an aggregated way are introduced. Table 3 gives an overview of all approaches in regard to which mechanisms they include and whether they account for easiness and decisiveness in their aggregation (with a weighting procedure), and gives some general notes on the indices' calculation.

### 4.1 Approaches assessing Rules in Form

This section introduces the three rules in form measures, which are compared in this paper: Gherghina (2016), Peters (2016), and the Legal Provisions for Direct Democracy subcomponent of the Democracy Barometer (Barometer 2016). A simple count measure of Legal Designs, for which data was derived from the Direct Democracy Navigator (which is also used as a comparison indicator), is not included here, as it is not considered a *measurement approach* in the actual sense and not derived from literature, but instead constructed in an ad hoc manner.

#### 4.1.1 Gherghina 2016

A current measurement approach explicitly dealing with rules in form comes from Gherghina (2016) pp. 7-9, who relies on data primarily from the IDEA Database. Five direct democratic institutions are covered: Obligatory and optional referendums, citizens and agenda initiative as well as recall. The index relies solely on their existence, ranging from 0 - 5, with each point standing for the existence of one of the institutions. A second index is constructed which measures the level of subnational direct democracy in the same manner, adding up the scores for regional and local direct democracy (ranging from 0-10). However, there is no differentiation between top-down and bottom-up institutions and a disadvantage of the approach taken by Gherghina is that neither decisiveness nor easiness are part of the measurement. On the upside, this also means that the indicator is rather effortlessly constructed and allows for a straightforward interpretation. On the other hand, one might call into question if the values of such a count index are equidistant and if each mechanism should count the same, for example, does the existence of a top-down mechanism imply the exact same importance than a bottom-up mechanism? Unfortunately, Gherghina does not aggregate both indices into one general index of direct democratic rules in form.

#### 4.1.2 Peters 2016

Another approach in assessing rules in form comes from Peters (2016), who examines the effects of direct democracy on party membership. She differentiates between bottom-up and top-down direct democracy, calculating an aggregated measure based on two sub-indices, all ranging between 0 and 1.5. Unfortunately, Peters' description of her construction of the bottom-up indicator is quite ambiguous and thus reproducing her measure for the empirical comparison became a difficult task. For this reason, there will be a thorough discussion of her measurement approach.

As described in the variable description in the study's appendix (Peters 2016: 155), the top-down direct democracy indicator is composed of the variables constitutional referendum (0 = no institution; 2 = institution exists) and legislative referendum (0 = no institution; 0.5 = institution exists, but referendum is not binding; 1 = institution exists and is binding). The result is described as five-point indicator ranging "from 0 to 1.5, where 0 implies that there are no top-down structures and 1.5 means that there are provisions for both constitutional and binding legislative referendums" (Peters 2016: 144). As it turns out, the indicator can take six possible values, though the difference could be caused by the fact that in Peters sample of 16 countries, there is no country that solely provides the possibility of a constitutional referendum, and therefore there is no empirical value of 2 or respectively 1 in the rescaled indicator (for the data description see the online appendix of Peters (2016), which can be found on the authors website). The other empirical values (as shown in Peters' Figure 1, pp. 147), correspond to an aggregation method of adding the first two variables and dividing the sum by 2.

According to the variable description, the bottom-up indicator of direct democracy is a "three-point index composed of 'citizen initiative' (0 = no institution; 3 = institution exists (1.5 = when it exists but is limited, e.g., to existing laws only)) and 'agenda initiative' (0 = no institution; 1.5 = institution exists)" (Peters 2016: 155). Again, there is a difference in possible combinations of the variables and the description of a "three-point index", which could also be explained by the fact that there are only three empirical values. On the other hand, when reading the descriptions, appendix and online appendix it's not easily comprehended how exactly the indicator was created. When interpreting the empirical results for the bottom-up indicator, Peters describes the values in parentheses (Peters 2016: 147): "popular initiatives (a score of 1.5)" and "agenda initiatives (a score of 0.75)". The interpretation does not correspond to the aforementioned description of the bottom-up index as "ranging between 0 and 1.5, where 0 implies that there are no provisions for either an agenda or citizen initiative and 1.5 means that there are provisions for both" [Peters (2016) pp. 145; emphasis by the authors].

Table 3: Overview of Used Direct Democracy Measures

Author/Main Source	Mechanisms	Easiness	Decisiveness	Calculation	Values
<i>Rules in Form</i>					
Gherghina 2016	Mandatory and Optional Referendum, Citizens' Initiative, Agenda Initiative, Recall	-	-	Counts existence of legal provisions for each mechanism.	0 - 5, discrete
Democracy Barometer	Mandatory Referendum, Veto-Player Referendum, Popular Veto, Popular Initiative	only ease of approval	only binding mechanisms	Counts existence of legal provisions for each mechanism, multiplied with a quorum variable.	0 - 4, continuous
Peters 2016	Referendums, Plebiscites, Citizens' Initiatives, Agenda Initiatives	-	non-binding plebiscites count half	Direct Democracy Index: <i>Bottom-Up + Top-Down</i>	0 - 1.5, discrete
<i>Rules in Use</i>					
Gherghina 2016	referendums ("issues put to a vote")	only ease of approval	non-binding referendums count half	Summed up over 19-year period. Non-binding mechanisms receive a 0.5 weight and score is multiplied with quorum variable.	continuous
Blume et al. 2007	Mandatory Referendum, Optional (Citizen Initiated) Referendum, Initiative	-	-	Summed up over 10-year period. Categorization into four categorie.	0-3, discrete
Either sum or average: Bauer & Fatke 2014, Peters 2016, Democracy Barometer	Depends on Measurement	-	-	Sum or average of the occurrence of mechanisms (within a determined time-frame), followed by logarithmization.	continuous
<i>Mixed Approaches</i>					
Fiorino 2017, DDI	Citizens' Initiative, Agenda Initiative, Facultative Referendum, Obligatory Referendum, Plebiscite	both ease of initiation and approval	accounted for	Qualitative rating into seven categories (exact aggregation method unclear).	1-7, discrete
Altman 2017, DPVI (V-Dem)	Obligatory Referendum, Plebiscite, (Facultative) Referendum, Citizens' Initiative	both ease of initiation and approval	accounted for	For each mechanism: <i>Ease of Initiation + Ease of Approval × Consequences</i> For detailed explanation see Altman 2017.	0-1, continuous



If the variable agenda initiative can take the values 0 and 1.5, and citizen initiative a maximum of 3, it could again be the case that the sample only had countries with either agenda or citizen initiatives, but not both. Table A in Peters online appendix lists two countries with agenda initiatives (Austria and Netherlands) and two countries (Italy and Switzerland) with the description “agenda/popular initiative”, which implies that both institutions exist there.<sup>6</sup> Unfortunately, there is no country in Peters’ sample that only offers the possibility of a citizen/popular initiative, but not the agenda initiative, which would enable a more precise reconstruction of the aggregation method.

Considering the variable descriptions and the appendices, the *first possibility* is that Peters did not aggregate the bottom-up indicator out of the variables citizens initiative and agenda initiative through addition (as the top-down indicator appears to be aggregated), but instead in a way that ignores agenda initiatives in the case that citizens initiatives are provided. Given sample data with only both institutions, the variable takes values from 0 for no institution, 1.5 for only agenda initiatives and 3 for citizens initiatives (or respective values of 0.75 and 1.5 on a variable ranging from 0 to 1.5), corresponding to the parenthesized values in Peters interpretation text. We are not sure whether this is the case, as the construction of the two variables for the bottom-up indicator is composed of different values for the existence of each institution, which implies a deliberate weighting. For example, agenda initiatives are given only half the weight of the initial values of citizens initiatives, much like the mandatory referendum which counts twice as much as plebiscites. Additionally, it would be completely unclear how the score for a hypothetical country with a citizens’ initiative value of 1.5 and an agenda initiative value of 1.5 would be aggregated, being the only value in the index actually reflecting “both institutions”. A construction ignoring agenda initiatives once citizens initiatives are provided implies that once there is the latter, the former does not matter. Mechanisms initiated by referendum or by parliament are usually considered as different institutions (e.g. IDEA and V-Dem). Moreover, the initial introduction of the bottom-up index describes the value 1.5 as the provision of both institutions, and the top-down indicator appears to be constructed additively. A *second possibility* is that Peters assigned Italy and Switzerland the value 1.5 on the citizen initiative variable, although in a cross-check with IDEA the institution did not appear to be strongly “limited” in both countries, neither does Peters’ online appendix hint at that. *Thirdly*, Peters could have made a mistake when aggregating the bottom-up indicator, assigning to countries with both provisions a value of 3 instead of the actual 4.5 (which would be the result of an addition). If so, the rescaling to a range of 1.5 resulted in higher values for the agenda initiative, which would result in a score of 0.5 (instead of the 0.75 specified by Peters).

<sup>6</sup>Cross-checking with the IDEA Database shows that Italy has both agenda and citizens initiative, while for Switzerland only the citizens initiative, but not the agenda initiative is present. The agenda initiative description of IDEA states that Switzerland provides “little more than a right to petition any government organ or legislature to do something - pass legislation, change the constitution, etc. The respective agency or organ is not under a legal obligation to consider such proposals”, see: <https://www.idea.int/node/205203>) Peters conception of bottom-up direct democracy includes “that citizens themselves can organise the collection of signatures in order to [...] force the legislature or government to address an issue they put forward” (Peters 2016: 144), which might be less restrictive than the IDEA classification.

After constructing the bottom-up and top-down indicators “the general measure of direct democratic institutions is simply a combination of the latter two measures; it is an eight-point index ranging from 0 (no institutions) to 1.5 (constitutional, legislative and popular referendum institutions)” (Peters 2016: 145). To sum up, the covered institutions are constitutional referendums, plebiscites, citizen initiatives and agenda initiatives. The decisiveness is captured by applying only half the weight if legislative referendums are non-binding. As both subindices range from 0 to 1.5, it appears that there is no weighting between top-down and bottom-up institutions. Within the subindices, obligatory referendums are assigned twice the value of a binding legislative referendum, the same applies for citizen compared to agenda initiatives. Neither ease of initiation nor ease of approval are covered in the indicator, nor is the local/regional dimension included. Whether citizens’ initiatives are restricted to certain issues is accounted for, as a “limited” provision of citizens initiative gives only half of the maximum value. In general the index has the advantage of a rather simple construction (besides the already discussed issues with replicability), and the provision of separate bottom-up and top-down indices.

#### 4.1.3 Democracy Barometer

In their index for the quality of democracy, the Democracy Barometer provides a subcomponent called *constitutional provisions for direct democracy*, available in country-year format between 1990 and 2014, covering 68 countries (Merkel et al. 2016: 49). The subcomponent is composed of an indicator measuring direct democracy provisions and as well as their respective required quora. The variables are then aggregated by taking the arithmetic mean for the standardized indicators. The first indicator consists of four direct democracy mechanisms, each present institution improves the score by one (Merkel et al. 2016: 49):

- 1.) *Mandatory referendum*
- 2.) *veto-player referendum*: referendum is triggered and question is asked by an existing veto-player
- 3.) *popular veto*: non veto-player triggers referendum, but question is asked by an existing veto player
- 4.) *popular initiative*: non veto-player asks question and triggers referendum

The underlying typology of institutions differs from the other discussed approaches, as is it also considers whether a top-down initiator is a political veto-player and only considers binding referendums. Easiness of approval is included in the second indicator, which measures constitutional provisions for approval or participation quorum in direct democratic votes (in case of different instruments, the quorum applying for most is used). The variable is calculated by subtracting the quorum from 1, which means smaller values indicate less ease of approval. In case of approval quorums, the value is multiplied with 2 before the subtraction, because in order to meet an approval quorum of 25%, at least 50% of the population must participate. Countries with no direct democracy are given the value of the country with the highest quorum.

## 4.2 Approaches assessing Rules in Use

The majority of approaches measure the rules in use by aggregating the number of referendums and initiatives in the past by different procedures, which will now be discussed. This section introduces the three measures considered for rules in use, which are derived from Gherghina (2016), summed measures (cf. Bauer & Fatke 2014) and categorized measures (cf. Blume et al. 2007).

### **Gherghina 2016**

Gherghina operationalizes the use of referendums at national level by calculating the number of referendums from 1990-2008 (defined as a question/issue put to a vote), but also including a weighting accounting for decisiveness and easiness (cf. Gherghina 2016: 9). Merely consultative referendums are divided by two. The quorums are assessed for binding and consultative referendums separately, with an ordinal variable of three values: 1 = both turnout and approval quorums, 2 = turnout quorum only, and 3 = no quorum required. This kind of construction is an exception, as most studies assessing rules in use and rules in form separately include easiness and decisiveness within their measurement of formal rules or not at all. It is noteworthy that Gherghina does not assess the ease of use, but only the ease of approval, as he does not differentiate between top-down referenda/plebiscites and bottom-up mechanisms when aggregating the rules in use indicator.

