Measuring Effective Democracy

Carl Henrik Knutsen

International Political Science Review 2010 31: 109
DOI: 10.1177/0192512110364736

International Political Science Review
31(2) 109–128

DOI: 10.1177/0192512110364736
http://ips.sagepub.com

Measuring Effective Democracy

Carl Henrik Knutsen

Abstract

This article discusses methodological problems related to operationalizing substantive definitions of democracy. The article argues that index-constructors need to be particularly conscious of measurement level issues. If not, their indexes may face severe reliability and validity problems, which in turn may bias empirical analyses utilizing the indexes. The article focuses particularly on the "effective democracy" measure developed by Ronald Inglehart and Christian Welzel. The measure has been used by Inglehart and Welzel in several studies, particularly for empirically testing hypotheses deduced from their version of modernization theory. These tests have generated very strong results in favor of the theory. The article is sympathetic to Inglehart and Welzel's goal of capturing "substantive" rather than "formal" democracy, but is critical of the specific measure proposed. The measure has several unfortunate theoretical and distributional properties; the empirical scores generated by the measure are often highly misleading. Empirical analysis suggests the index is biased, and that rich, Western countries are particularly favored. Utilization of the measure in statistical analysis may therefore lead to false inferences.

Keywords

Democracy, effective democracy, measurement, index

1. Introduction

The causes and consequences of democracy are among the most studied subjects in the social sciences. No consensus exists, however, on how to define democracy, let alone on how to measure it. There is an ongoing debate on the effects of "mass attitudes" on democratization (Hadenius and Teorell, 2006; Inglehart and Welzel, 2005; Welzel, 2007; Welzel and Inglehart, 2006). Linked to that substantial debate is a methodological debate on how to measure democracy. In particular, there is disagreement over whether Inglehart and Welzel's measure of effective democracy, here labeled the effective democracy index (EDI), is a valid or even meaningful measure. EDI is constructed by multiplying the Freedom House Index (FHI) by a measure of corruption, either Transparency International's Corruption Perception Index (CPI) or the Control of Corruption index (CCI) from the World Governance Indicators (WGI). The EDI, by including lack of corruption as

Corresponding author:
Carl Henrik Knutsen, Department of Political Science, University of Oslo,
Boks 1097 Blindern, 0317 Oslo, Norway
[email: c.h.knutsen@stv.uio.no]

an indicator of democratic quality, is supposed to measure not only the existence but also the functioning of democratic institutions. According to Welzel (2007: 421), "the dispute about the usefulness of measuring effective democracy" is still unresolved. This article analyzes the meaningfulness, reliability, and validity of the EDI. The main conclusion is that the EDI is beset with serious methodological problems. We should therefore view empirical studies of democracy's causes and consequences based on this measure with skepticism. In addition to analyzing the EDI, this article contributes generally to the literature on how to define and measure democracy, and to the methodology of index construction.

In Section 2.1, I present the debate on how to define democracy, and I discuss narrower, institutional definitions of democracy and broader, substantive definitions. In 2.2, I argue that the EDI is a *meaningful* operationalization of democracy defined substantively, but not democracy defined institutionally. In 2.3, I describe the FHI and the EDI. In 3.1, I comment on some benefits of the EDI. In 3.2 and 3.3, however, I show that the EDI's reliability and validity problems are grave. In 3.2, I present five concrete methodological problems with the EDI. In 3.3, I show how these problems affect particular empirical EDI scores, with a focus on rank and distance between cases. I also use statistical analysis to show the EDI's systematic biases: rich, Western, and Protestant countries achieve undeservedly high EDI scores. In the conclusion, I show how the EDI can be adjusted to mitigate some of the methodological problems. Nevertheless, the FHI might still be a better operationalization of substantive democracy in quantitative studies.

2. Defining and Measuring Democracy

2.1 The Conceptual Debate: What Is Democracy?

According to Schumpeter (1976: 250):

> [T]he eighteenth century philosophy of democracy may be couched in the following definition: the democratic method is that institutional arrangement for arriving at decisions which realizes the common good by making the people itself decide issues through the election of individuals who are to assemble in order to carry out its will.

Schumpeter's critique of this "classical definition" of democracy has proved to be very influential. His main point was that defining democracy in terms of institutions' ability to secure the implementation of a "general will" or a "common good" is unsatisfactory, because these constructs lack a meaningful reference. Schumpeter pointed out that for "different individuals and groups, the common good is bound to mean different things" (1976: 251). Some years later Kenneth Arrow (1951) showed the impossibility of aggregating individual fixed and well-defined preferences to a determinate, well-defined collective preference, when the number of individuals and issue dimensions increase sufficiently. One implication of this theoretical proof was that democracy viewed as an implementation of a "general will" was nonsensical.

Schumpeter's alternative was defining democracy, or more precisely the democratic method, as the "institutional arrangement for arriving at political decisions in which individuals acquire the power to decide by means of a competitive struggle for the people's vote" (1976: 269). In other words, Schumpeter bypassed the issue of a popular will by referring directly to an institutional mechanism, namely competitive elections. He thereby escaped the problems of fuzziness, indeterminacy, and even emptiness of classical democracy definitions. More recently, Adam Przeworski has been a vocal defender of a "minimalist," institutionally based democracy definition. Przeworski et al. (2000), for example, define democracy as a political regime that holds contested elections. The comparative benefits of a minimalist, institutionally based definition are analytical stringency, precision, and clarity (Przeworski et al., 2000).