### **Summed Measures (e. g. Bauer/Fatke 2014)**

Besides Gherghina, the examined approaches assessing rules in use separately all use either the sum or average of the occurrence of direct democratic mechanisms within a determined time-frame. Bauer and Fatke assess rules in use for the Swiss cantons by averaging the number of all cantonal initiatives and optional referendums per year in the period 2002 to 2006 (mandatory referendums are excluded, as they don't fit the authors' theoretical argument) (cf. Bauer & Fatke 2014: 54). As Peters' study relies on time-series data with four time-points in four decades, she counted the number of referendums held for each decade preceding the measurement of her dependent variable (cf. Peters 2016: 155). A rather similar procedure is used in the Democracy Barometer when assessing the effective use of direct democratic instruments, which sums up the number of national non-mandatory referendums per year (cf. Merkel et al. 2016: 47). As a last step, Peters' indicator as well as the one calculated by the Democracy Barometer are logarithmized. For count measures it might be recommendable to use logged variables, especially since additional referendums might be less relevant in countries where there already are many referenda [cf. dembarcodebook pp. 47].

### Categorized Measures (e. g. Blume 2007)

The approach of Blume et al. differs from the already discussed in that it categorizes the number of direct democratic votes in the timeframe of 1996 to 2005) into four categories, ranging from “0 = no factually observed direct democracy” to “1 = low level of factually observed direct democracy (i.e. one or two votes)” to “2 = medium level of factually observed direct democracy (i.e. three to five votes)” to “3 = high level of direct democracy (i.e. more than five votes)” (Blume et al. 2007: 14).

## 4.3 Approaches assessing Rules in Form and Rules in Use

This section introduces measures that explicitly include both rules in form and use. Two of such indices will be utilized: the measure created by Fiorino et al. (2017) and the *Direct Popular Vote Index* in the V-Dem dataset.

### Direct Democracy Index

An index that is supposed to measure direct democracy is the Direct Democracy Index (DDI) by Fiorino et al. (2017), who use data from Kaufmann (2004), Madroñal (2005) and Hwang (2005). Fiorino et al. apply Kaufmanns’ seven-point rating of direct democracy by consulting country reports, covering a dataset of 87 countries for the period of 2000-2005 (cf. Fiorino et al. 2017: 14; see appendix A1 for the original description by Kaufmann 2004).<sup>7</sup> The seven categories are:

- 1.) *radical democrat*
- 2.) *progressive*
- 3.) *cautious*
- 4.) *hesitant*
- 5.) *fearful*
- 6.) *beginner*
- 7.) *authoritarian*

Each country’s score is based on a qualitative assessment of the direct democratic quality of the political system. The mechanisms explicitly referred to in the coding scheme, as described by Kaufmann (2004) pp. 26-29, are *citizens’ and agenda initiatives*, *obligatory referendums*, *facultative referendums* and *plebiscites*. The actual use of direct democratic institutions is explicitly included. The index covers “the procedures a political system provides in order to propose, approve, amend, and delete laws through popular initiative and referendums, as well the actual practices of direct democracy and the general political condition a country experiences” (Fiorino et al. 2017: 148). The easiness of initiation and approval is covered by the consideration of entry hurdles, time limits and majority requirements/quora. As Blume et al. point out, an important disadvantage of the DDI

<sup>7</sup>See Kaufmann & Waters (2004) for a different version of the index applied to European countries.

is that “the criteria used for weighing the different criteria remain completely opaque and that it does not tell anything on the relevance of institutional details” (Blume et al. 2007: 12). The actual weight of the dimensions bottom-up/top-down, subnational direct democracy, easiness, decisiveness and de jure/de facto measures for the index construction is therefore unclear, though some hints can be drawn when consulting the categories descriptions. For example, the categories 7 and 6 are only ascribed to countries that embody bottom-up mechanisms as well as obligatory referendums. Category 5 consists of countries that have practical experience, but where the procedures have only plebiscitary character (cf. Kaufmann 2004: 39), so bindingness as well as bottom-up mechanisms seem to be given a greater weight.

### Direct Popular Vote Index (DPVI)

The most complex of the introduced measurements is constructed by Altman, using V-Dem data (Altman 2017; Coppedge Michael et al. 2017: 62f., 138-153). The concept to be measured consists not only of de jure measures but also includes the actual use of direct democracy. In his paper, Altman names his index *Direct Democratic Practice Potential* (DDPP), however the corresponding V-Dem index is referred to *Direct Popular Vote Index* (DPVI), which will be used in this paper for empirical comparison. Mostly, the described variables and aggregation method of the DDPP resembles the one described in the V-Dem codebook for the DPVI, with one major exception that will be discussed later on. The mechanisms covered are *popular initiatives*, *popular referendums*, *obligatory referendums* and *plebiscites*, whereas the first two are citizen initiated (in our terms bottom-up) and the latter two are considered top-down mechanisms.

Altman/V-Dem provide not only an aggregated index, but also sub-indices for the bottom-up/top-down dimensions as well as a measure for each mechanism itself. The aggregated score for each mechanism results from the addition of the values for the variables ease of initiation and ease of approval, which is then multiplied with a term called *consequentiality*. The ease of initiation consists of the existence of the mechanism and, in case that it is initiated from the people, the number of signature and time limits for gathering them. The major difference between the DDPP and DPVI is that the latter also considers whether the respective institutions exist on the national, subnational level or both (Coppedge Michael et al. 2017: 62).<sup>8</sup> Ease of approval is measured by calculating the surface of the polygon determined by the quora/thresholds for participation, approval and supermajority, and multiplying the result with district majority. Altman chooses this intricate method of calculating a polygon area for the construction of his index because many less complex aggregation methods would not adequately account for the interaction between different quora (for example, an approval quorum of 20% already implies a participation/turnout quorum of 40%) (cf. Altman 2017: 1213–5). The consequentiality variable consists of the decisiveness,

<sup>8</sup>There is no hint to the reason why the term is only included in the DPVI and not the DDPP. The subindices for the respective institutions, as described in the V-Dem codebook, do not include the variable in their description, which makes it even more puzzling.

whether the vote is binding (a value of 1) or consultative (0.75), as well as another variable called credible threat. Credible threat is calculated based on the frequency and success of direct popular votes in the past. The aggregated index results from the addition of the weighted scores for each mechanism. The bottom up mechanisms receive a weighting of 1.5 to account for the penalty the index construction applies for signature thresholds and time gathering periods (which are not relevant for top-down initiated mechanisms). Lastly, the DPVI is standardized to range between 0 and 1. As the construction of the DPVI is rather complex, we refrain from describing it thoroughly and recommend the reader to consult Altman (2017) for further explanation.

A great advantage of the DPVI is the scope of the considered dimensions, as well as the yearly coverage of 178 countries since 1900 and the availability of these data. The decisiveness is accounted for, and also the easiness, with an explicit distinction between ease of initiation and approval and an elaborate aggregation method for different quora. Moreover, separate indices for bottom-up and top-down mechanisms are provided. Rules in use are not separately measured but implicitly included in the variable measuring credible threat, although the purpose of this variable is not to directly measure de facto direct democracy, but rather the usage potential of a mechanism. With the advantage of the broad scope of dimensions and complexity of the index aggregation comes the drawback of a less straightforward interpretation of the aggregated index and sub-indices.

#### 4.4 Summary of Measurement Approaches

In conclusion, all of the examined approaches have in common that their measurements rely to some degree on the existence of a defined set of institutions, and in case of rules in use assessments, whether these institutions are used in practice within a certain time-frame. Some of the approaches give these mechanisms equal weighting in aggregation, but sometimes bottom-up institutions get a greater weight than top-down mechanisms, binding referendums get assigned higher values than consultative ones, or obligatory referendums weigh more than plebiscites. Mostly, those values seem to be chosen rather arbitrarily, as there is, for example no reference to draw from when to decide if consultative mechanisms (in comparison with binding ones) should be given the same value (cf. Gherghina 2016), a 0.75 weighting (cf. Altman 2017), half the weight (cf. Peters 2016), or be not considered at all (cf. Barometer 2016). Moreover, the approaches differ in regard to whether they consider the easiness of approval and/or initiation and if so, in which way they are included in the index aggregation. For example, the DPVI is constructed by assessing ease of approval with a complex calculation of the surface of a polygon, which is determined by three quorum variables, while Gherghina (2016) classifies quora with a categorical three-point variable. It is far from clear, if and how these dimensions should be weighted to assess the degree of direct democracy in a country most adequately, and besides Altman (2017), most authors do not spend much words on elaborating on their index construction. In general, there is no recommendation in regard to which

index construction to use in one’s empirical research, as it highly depends on the research question.

Lastly, it has to be noted, that most of the discussed approaches are only applied to democracies (e.g. Peters 2016; Gherghina 2016), which makes the discussion of the *democracy* part of direct democracy unnecessary. Since we reconstruct the indicators for a range of all available countries, this question has to be addressed. The direct democratic mechanisms mentioned in this paper, even though called *democratic* are not necessarily limited to democratic countries. Interestingly, some autocratic or semi-democratic regimes have adopted direct popular vote instruments, using them to seemingly legitimate their rule, a group of countries Altman labels the “nightmare team” (Altman 2011: 92). Of the discussed measurement approaches, which cover a wider range of countries, the rather qualitative rating of Fiorino et al. (2017) accounts for the general state of democracy a country experiences. The index provided by V-Dem itself is a sub-index of a much broader assessment of democratic quality, and not accidentally called Direct Popular Vote Index. Therefore, in the empirical comparison, we examine available as well as reconstructed indices in relation to their three-fold freedom house classification. An important notion for further research is to find valid ways to aggregate indices assessing direct popular votes and indicators representing the degree of democracy of a given country. Some might argue that any direct democracy index must account for the degree of authoritarianism or democracy in a country, or else it does not capture the concept well. Others could argue that authoritarian and semi-democratic regimes use direct popular votes for other reasons than democracies but that doesn’t mean that they should be generally excluded from the analysis. It’s not necessarily intuitive to determine the weighting of the two dimensions in regard to each other, and to decide whether the aggregation should be constructed multiplicatively, additively or by another mathematical operation.

## 5 Data Wrangling & Empirical Comparison

After discussing available data sources and measurement approaches, it seems worthwhile to compare them empirically as well. This chapter introduces the examined measurements, explains the many ways the data had to be transformed in order to reproduce some of the previously discussed measurements, and therefore may serve as an example to researchers interested in measuring direct democracy with the given data. Subsequently the reconstruction of the selected direct democracy measures is described (Section 5.1). Following that, we empirically compare the indicators of rules in form (5.2) and rules in use (Section 5.3) as well as measures that include both (Section 5.4) separately.<sup>9</sup>

### 5.1 Data Wrangling

This section elaborates on the various transformations that had to be applied in order to reproduce the selected measurement approaches. It further introduces the ready-to-use variables taken from other sources. First, the rules in form measures are discussed, followed by the rules in use indicators and the mixed approaches.

#### 5.1.1 Construction of Rules in Form Measures

First, we consider how the different rules in form measures were constructed. In order to make the measures comparable, we use the most recent data available for each dataset. In total, there are three rules in form measures to be constructed: *Gherghina - RiF*, *Peters - RiF* and *V-Dem RiF*, with respective bottom-up and top-down components when available. Additionally, two indicators will also be used in order to enhance the comparison: *Direct Democracy Provisions* from the Democracy Barometer and *Direct Democracy Legal Designs* from the Direct Democracy Navigator. Table 4 shows descriptives statistics for the used measures.

Table 4: Rules in Form

Rules in Form Measures	N(=Countries)	Mean	SD	Median	Range	Time	Data
Gherghina - RiF	192	1.76	1.24	2.00	0 - 5	Most Recent	IDEA
Peters - RiF	196	0.62	0.46	0.50	0 - 1.5	"	"
<i>Top-Down</i>	196	0.80	0.53	1.00	0 - 1.5	"	"
<i>Bottom-Up</i>	196	0.44	0.64	0.00	0 - 1.5	"	"
V-Dem - RiF	170	1.64	1.06	2.00	0 - 4	2016	V-Dem
<i>Top-Down</i>	172	1.26	0.69	1.00	0 - 2	"	"
<i>Bottom-Up</i>	170	0.38	0.71	0.00	0 - 2	"	"
Direct Democracy Provisions	68	1.65	1.38	1.33	0 - 4	2014	Democracy Barometer
Direct Democracy Legal Designs	107	2.86	1.94	2.00	1 - 12	Most Recent	Direct Democracy Navigator

<sup>9</sup>The entire code for all transformations and visualizations can be found on the following GitHub Repository: <https://github.com/favstats/DirectDemoracyPaper>



#### 5.1.1.1 Gherghina - RiF

As already discussed in the previous chapter, the Gherghina Rules in Form (referred to as: *Gherghina - RiF*) measure is a simple count of the given direct democracy mechanisms available in the IDEA dataset: mandatory/obligatory referendums, optional referendums, citizens initiatives, agenda initiatives and recall. This is easily reproduced for all available countries in the IDEA dataset. IDEA includes a column for each mechanism "*Legal provisions for [mechanism] at national level*", indicating whether "Yes" the respective provision is available in that country on the national level and "No" for when the measure is not available. In accordance with general practice, cases with "Yes" are coded as 1 and cases with "No" are coded as 0. A special case is when certain countries specify a condition such as "No, but ad hoc referendums are possible" (for example in Norway for optional referendums). It was decided to classify such cases as a "No" (= 0), as an ad hoc referendum implies that there generally is no legal basis on which a referendum can occur and therefore it is not an indicator of rules in form. Cases indicating "No data" are coded appropriately as missing value. As a next step the five variables are simply counted, leading to a 6-point scale ranging from 0 to 5, where 0 indicates that there are no provisions in place in that country and 5 means that all 5 of them are in place. Data was available for 192 countries and the average amount of provisions in place was 1.76 (SD = 1.24).<sup>10</sup>

#### 5.1.1.2 Peters - RiF

Next, it is described how the rules in form measure by Peters (2016) is constructed (referred to as: *Peters - RiF*). In order to calculate the Peters - RiF measure, IDEA is again the most appropriate dataset as it offers information on agenda initiatives and the aim of this reconstruction is to stay as close as possible to the original measure. As a first step, the column "*Who can initiate an optional referendum at the national level?*" will be parsed in order to extract information on the bottom-up or top-down nature of a optional referendum in a given country. Whenever a case includes the phrase "Registered Electors", the variable is coded as *Bottom-Up* and whenever a case includes a state institution ("Government", "Supreme Court", "Parliament" etc.) it is coded as *Top-Down*. The phrase "Not applicable" is coded as 0 in either case, as it implies that there is no obligatory referendum in that country in the first place, and again "No data" is coded as missing value.

Next, the bindingness of a given referendum is coded. The column "*Are referendum results binding?*" includes information on the bindingness of referendums. Whenever the phrase "Optional referendum - always" appears, the specific country receives a 1 for optional referendum decisiveness, whereas "Optional referendum - never" and "Optional referendum - sometimes" is classified as non-binding. We consider sometimes binding referendums as generally non-binding, as bindingness in that case

<sup>10</sup>As this paper focuses on the national level, we refer from reconstructing Gherghinas' index for the local/regional level, but recommend researchers to use the IDEA dataset if interested in doing so.

depends on other factors (such as ease of approval).<sup>11</sup> Now the variables are in place in order to construct the top-down following Peters (2016). Whenever a country has a *binding* optional referendum that is initiated by state institutions, it receives a score of 1, while a country with a *non-binding* top-down initiated optional referendum receives a score of 0.5, creating the variable *legislative referendum* equivalent to the Peters (2016) measure with the same name. Similarly, whenever there is a mandatory referendum in a given country, it receives the score of 2, otherwise a 0, creating the variable *constitutional referendum* equally constructed as the Peters (2016) variable with the same name. Subsequently, the the top-down measure can be constructed by adding up constitutional referendum and legislative referendum divided by 2, which results in a 0 to 1.5 scale.

For the bottom-up measure, a variable “citizens initiative” is calculated, which takes the value 3 if a citizens’ initiative exists and 0 if not.<sup>12</sup> The variable “agenda initiative” takes the value 1 if an agenda initiative is provided and 0 if not. The aggregated bottom up measure is constructed by adding both variables and dividing by 3, also resulting in a 0 to 1.5 scale. It has to be noted, that we are not sure whether this construction corresponds to the calculation as intended by Peters (see Chapter 4 for the discussion of the other possible construction method).

As a final step, Peters - RiF is constructed by adding up the top-down and the bottom-up measures and dividing them by two, creating yet another 0 to 1.5 scale. The average Peters - RiF score is 0.62 (SD = 0.46). The average top-down score of the Peters - RiF measure is 0.80 (SD = 0.53), whereas the average Bottom-Up score is 0.44 (SD = 0.64), already implying that there are more top-down rules in form on average. The three measures are available for 196 countries.

### 5.1.1.3 V-Dem RiF

The V-Dem rules in form measures (referred to as: *V-Dem RiF*) are an attempt to emulate the approach taken in Gherghina (2016), meaning, it’s a simple count index of the four direct democracy mechanisms that are available in the V-Dem dataset: obligatory referendums, plebiscites, initiatives and (facultative) referendums. This leads to a four-point scale, where 0 means no democracy mechanism is in place, whereas 4 means that all four legal provisions are in place. An advantage of the V-Dem data is that we can clearly distinguish between top-down and bottom-up mechanism. Therefore, we calculated a V-Dem - RiF top-down measure by summing obligatory referendums and plebiscites provisions (3-point measure) and a bottom-up measure by summing initiatives and facultative referendums (also a 3-point measure). On average there were 1.64 provisions in place (SD = 1.06), while the higher average value for top-down (Mean = 1.26, SD = 0.69) compared

<sup>11</sup>A different approach worth considering would be to give sometimes binding referendums a medium score. As we want to stay as close as possible to the original index construction by Peters - RiF, we decided to use the mentioned approach.

<sup>12</sup>As a departure from Peters (2016), we refrain from assigning a 1.5 category, as parsing and coding individual issues for each citizen or agenda initiative is a task that would require researching each initiative individually, something that is beyond the scope of this paper.

to bottom-up (Mean = 0.38, SD = 0.71) implies a greater prevalence of top-down rules in form measures. The V-Dem RiF measure and its bottom-up component are available for 170 countries, whereas the top-down component is available for 172 countries.

#### 5.1.1.4 Democracy Barometer and Direct Democracy Navigator Indicators

The following indicators are taken from their respective data sources: *Direct Democracy Provisions* from the Democracy Barometer and *Direct Democracy Legal Designs* from the Direct Democracy Navigator. Direct Democracy Provisions are available for 68 countries and similarly to Gherghina (2016) it is a simple count measure of direct democratic provisions, with an average of 1.65 (SD = 1.38) measures in place. Direct Democracy Legal Designs is a simple count of legal designs for direct democratic institutions available in a given country. As already discussed in Chapter 3, some inconsistencies can be observed when delving deeper into the database. However, we still consider it to be a valuable data source that is at least worthy of comparison. The average use of legal designs is 2.86 (SD = 1.94) and the measure is available for 107 countries. Note that the Direct Democracy Navigator does not distinguish between a country not having any provisions at all or a completely missing case, therefore there is no 0 category and only countries are recorded that have at least one legal design on the national level.

#### 5.1.2 Construction of Rules in Use Measures

Next, we will discuss how the rules in use measures were created. In order to facilitate comparison, we limit the data to a common (rather large) time range, as referendums are not very frequent in some countries, but still practiced occasionally. All rules in use measures were thus aggregated for a time-frame of 21 years (1996 - 2016), except for the Democracy Barometer data which was only available from 1996 to 2014. Table 5 shows descriptives statistics for the constructed indicators.

Table 5: Rules in Use

Rules in Use Measures	N(=Countries)	Mean	SD	Median	Range	Time	Data
Gherghina - RiU	161	9.38	46.62	1.00	0 - 552	1996 - 2016	V-Dem
V-Dem - RiU - Sum	172	3.74	15.43	1.00	0 - 184	"	"
<i>Top-Down</i>	172	2.38	6.80	1.00	0 - 66	"	"
<i>Bottom-Up</i>	172	1.37	11.25	0.00	0 - 144	"	"
Sudd - RiU - Sum	96	8.03	20.76	2.00	1 - 187	"	sudd
<i>Top-Down</i>	93	5.14	9.10	2.00	1 - 66	"	"
<i>Bottom-Up</i>	24	12.21	29.50	4.00	1 - 146	"	"
V-Dem - RiU - Cat.	172	1.10	1.52	1.00	0 - 5	"	V-Dem
<i>Top-Down</i>	172	0.94	1.32	1.00	0 - 5	"	"
<i>Bottom-Up</i>	172	0.26	0.92	0.00	0 - 5	"	"
Sudd - RiU - Cat.	96	2.29	1.64	1.00	1 - 5	"	sudd
<i>Top-Down</i>	93	1.99	1.46	1.00	1 - 5	"	"
<i>Bottom-Up</i>	24	2.54	1.64	2.00	1 - 5	"	"
Effective Use	71	0.80	1.88	0.00	0 - 13	1996 - 2014	Democracy Barometer
Credible Use	172	1.20	2.65	0.00	0 - 21	1996 - 2016	V-Dem

### 5.1.2.1 Gherghina - RiU

The Gherghina Rules in Use measure (referred to as: *Gherghina - RiU*) will be constructed with the help of the V-Dem dataset. As a first step, all occurrences are counted for each available direct democracy institution: obligatory referendums, plebiscites, initiatives and referendums. Table 6 shows all the used variables of the V-Dem dataset.

Table 6: V-Dem Variables used for Gherghina - RiU Indicator Construction

Mechanism	V-Dem Variables		
	Occurences	Quora	Decisiveness
Obligatory Referendum	v2ddyror	Turnout Quorum: v2ddpartor Approval Quorum: v2ddappor	v2ddlexor
Plebiscite	v2ddyrrp	Turnout Quorum: v2ddpartpl Approval Quorum: v2ddapprpl	v2ddlexpl
Initiative	v2ddyrci	Turnout Quorum: v2ddpartci Approval Quorum: v2ddapprci	v2ddlexci
(Facultative) Referendum	v2ddyrrf	Turnout Quorum: v2ddpartrf Approval Quorum: v2ddapprrf	v2ddlexrf

Source: See V-Dem Codebook (cf. Coppedge Michael et al. 2017)

As a next step, turnout and approval quora are coded into 1, when a quorum existed, and 0, when it did not, for each of the four mechanisms separately. Now, the quorum variable can be calculated as Gherghina (2016) suggested: when both a turnout and an initiation quorum are present, the variable is coded as 1. When a turnout quorum was present but no approval quorum, the variable is coded as 2. And when no turnout and no approval quorum was present, the variable was coded as 3. This leaves us with a case that wasn't considered by Gherghina: what if there is an approval but no turnout quorum? It was decided that this value set will also receive a score of 2, equivalent to the opposite setup.

Following that, the decisiveness for each mechanism is also coded into 1, when a mechanism was binding, and 0, when it wasn't, so we are able to distinguish between consultative and binding referendums. However, given the different data structures, the formula had to be applied to each mechanism individually. So when a mechanism was binding, the use of that mechanism was multiplied by the respective quorum variable calculated above. When the mechanism wasn't binding, it was first divided by two and then multiplied by the quorum. Finally, all four mechanisms that have been transformed individually can now be summed up, creating the *Gherghina - RiU* measure. The average score for the Gherghina - RiU measure is 9.38 (SD = 46.62) and is available for 161 countries.<sup>13</sup>

<sup>13</sup>Although the interpretation of the mean should be done carefully here, as serious outliers like Switzerland are included in the data.

### 5.1.2.2 Summed, Categorized and other Rules in Use Measures

#### V-Dem - RiU - Sum

The summed V-Dem Rules in Use measure (referred to as: *V-Dem - RiU - Sum*) is constructed in a straightforward fashion. It merely consists of summing up the occurrence of any direct democracy mechanism in the V-Dem data in the time range from 1996 to 2016. Given that we are dealing with V-Dem data, at this place we can use the same procedure for the bottom-up and top-down measures. All measures are available for 172 countries and the V-Dem - RiU - Sum shows an average use of 3.74 (SD = 15.43) mechanisms, whereas Top-Down mechanisms were used 2.38 (SD = 6.80) and Bottom-Up mechanisms were used 1.37 (SD = 11.25) times on average.<sup>14</sup>

#### Sudd - RiU - Sum

The most extensive transformation had to be done on the sudd data. This dataset was not available for easy download at the time of writing of this paper. The creator of the database, Beat Müller, is currently working on simplifying the structure and making it available in XML format, though at the time of finishing this paper, this task wasn't accomplished yet. In the meantime, the sudd data had to be web scraped from the website directly, which was done using the R programming language. Given that our main interest in this paper is the spread of direct democracy on the national level, we first excluded all referendums from regions that were dependent at the time (mostly colonized entities) or were held in only part of a sovereign country. Votes held in disputed territories or regions generally not recognized as sovereign states such as Southeast Ossetia, Crimea or the (hoax) micronation of Sealand are also excluded. Further, the sudd data also includes popular votes that were used to elect Prime Ministers, Presidents, Parliamentarians and even Kings, which for the purposes of this comparison is not relevant and thus excluded. Lastly, we also excluded those referendums that were categorized as "unofficial vote", as there is no such category for the other compared datasets.

In order to capture the bottom-up and top-down dimensions for the sudd data we differentiated between referendums that were initiated "by the people", which were labelled as bottom-up referendums/citizens' initiatives, and those that were initiated by either a governmental, institutional or bureaucratic entity, which were categorized as plebiscite/top-down referendum<sup>15</sup>.

Lastly, all occurrences of mechanisms in the time-frame between 1996 and 2016 were summed up. On average, there were 8.03 (SD = 20.76) occurrences of direct democracy mechanisms in 96 countries, whereas there were 93 countries that used 5.14 (SD = 9.10) Top-Down mechanisms and 24 countries that used 12.21 (SD = 29.50) Bottom-Up mechanisms on average.<sup>16</sup>

<sup>14</sup>Again the interpretation of the mean should be done carefully as outliers are included.

<sup>15</sup>For example Uruguay has referendums labeled as initiative, even though they weren't initiated by the citizens. According to our typology, these were labelled as top-down mechanisms

<sup>16</sup>Once more the interpretation of the mean should be done carefully as outliers are included.

Note that the sudd data is recorded on the referendum level and there is no “0” category in the measurements, meaning when a direct democracy mechanism didn’t occur, it’s not recorded in the data. Therefore, all rules in use measures that use the sudd data start with 1 rather than 0 and is therefore missing for countries in which no referendum took place.<sup>17</sup> That is also the reason why there is such a high discrepancy between sample size when differentiating between top-down and bottom-up measures, as the latter only occurs in 24 countries in the sudd data.

### V-Dem - RiU - Categorized/Sudd - RiU - Categorized

The V-Dem/Sudd - RiU - Categorized measures follow the example of Bauer & Fatke (2014), and categorize the summed occurrences of all mechanisms into five categories (due to the longer time-frame, we extended the threefold classification to five categories):

0 = 0 occurrences

1 = 1 or 2 occurrences

2 = 3 or 4 occurrences

3 = 5 or 6 occurrences

4 = 7 or 8 occurrences

5 = 9+ occurrences

Both the categorized sudd and V-Dem measures are available in their top-down and bottom-up format. The average category for the overall measures is 1.10 (SD = 1.52) for the categorized V-Dem - RiU and 2.29 (SD = 1.64) for the categorized Sudd - RiU measure. The respective means for the top-down measures are 0.94 (SD = 1.32) for the categorized V-Dem - RiU and 1.99 (SD = 1.46) for the categorized Sudd - RiU measures. The respective means for the bottom-up measures are 0.26 (SD = 0.92) for the categorized V-Dem - RiU and 2.54 (SD = 1.64) for the categorized Sudd - RiU measures. Given that only transformations were applied, data availability for these measures is the same as in the two sections above that use V-Dem and sudd data.

### Other Indicators

As a means of comparison *Effective Use* and *Credible Use* have been taken from their respective data sources (Democracy Barometer and V-Dem). The effective use measure sums up the use of non-mandatory referendums in a given year and ultimately logarithmizes the measure “to account for the fact that an additional referendum is less important in countries with many referenda than in countries with few referenda” (Merkel et al. 2016: 47). The average effective use is 0.80 (SD = 1.88) and the measure is available for 71 countries. Finally, the credible use measure is meant to capture the occurrence of a direct democratic vote that is *credible*, meaning that “the official

<sup>17</sup>The sudd website states that information is available for almost all countries in the world, indicating that a country not appearing in a given year really means that no referendums have taken place and that the case is in fact not missing. However, there is no obvious way to distinguish whether a case was truly missing or “correctly not recorded”. On the other hand, the approach used in this paper has the advantage of making it possible to depict rules in use only for countries where they took place at least once. Especially the bottom-up rules in use measures are highly skewed with a large amount of cases having a value of 0

results of the vote(s) reflect the actual vote (leaving aside issues of voter exclusion, intimidation, or vote-buying)” (Coppedge Michael et al. 2017: 151). As the variable can only take the value 0 or 1 for a given year, the maximum value for a country is 21, indicating that every year between 1996-2016 at least one credible direct democracy mechanism occurred. The average credible use is 1.20 (SD = 2.65) and the variable is available for 172 countries.

### 5.1.3 Mixed Measures and Additional Variables

The *Direct Popular Vote Index (DPVI)* and the *Direct Democracy Index (DDI)* are taken from their respective data sources (V-Dem and Fiorino et al. 2017). Given that Fiorino et al. (2017) created their DDI measurement for countries in the range between 2000 and 2005, the DPVI variable has also been averaged for this timeframe, in order to ensure comparability. Further, the DPVI measure is also available for its top-down (Mean = 0.20, SD = 0.20) and bottom-up components (Mean = 0.04, SD = 0.12). The overall mean DPVI score is 0.10 (SD = 0.12), whereas the mean DDI score is 3.03 (SD = 1.83). Table 7 shows descriptives statistics for the used measures.

Table 7: Mixed Indices

Mixed Measures	N(=Countries)	Mean	SD	Median	Range	Time	Data
Direct Popular Vote Index (DPVI)	172	0.10	0.12	0.06	0 - 1	2000 - 2005	V-Dem
<i>Top-Down</i>	172	0.20	0.20	0.11	0 - 1	2000 - 2005	”
<i>Bottom-Up</i>	172	0.04	0.12	0.00	0 - 1	2000 - 2005	”
Direct Democracy Index	87	3.03	1.83	3.00	1 - 7	2000 - 2005	Fiorino et al. 2017

### Additional variable: Freedom House Index

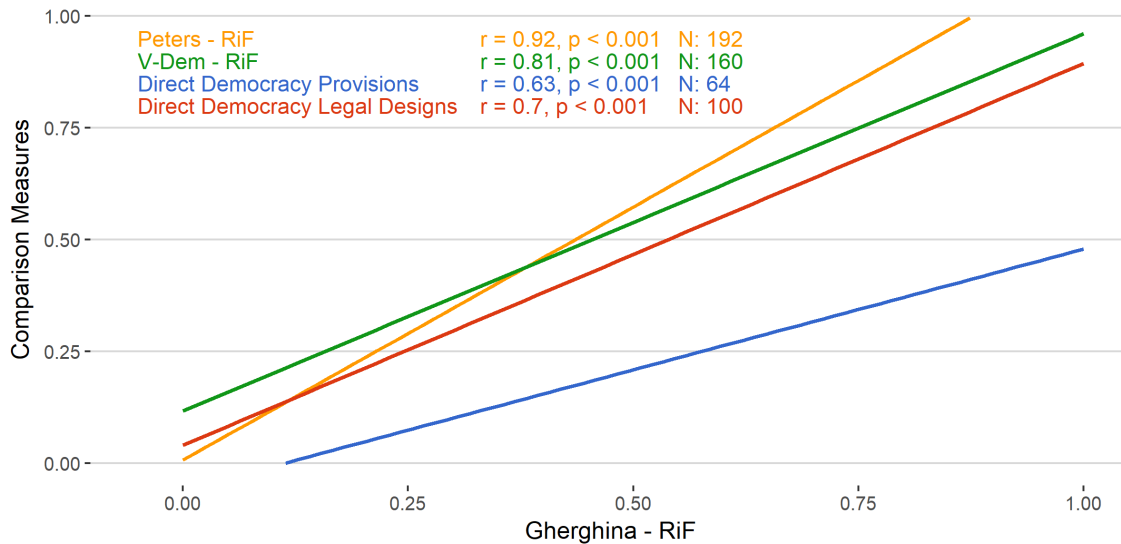
As stated in the previous chapter, some autocratic or semi-democratic regimes have adopted direct popular vote instruments, a group of countries Altman labels the “nightmare team” (Altman 2011: 92). As a first exploratory application of the measures presented in this paper, we seek to examine the rules in form, rules in use and mixed measurement approaches for direct democracy for different regime types. In order to do that, we use the Freedom House index (Freedom House 2018) and its three-category classification of “Free”, “Partly Free” and “Not Free”. We then seek to descriptively evaluate whether certain mechanism types occur in different regime types. For the rules in form and rules in use measure, the Freedom House index is taken for the year 2016 while the mixed measures are compared to the Freedom House index of the year 2005.

The following section compares the variables with each other: first rules rules in form are compared, followed by a comparison of rules in use and mixed measurements, while also accounting for top-down and bottom-up dimensions whenever possible. In order to allow for an intuitive visual comparison, all variables are normalized to range from 0-1. A heatmap, show the correlation between all direct democracy assessments can be found in Appendix A1.

## 5.2 Comparing Rules in Form Measures

In order to compare the indicators for rules in form, the bivariate correlations are examined at first. To keep the comparison between the measures simple, we do not visualize the relationship between all indicators and only compare the Peters and Gherghina rules in form measures to the other indices by depicting the regression lines, with the drawback of some loss of information. Considering the correlations depicted in Figure 1 and Figure 2, it can be observed that Peters - RiF corresponds the most to Gherghina - RiF and vice versa ( $r = 0.92$ ).

Figure 1: Gherghina RiF compared to other Rules in Form Measures



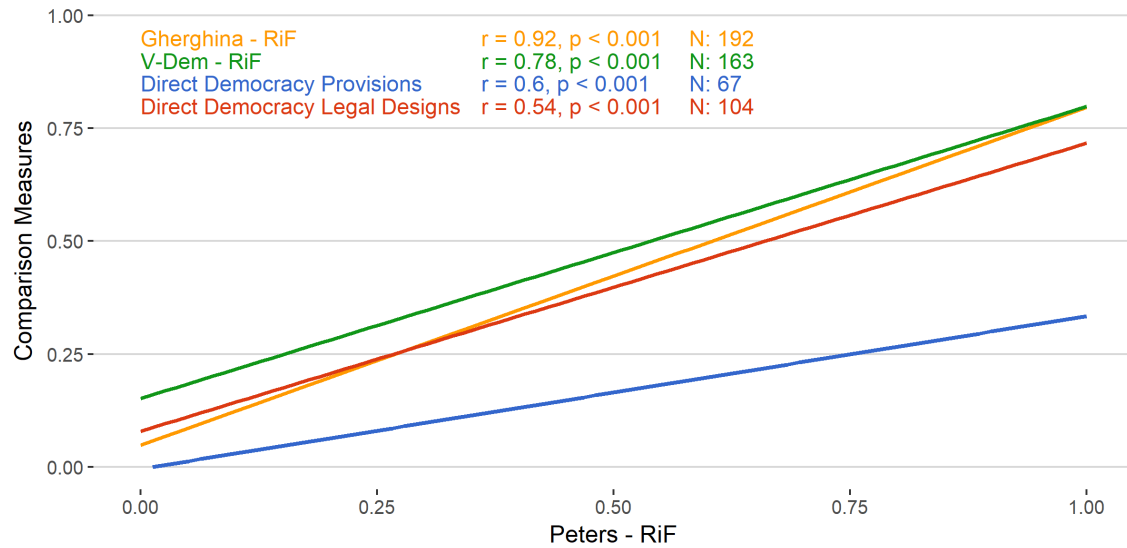
Data sources: IDEA (2018), Coppedge et al. (2017), Navigator (2018b), Barometer (2016).

This is not very surprising, as they are both reconstructed using IDEA data, even though Gherghina also considers the recall institution, while the Peters measures gives different weights to each institution and accounts for bindingness. The count measure constructed on the basis of the V-Dem mechanisms (V-Dem RiF) is also closely associated with the Gherghina index ( $r = 0.81$ ), as well as the Peters measure ( $r = 0.78$ ). The Direct Democracy Provisions measure by the Democracy Barometer is also rather strongly related to both measures, although smaller values can be observed on the respective regression line ( $r > 0.6$ ). This could be explained by the differing underlying typology, the different data collection, the fact that ease of approval is accounted for as well as the consideration of only binding mechanisms. Rather strong correlations can also be observed with the Direct Democracy Legal Designs variable, which is rather surprising given the fact that the data is partly incomplete and the completely different underlying definition of mechanisms ( $r > 0.5$ ).

To get a better overview in regard to how the indices differ across regime types, we compare their distributions visualized with density plots (Figure 3), grouped according to the Freedom House classification into free, partly free and not free countries (for the construction of the Freedom



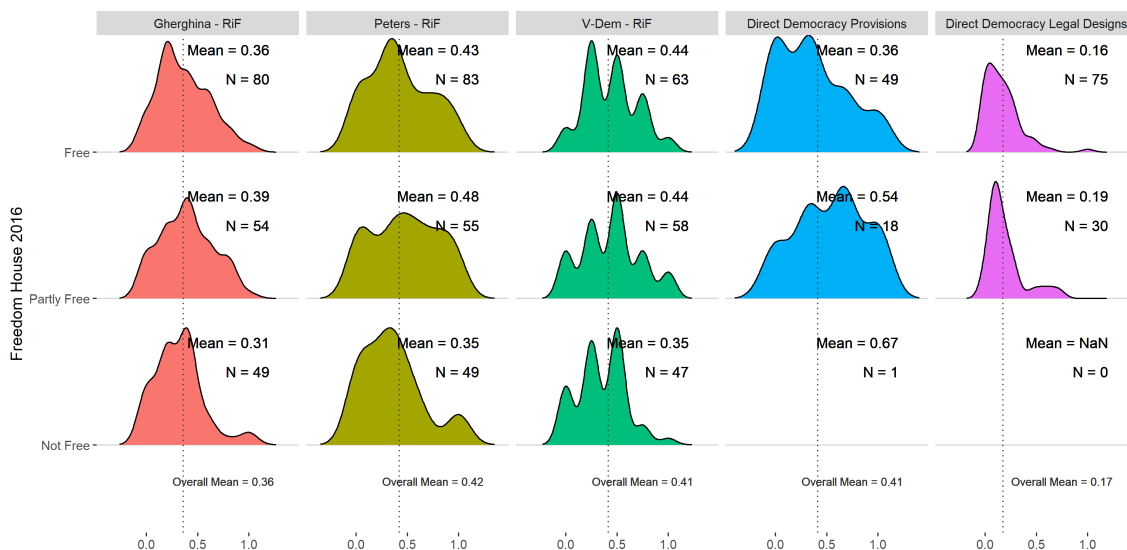
Figure 2: Peters RiF compared to other Rules in Form Measures



Data sources: IDEA (2018), Coppedge et al. (2017), Navigator (2018b), Barometer (2016).

House index and the classification see Freedom House 2018). It has to be noted that the visualized distributions are rather smooth, though in case of discrete variables the original categories can be recognized at the peaks. Most of the rules in form measures are rather normally distributed, with many countries reaching medium scores and few countries with high and low scores. Concerning not free countries (if available), the distributions become more skewed, with less countries having higher values, a pattern which appears with all of the applicable measures. It becomes visible that the simple count variable of Legal Designs is not very accurate, as on average even less countries

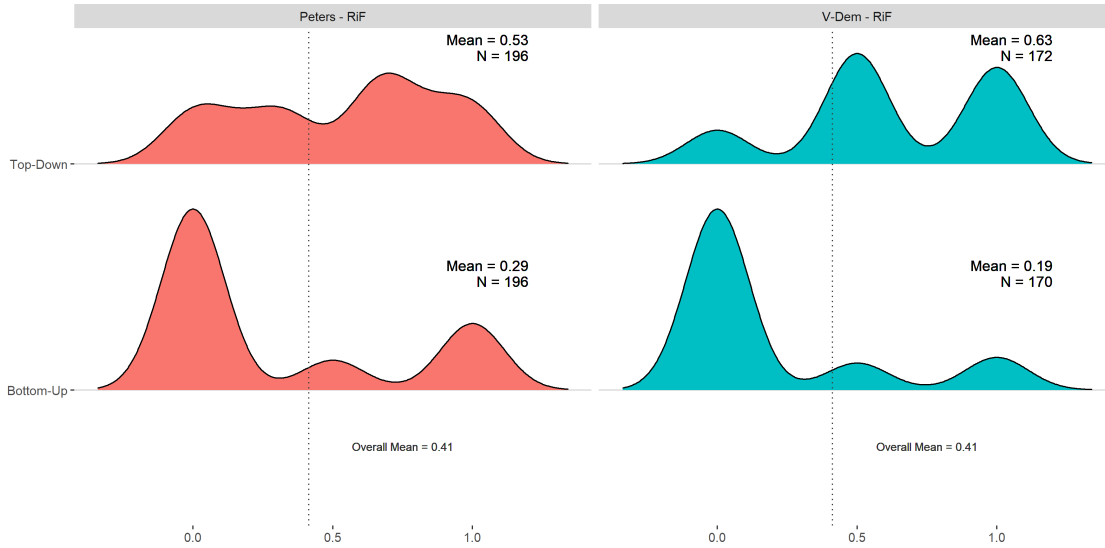
Figure 3: Rules in Form Measures by Freedom House Index



Data sources: IDEA (2018), Coppedge et al. (2017), Navigator (2018b), Barometer (2016), Freedom House (2018).

score on higher values. In general, partly free countries score as high or even higher as the ones categorized as free on all indicators. This implicates that the previously discussed importance of the level of democracy should not be neglected when comparing direct democracy across a wider range of countries. One might argue that, given equal direct popular vote mechanisms, a semi-democratic country is still less direct *democratic*, than a “full” democracy. Nevertheless, this pattern is a noteworthy discovery that should be investigated further by future research.

Figure 4: Bottom-Up and Top-Down Rules in Form Measures

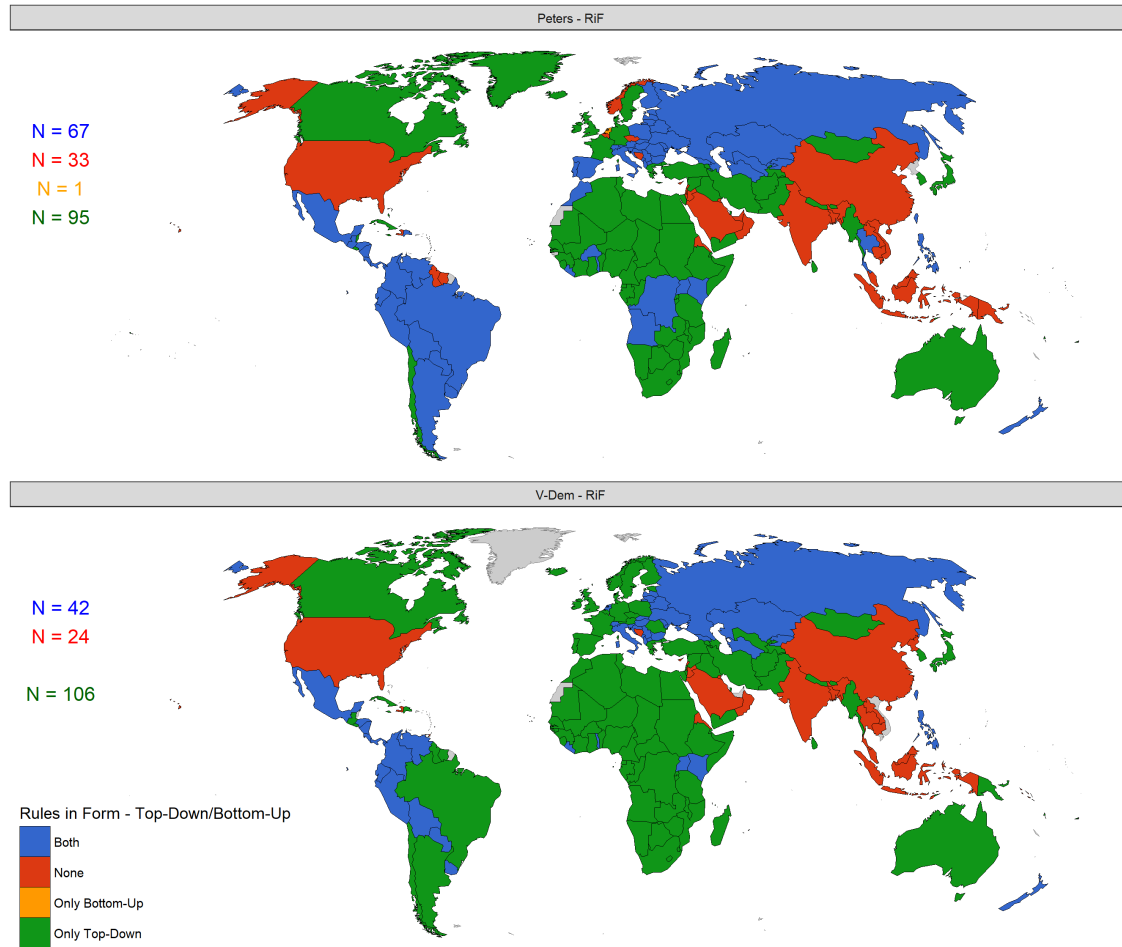


Data sources: IDEA (2018), Coppedge et al. (2017).

Next, the distributions for the two bottom-up and top-down rules in form measures are depicted separately in Figure 4. Both indicators are rather similar, as they are both constructed using V-Dem data. It can be observed, that top-down provisions for direct democracy (Peters RiF Mean = 0.29; V-Dem RiF Mean = 0.53) are by far more common than citizen initiated institutions (Peters RiF Mean = 0.19; V-Dem RiF Mean = 0.63), which was already found by Serdült & Welp (2012) pp. 77.

Figure 5 depicts a map, coloured according to a simple variable, indicating whether there are provisions only for bottom-up or top down direct democratic mechanisms, for both, or for none. Indication of such provisions is simply an index score above 0 on the respective measurement. Differences in classification are due to the consideration of different mechanisms, as well as different data sources for each measure. Here, the previous finding becomes apparent as well: a majority of countries only provide for top-down institutions (Peters - RiF N = 95; V-Dem RiF N = 106), while, depending on the coding scheme, a much smaller part has provisions for both (Peters - RiF N = 67; V-Dem RiF N = 42). Interestingly, those countries can be found mostly in Eastern Europe and Latin America. There are 33 (Peters - RiF) or 24 (V-Dem - RiF) countries that provide neither top-down nor bottom-up mechanisms. Interestingly enough, there is a single case (the Netherlands) that has only bottom-up mechanisms in place, as indicated by the Peters - RiF measure.

Figure 5: Rules in Form Measures - World Map



Data sources: IDEA (2018), Coppedge et al. (2017).

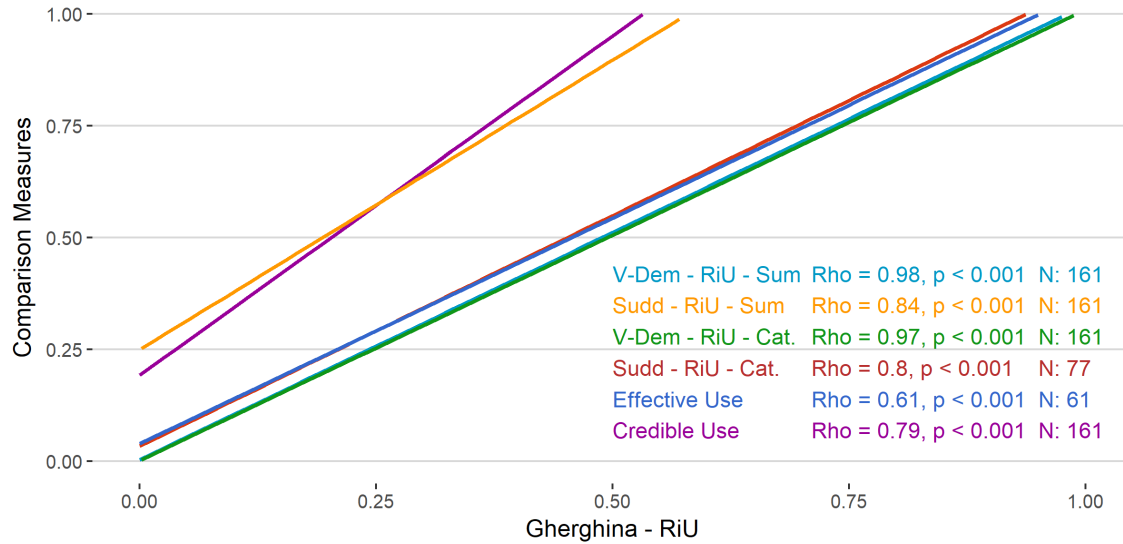
### 5.3 Comparing Rules in Use Measures

As we examine rules in use indicators, only one approach is compared with the others, namely Gherghina - RiU, the measure constructed following Gherghina (2016), as it differs the most in regard to construction (easiness of approval and decisiveness are accounted for). Since most of the indices rely on some kind of count measure, and empirically there are many countries in which direct democracy is rarely used and very few in which it is used extensively (the most prominent example being Switzerland, but also Uruguay and Azerbaijan stand out as outliers), we use the rank-based correlation coefficient Spearman's Rho to account for the heavily skewed data.

The summed up indicator based on V-Dem data shows the strongest relation to the Gherghina measure ( $\text{Rho} = 0.98$ ), which was also constructed using V-Dem data (Figure 6). All the other indicators have strong correlations as well, the least being the Effective Use count measure ( $\text{Rho} = 0.65$ ). In average, the indices seem to be in line with the Gherghina - RiU, although the Sudd - RiU

Sum as well as the Credible Use indicators are on average observed at higher values.

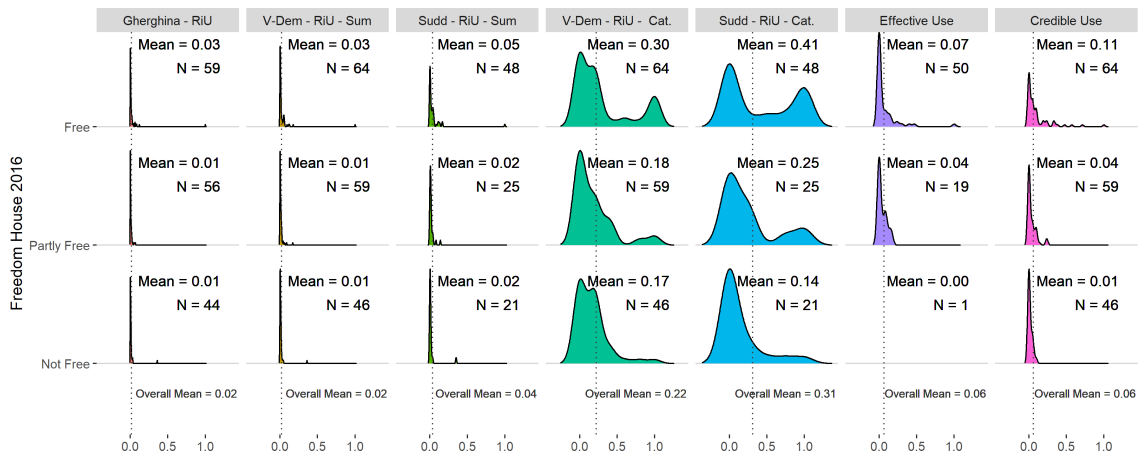
Figure 6: Gherghina - RiU compared to other Rules in Use Measures



Data sources: sudd (2018), Coppedge et al. (2017), Democracy Baromete (2016).

Figure 7 depicts the distributions for the rules in use indicators, for which it can be observed that the summed up indicators, and also the Gherghina - RiU measure are heavily skewed, with very few countries ranging very high, while most are located at the lowest scores. Due to construction, this does not apply as much to the Credible Use indicator, and even less to the categorical classifications, illustrating an important advantage of such measurement approaches, even if they inhere a loss of empirical information.

Figure 7: Rules in Use Measures by Freedom House Index



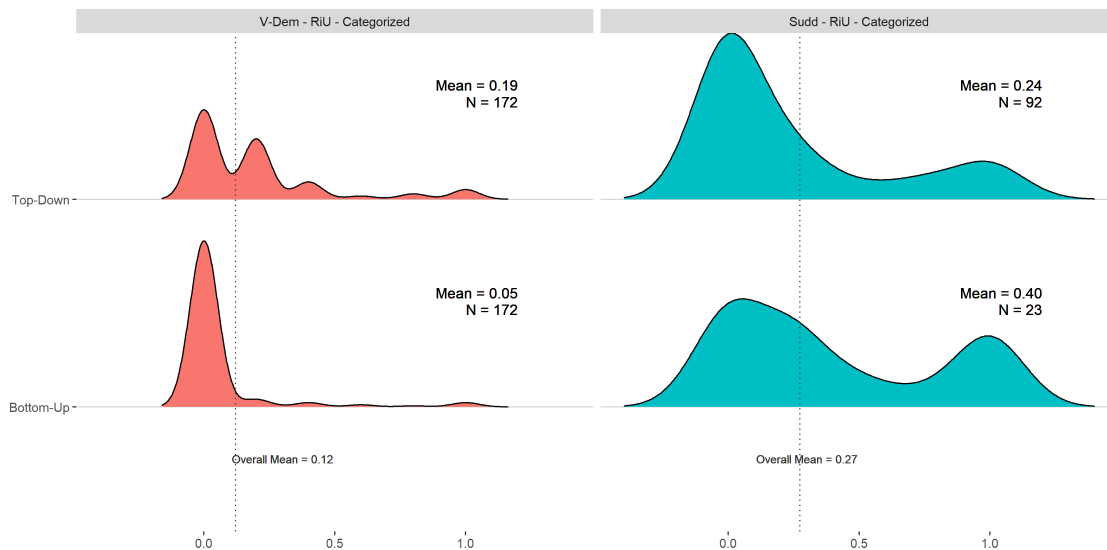
Data sources: sudd (2018), Coppedge et al. (2017), Democracy Baromete (2016), Freedom House (2018).

In Figure 8, the categorical variables for top-down and bottom-up rules in use are depicted. For the V-Dem - RiU Categorized measure, it is evident that mostly top-down mechanisms are used

in practice (Mean = 0.19), while bottom-up mechanisms remain mainly unused (Mean = 0.05), which was also observed by Serdült & Welp (2012) pp. 78. Of course, this is partially due to the fact that top-down mechanisms are much less common as rules in form, and citizens can't use bottom-up mechanisms if they are not provided. On the contrary, the Sudd - RiU Categorized measure implies an opposite trend: bottom-up usage has a mean of 0.40, while top-down usage has a mean of 0.24. However, this discrepancy arises from the structure of the sudd data, that only records direct democracy mechanisms that actually occurred. Correctly interpreted, it can be said fewer bottom-up measures have taken place ( $N = 23$ ) compared to top-down measures ( $N = 92$ ).

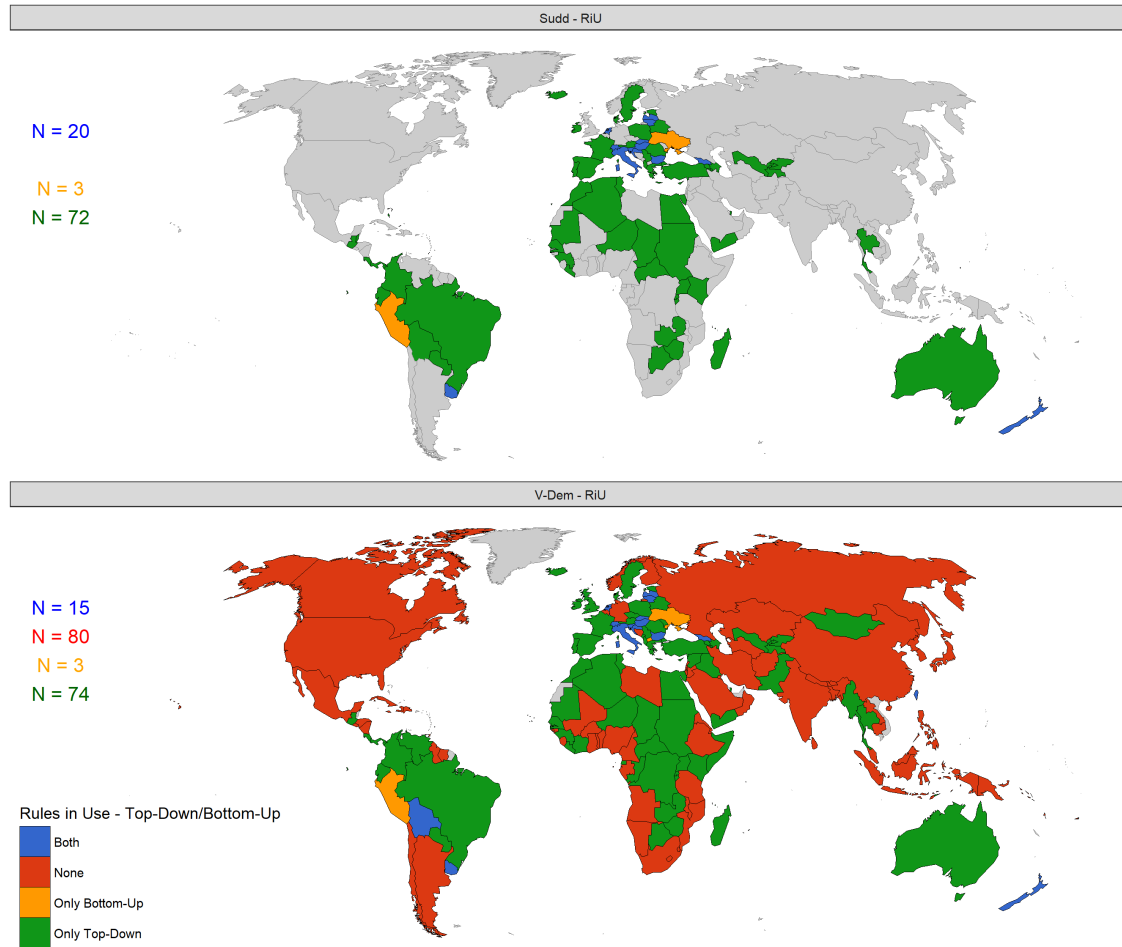
Similar to the map depicting rules in form, Figure 9 depicts a world map indicating whether, a country had at least one occurrence of only top-down or bottom up direct popular votes in between 1996 and 2016, or whether both occurred at least once. As discussed before, the sudd measures only embody countries with at least one occurrence, which becomes obvious in comparison with the V-Dem measure. Most countries are assigned to the same category, for example both measures show that Ukraine, Macedonia and Peru are the three countries where only bottom-up measures occurred. A few exceptions can also be noted, for example Bolivia, which had only a top-down occurrence according to sudd, while V-Dem counts both top-down and bottom-up rules in use. In line with the previously described results, the biggest group consists of countries that used neither bottom-up nor top-down direct democracy (V-Dem - RiU  $N = 80$ ), following such countries in which only top-down mechanisms are used (Sudd - RiU  $N = 72$ ; V-Dem - RiU  $N = 74$ ). Countries in which both mechanisms were used (Sudd - RiU  $N = 20$ ; V-Dem - RiU  $N = 15$ ) are mostly centered in Europe, while in Africa and Latin America mostly top-down mechanisms are practiced exclusively (which is not very surprising, as many countries in this regions only provide for such mechanisms).

Figure 8: Bottom-Up and Top-Down Rules in Use Measures



Data sources: sudd (2018), Coppedge et al. (2017)

Figure 9: Rules in Use Measures - World Map



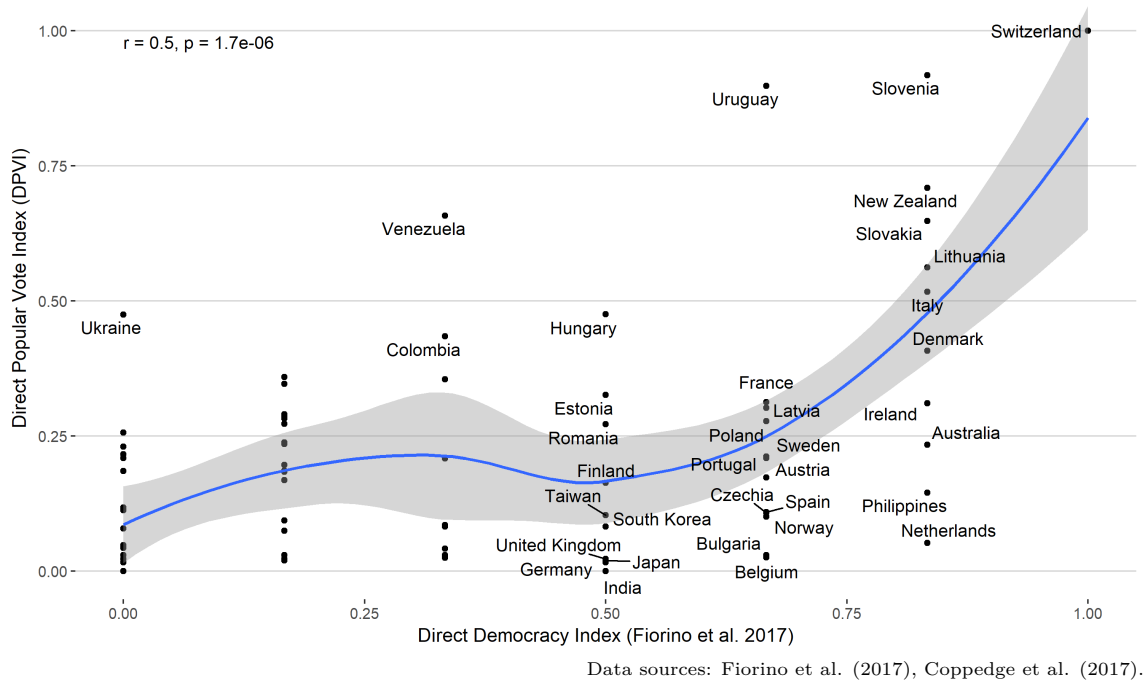
Data sources: sudd (2018), Coppedge et al. (2017).

## 5.4 Comparing Mixed Measures

As we are only assessing two mixed measurement approaches, we compare them in a more detailed way. Figure 10 depicts a scatterplot of Fiorino's Direct Democracy Index and the Direct Popular Vote Index from Altman/V-Dem. It becomes visible, that the Fiorino Index also considers the general level of democracy, while the other measure only accounts for the direct popular vote dimension. For example, countries like Ukraine or Venezuela score high on the DPVI, but not on the Direct Democracy Index. In general, there are more countries that get relatively higher scores in the Fiorino et al. (2017) measure than the DPVI, for example the Netherlands, Philippines or Belgium. As both indices capture a wide range of variables, and the DPVI is aggregated in a complex fashion while the Direct Democracy Index is a qualitative assessment, diverging scores are not very surprising. In regard to the Freedom House classification, countries labeled as free score higher on average on the Direct Democracy Index than countries of the other two Freedom House categories (see Figure

11). As expected, the DPVI shows no such pattern. On the contrary, the median is highest for the countries categorized as not free, though it has to be noted that the boxplots only depict results for the sample available for both indices, which means a proportionally large number of autocracies are not considered.

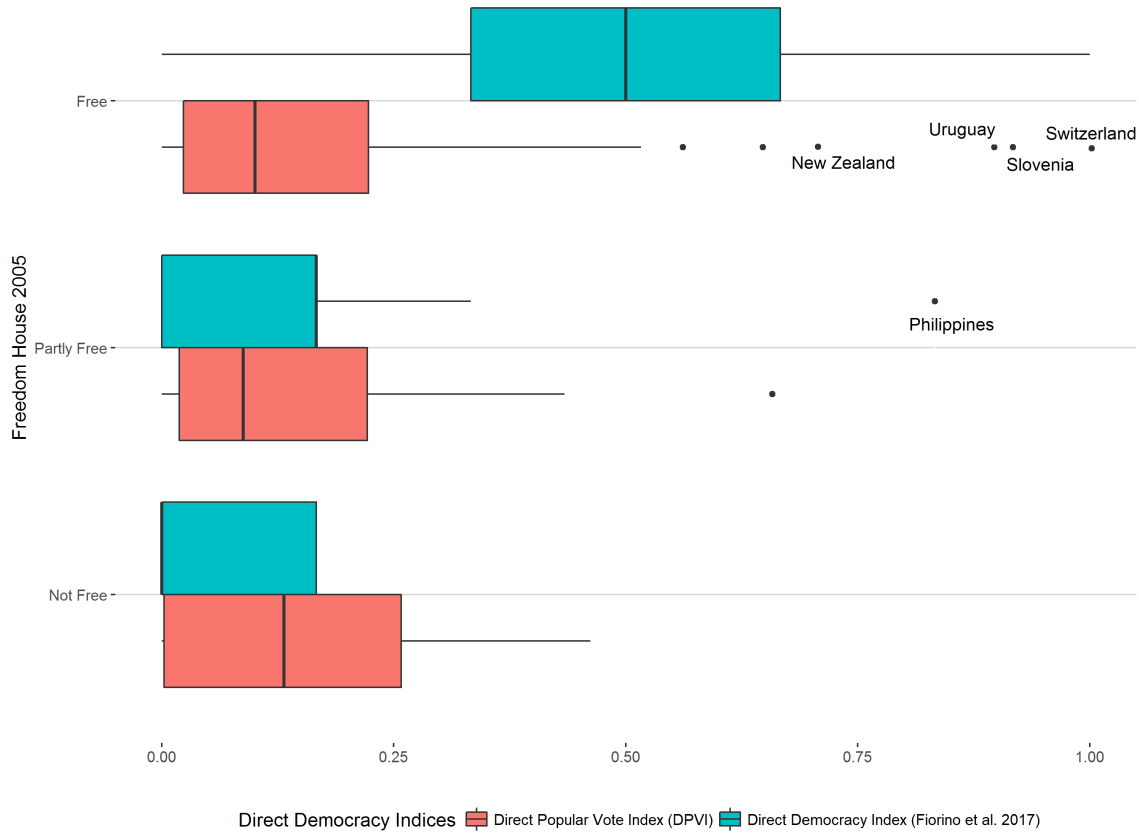
Figure 10: Scatterplot - DDI vs. DPVI



World maps for both the DDI and the DPVI are depicted in Figures 12 and 13 respectively. The Fiorino measure indicate high levels of direct democracy especially in Europe (and Australia), while the pattern for the DPVI is quite different. Here, only few countries stand out with high values (for example Switzerland, Uruguay, Venezuela), while Central Europe is shaded darker than in the Fiorino map. So, once again, it becomes visible that only the Fiorino index considers the general level of democracy.

Lastly, we take a look at the DPVI subindices for bottom-up and top-down, as depicted in Figure 14. Countries scoring high primarily on the top-down measure appear in the top-left corner, most notably Venezuela, but for example also Algeria, Denmark and Tajikistan. The countries with the highest levels of bottom-up direct popular votes are Slovenia, Switzerland and Uruguay. Comparatively low levels of top-down but high levels of bottom-up direct popular votes can be found for example in Latvia, Georgia and Ukraine. In general, it becomes visible that direct democratic institutions are primarily common in their top-down variant, with only a small number of countries scoring high on the bottom-up sub-index, and most countries having scores of zero. A correlation between the two measures implies an association between top-down and bottom-up indices, however this is mostly

Figure 11: Mixed Measures by Freedom House Index



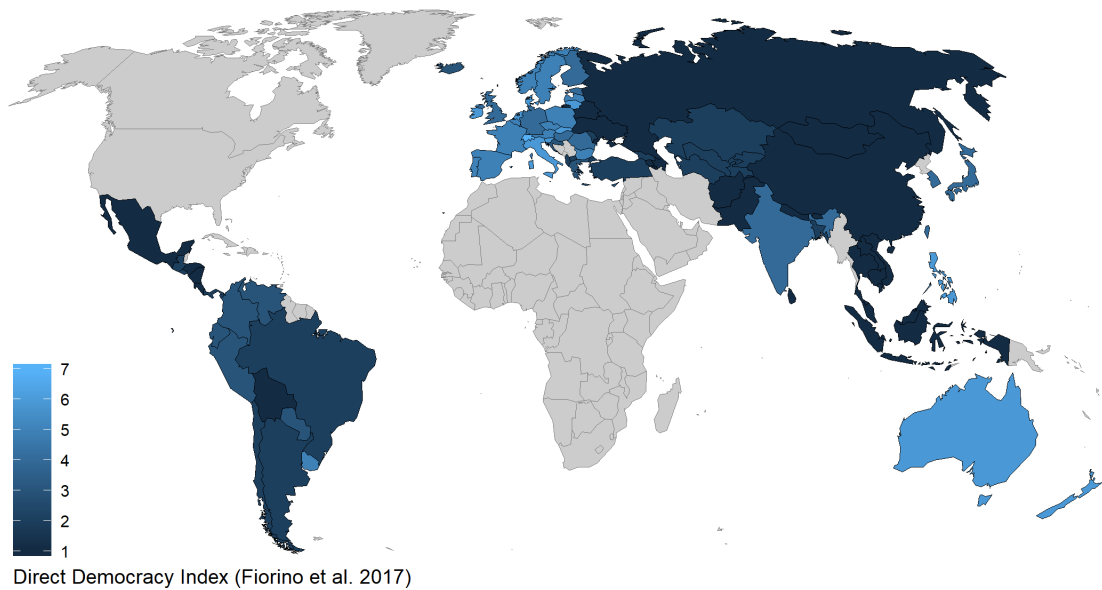
DDI: Not Free = 16, Partly Free = 22, Free = 49; DPVI: Not Free = 44, Partly Free = 56, Free = 67.  
 Data sources: Fiorino et al. (2017), Coppedge et al. (2017), Freedom House (2018).

based on a few outlier countries that were already discussed ( $r = 0.24$ ).

To sum up the empirical comparison, all the measurement approaches - even though they differ in regard to many criteria - seem to be fairly close in their respective assessments. The rules in form measurements generally correlate high with each other, especially if they are constructed based on the same data source (IDEA or V-Dem). The same applies to rules in use measurements and their respective data sources (mainly sudd and V-Dem), for which it became especially clear, that direct democracy is mainly a phenomenon existing in form, but not in practice, with the exception of a few cases (the most prominent being Switzerland). When examining the effects of rules in use, it could therefore be recommended to only assess such countries, in which at least rules in form exist. In general, but especially when examining the mixed measurements and the rather qualitative assessment by Fiorino et al. (2017), it became clear that most of the other approaches originally aim at assessing direct democracy in democratic countries or explicitly refrain from measuring direct *democracy*. As our comparison and index construction was conducted across a much wider range of countries, we found that the measurements were mostly non-sensitive to the *democracy* dimension of direct democracy. This emphasizes the already discussed importance of a

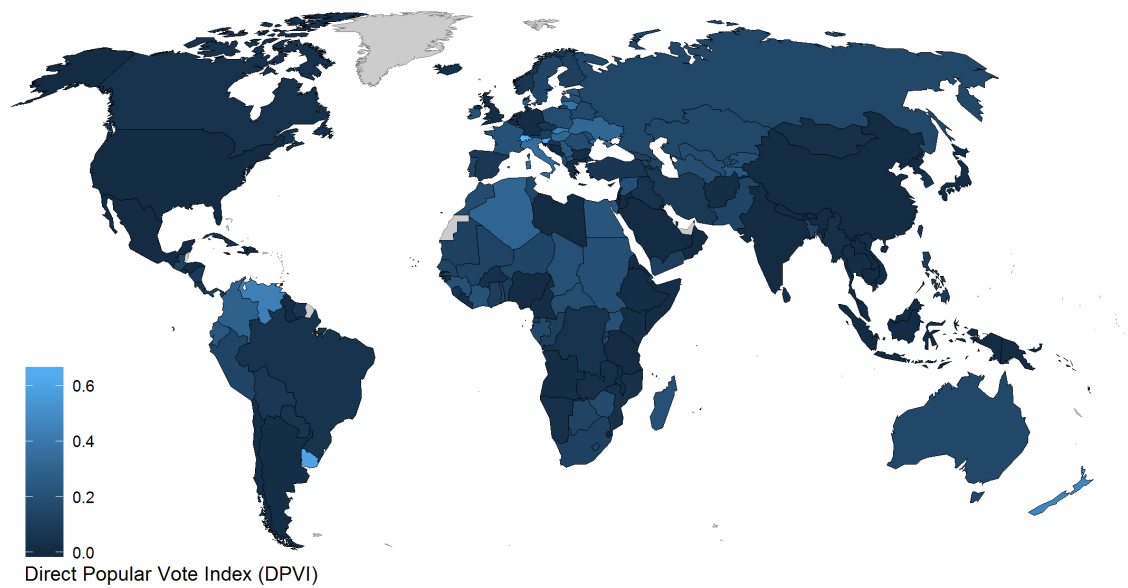


Figure 12: Direct Democracy Index - World Map



Data source: Fiorino et al. (2017).

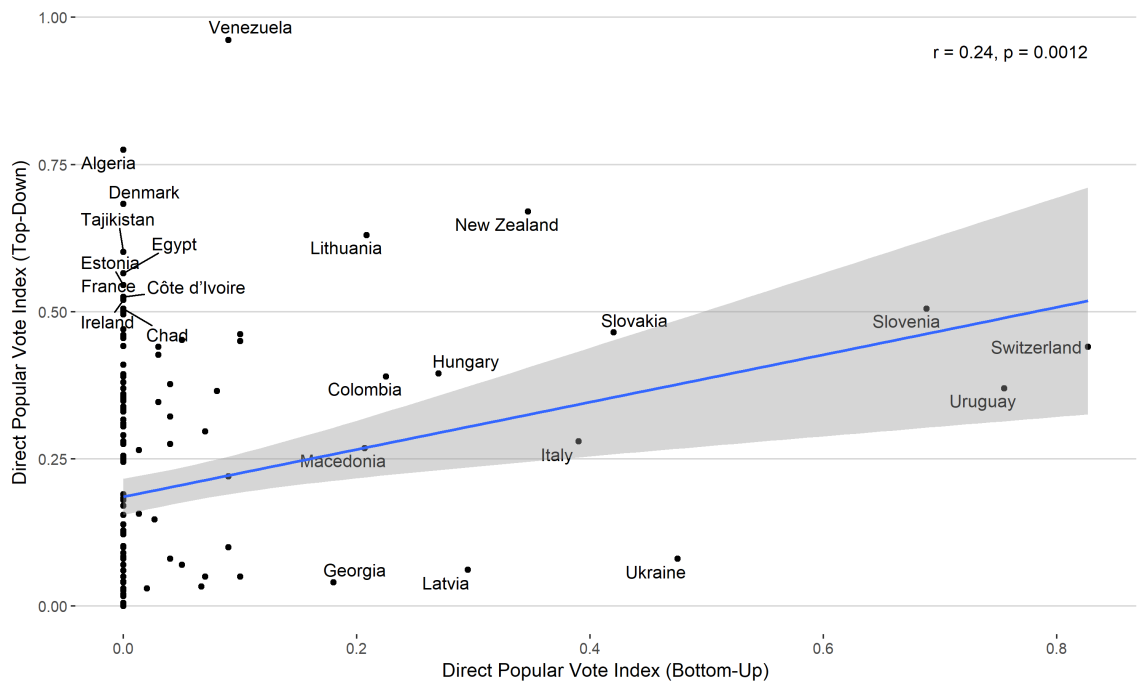
Figure 13: Direct Popular Vote Index - World Map



For comparison with the DDI, the DPVI is averaged from 2000-2005. Data source: Coppedge et al. (2017).

theoretical and methodological discussion by further research on how these two dimensions should each be measured and, especially, in turn aggregated.

Figure 14: Bottom-Up and Top-Down Components of DPVI



Data source: Coppedge et al. (2017).

## 6 Conclusion and Future Research

The aim of this work was to systematically compare selected approaches to the measurement of direct democracy theoretically and empirically, as well as to examine available data sources on direct democratic institutions in regard to their usability for quantitative research (IDEA, V-Dem, sudd (Database and Search Engine for Direct Democracy) and the Direct Democracy Navigator. To this end, we first identified relevant criteria captured in contemporary measurements and then examined different Databases for direct democratic institutions in relation to these dimensions. Next, contemporary approaches assessing direct democracy on the national level (Gherghina 2016, a subcomponent of the Democracy Barometer (Barometer 2016), Peters (2016), Fiorino et al. (2017), and Altman (2017) / ‘Varieties of Democracy’ (M. Coppedge et al. 2017)) are examined in whether they cover certain criteria, like decisiveness and easiness and if so, in which way they are accounted for in the respective index aggregation. During the examination, we differentiated between approaches assessing rules in form, rules in use and measures that include a combination of both. Another crucial differentiation is the one between bottom-up and top-down mechanisms of direct democracy, which was accounted for whenever possible.

In the empirical comparison, we found that different direct democracy measures generally overlap substantially with mostly minor differences, although it has to be noted that we only compared rules in form, rules in use, and mixed assessments within each type of measurement as this is where our focus lied. Nevertheless, some of the differences might become relevant when assessing the effects of direct democracy empirically. In general, there is no recommendation in regard to which database should be used or which measurement approach should be taken, as this relies heavily on the research question one wishes to answer. For some questions, it might be highly relevant whether direct democratic mechanisms are provided only in their top-down variants or whether bottom-up provisions exist as well, and whether these variants are used in practice. The V-Dem data provides information which is most easily accessed, with a rather extensive range of variables on rules in form, and some superficial variables for rules in use. If one is interested in more detailed information, IDEA and the Direct Democracy Navigator provide for such regarding rules in form, while sudd provides rather in depth information on actual referendums. Regarding the operationalization of direct democracy, one has to consider whether easiness and decisiveness should be accounted for as well. In regard to how these dimensions should be weighted, the discussed approaches are far from being unified, which points out an important issue needing to be assessed by further research: Which aggregation method most accurately captures the actual impact of different quorums and thresholds on rules in form and rules in use? For example, if one wishes to assess the metaphorical “Sword of Damocles” that direct democratic rules in form are said to imply, it might be highly relevant to gather information on how high/low ease of approval and initiation are.

In regard to our goal of comparing approaches in measuring direct democracy, we have some limitations that have to be addressed and discussed in regard to implications for further research. First, we chose to only examine some of the contemporary measurement approaches and data sources more thoroughly and therefore neglected others completely. Moreover, due to the scope of this work, we refrained from comparing the measurement types with each other. For example, we could have analyzed relationships between rules in use and rules in form, as Figure A1 in the appendix implies that overall there isn't a very strong correlation between them. Another possibly interesting approach would have been to conduct a factor analysis with the derived variables, to assess whether the dimensions bottom-up, top-down or rules in use and rules in form are actually present in the data, or whether specific measurements have other underlying dimensions. All of these limitations can at the same time be understood as recommendations for future research. Moreover, our study did not discuss how to aggregate direct popular vote into direct *democracy* measurements, when assessing a broader scope of countries including autocracies, which is also a topic worthy of further discussion. In this context, it could also be interesting to group top-down or bottom-up measures per regime type, as it would be plausible to expect different distributions between autocracies and democracies. A limitation of this study, but also one of many contemporary measurement approaches, is the neglect of two important factors: subnational democracy and the content and scope of issues for which direct democratic mechanisms are allowed/excluded, mostly because their assessment is an even more difficult task than the one of easiness or decisiveness. An important recommendation for further research is to arrive at a more unified, but also detailed scheme of classifying direct democracy institutions, as well as a common terminology as to ensure comparability and reproducibility between different kinds of research.

## References

- Altman, David 2011:** *Direct democracy worldwide*. Durham, North Carolina: Cambridge University Press.
- Altman, David 2017:** The potential of direct democracy: A global measure (1900–2014), *Social Indicators Research* 133, pp. 1207–27.
- Auer, Andreas & Bützer, Michael 2001:** *Direct democracy: The eastern and central european experience*. Aldershot: Ashgate Pub Limited.
- Barometer, Democracy 2016:** Data, Available at: [http://www.democracybarometer.org/dataset\\_en.html](http://www.democracybarometer.org/dataset_en.html).
- Bauer, Paul C & Fatke, Matthias 2014:** Direct democracy and political trust: Enhancing trust, initiating distrust—or both?, *Swiss Political Science Review* 20, pp. 49–69.
- Beramendi, Virginia et al. 2008:** *Direct democracy: The international idea handbook*. IDEA, International Institute for Democracy; Electoral Assistance.
- Bernard, Laurent & Bühlmann, Marc 2014:** Beeinflusst die direkte Demokratie das Wirksamkeitsgefühl von BürgerInnen?, in Münch, Ursula and Hornig, Eike-Christian and Kranenpohl, Uwe, ed. *Direkte demokratie. analysen im internationalen vergleich*, Baden-Baden: Nomos.
- Blume, Lorenz, Müller, Jend & Voigt, Stefan 2007:** Direct democracy and political trust: Enhancing trust, initiating distrust—or both?, *CESifo Working Paper No. 2149*.
- Butler, David & Ranney, Austin 1978:** *Referendums: A comparative study of practice and theory*. Washington, DC: American Enterprise Institute.
- Clark, William Roberts, Golder, Matt & Golder, Sona Nadenichek 2017:** *Principles of comparative politics*. 3rd ed., CQ Press.
- Coppedge, Michael et al. 2017:** V-dem country-year data, version 7.1, *Varieties of Democracy (V-Dem) Project*. Available at: <https://www.v-dem.net/en/data/data-version-7-1/>.
- Coppedge, Michael et al. 2017:** V-dem codebook v7, *Varieties of Democracy (V-Dem) Project*.
- Fiorino, Nadia, Ricciuti, Roberto & Venturino, Fulvio 2017:** Measuring direct democracy, in *State, institutions and democracy*, Springer.
- Freedom House 2018:** *Freedom in the world 2018*. New York. Available at: <https://freedomhouse.org/report/freedom-world/freedom-world-2018>.
- Gherghina, Sergiu 2016:** Direct democracy and subjective regime legitimacy in europe, *Democratization* 24, pp. 613–31.
- Hwang, Jau-Yuan 2005:** *Direct democracy in asia: A reference guide to the legislations and practices*. Taipei: Taiwan Foundation for Democracy.
- Kaufmann, Bruno 2004:** Initiative and referendum monitor 2004/2005, *Initiative and Referendum Institute Europe*.
- Kaufmann, Bruno & Waters, M. Dane 2004:** *Direct democracy in europe: A comprehensive guide to the initiative and referendum process in europe*. Durham, North Carolina: Carolina Academic Press.
- Lissidini, Alicia 2011:** Democracia directa en latinoamérica: Entre la delegación y la participación,
- Madroñal, Juan Carlos 2005:** Direct democracy in latin america, *Mas Democracia and Democ-*

racy Internatiuonal.

**Merkel, Wolfgang & Ritzi, Claudia 2017:** Fazit, in *Die legitimität direkter demokratie*, Springer.

**Merkel, Wolfgang et al. 2016:** Democracy barometer: Codebook for core set (blueprint) and extended sample, version 4, Available at: [http://www.democracybarometer.org/documentation\\_en.html](http://www.democracybarometer.org/documentation_en.html).

**Navigator, Direct Democracy 2018:** Typology, Available at: <http://www.direct-democracy-navigator.org/typology>.

**Peters, Yvette 2016:** (Re-) join the party! The effects of direct democracy on party membership in europe, *European Journal of Political Research* 55, pp. 138–59.

**Pogrebinschi, Thamy 2015:** Mehr partizipation- ein heilmittel gegen die ‚krise der demokratie’?, in *Demokratie und krise*, Springer.

**Serdült, Uwe & Welp, Yanina 2012:** Direct democracy upside down, *Taiwan Journal of Democracy* 8.

**Stadelmann-Steffen, Isabelle & Vatter, Adrian 2012:** Does satisfaction with democracy really increase happiness? Direct democracy and individual satisfaction in switzerland, *Political Behavior* 34, pp. 535–59.

**Zovatto, Daniel 2015:** Las instituciones de la democracia directa, *Revista de Derecho Electoral*, pp. 34–75.

Figure A1: Heatmap of all Variable Used



Table A1: Description of Direct Democracy Index - Kaufmann

Score	Description
1.) <i>The Radical Democrat</i>	Citizens have access to a broad spectrum of direct-democratic procedures. As well as the binding popular initiative, these include the right of facultative referendum and obligatory referendums for alterations to the Constitution and state treaties.
2.) <i>The Progressive</i>	Citizens have, at least in part, the possibility of initiating national referendums without the express permission of the organs of the state (parliament, government, president). There are also procedures for obligatory referendums.
3.) <i>The Cautious</i>	The electorate does have practical experience of popular initiatives and /or national referendums. But these procedures are essentially plebiscitary in nature, i.e. they are not protected or controlled by the citizens themselves or by the law, but are controlled "from above" by parliament (political parties) or by the executive.
4.) <i>The Hesitant</i>	The political elites in the countries of this category appear to be afraid of popular participation in political decision-making, whether out of fear of having to share power or because of concrete historical experiences. Even here, however, there are still some traces of statutory I&R procedures, which may form the basis for future improvement .
5.) <i>The Fearful</i>	Almost entirely lacking institutional procedures and practical experience, the countries in this category make it very hard for themselves to complement indirect democracy. In addition, the political and cultural circumstances scarcely provide a stimulus for the introduction or the strengthening of elements of popular decision-making. Nonetheless, the issue is occasionally debated.
6.) <i>The Beginners</i>	These countries have only recently started their democratization process, including a respect for basic freedoms and human rights. Parliaments have been elected by the people, but there is still a great deal of mistrust between governments and governed, making the introduction of additional instruments like direct democracy extremely difficult. The Authoritarians: In the countries belonging to this category, there is at present no basis at all for the development of direct democracy.
7. <i>The Authoritarians</i>	In the countries belonging to this category, there is at present no basis at all for the development of direct democracy.

Source: Kaufmann (2004)