At least two powerful criticisms can be made of the "Schumpeter–Przeworski" view of democracy. The first criticism takes the institutional point of departure as given, but argues that proponents of the Schumpeter–Przeworski view commit a "fallacy of electoralism" (Diamond, 1999: 9). The argument is that elections are not sufficient for securing a democratic political regime. One needs additional institutional guarantees, and therefore the democracy definition needs to be broadened. One paradigmatic list of such "institutional guarantees" can be found in Robert Dahl's *Polyarchy* (1971: 3), which includes, for example, freedom of expression and alternative sources of information.

The second criticism of the electorally based definition disagrees with the underlying premise of defining democracy institutionally, and takes the role of the populace in political decision-making as the point of departure. It is problematic, according to Beetham (1999), to claim that democracy merely *consists* of a matrix of certain rights, liberties, and institutions. The important question is *why* particular institutions and rights are considered democratic. Any answer risks ending up in a tautologous argument: "The only way to avoid circularity is by specifying the underlying principles which these institutions embody or help to realize, and in terms of which they can plausibly be characterized as democratic" (Beetham, 1999: 90). Beetham claims that "the core idea of democracy is that of popular rule or popular control over collective decision making" (1999: 90), and adds political equality as a second criterion. Institutions are considered democratic *only* if they contribute to realizing popular control over politics and political equality. One can still be interested in particular institutions when studying democracy, but these are only the instruments that underpin democracy. The argument for a substantive-democracy definition is strong.

One should notice that this definition does not refer to a "general will" or to aggregate preferences of the collective, but only to the populace's *opportunities* for exerting actual influence in public decision-making, and the distribution of these opportunities among citizens. By not referring to a "general will," the definition therefore evades some of Schumpeter's criticisms of classical democracy definitions. However, the "vagueness" and "indeterminacy" criticisms still apply to a certain degree. What does "popular control" actually mean, and who is the relevant demos? Nevertheless, even though we cannot make precise scales for democracy in the Beethamian sense that are *perfectly* valid and reliable, we should not refrain from trying to construct *good* measures.

2.2 Democracy Definitions and Implications for the Effective Democracy Index

The choice of conceptual democracy definition has important implications for the meaningfulness of the EDI. If one defines democracy institutionally, it is problematic to make extra claims about what constitutes "effective democracy" within the limits of the definition. Effectiveness can, of course, be defined as an independent concept, and one could measure whether democratic institutions were empirically effective in providing, for example, economic equity or growth. The mere existence, or possibly the degree of existence, of one or more institutions is, however, the sole criterion looked at when *defining* democracy. But if one defines democracy substantively, the task of operationalizing effective democracy is not only sensible but also vital. A first approximation to operationalizing "popular control over politics combined with political equality" can be made by looking at the empirical existence of certain formal institutions such as elections and constitutional guarantees of freedom of speech and association, which have historically proved to be efficient in fulfilling these criteria. However, effective democracy can now plausibly be interpreted as a measure that also incorporates how these

institutions actually *function* in translating the demands of broad population segments into legislation and ultimately into implemented policies.

Viewed in this light, effective democracy is *not* a combination of two analytical concepts; it is rather a *better* empirical approximation to the theoretical definition than a "formal" measure. As Dahl (1971) suggested, democracy should be viewed as an ideal concept. Egalitarian distribution of power over collective decision-making among citizens provides the proper ideal. The empirical possibility of relatively equal political influence hinges not only on the existence of formal institutions, but also on how these institutions actually function in practice when transforming demands and other "inputs" into political outcomes. "Effective democracy" therefore becomes something of a misnomer, since it is actually here intended as a more comprehensive and better measure of democracy defined substantively.

The need to take the actual functioning of rights and institutions into account when measuring democracy might have become even more urgent in the post-Cold War world. Przeworski et al. partly legitimize their minimalist theoretical definition and consequent operationalization by pointing out that: "[D]ifferent views of democracy, including those that entail highly subjective judgments, yield a robust classification" (2000: 57). They find that their measure and the FHI have a correlation of .94. However, their analyses end in 1990 and, as Diamond (1999: 286) notes, there might be a larger divergence between the formal properties of institutions and the actual democratic character of political regimes after the Cold War. Inglehart and Welzel note, referring to the new democracies of the "third wave," that "many of the new democracies show severe deficiencies in their actual practice of civil and political liberties" (2005: 149). Elections are no longer (if they ever were) a sufficient proxy for substantive democracy. The construction of a broad democracy indicator is therefore of the utmost importance.

2.3 Two Operationalizations of Substantive Democracy

I. The Freedom House Index. The FHI is one of the most utilized democracy indicators in contemporary academic research. Diamond calls it the "best available empirical indicator of liberal democracy" (1999: 12). Freedom House scores countries on a scale from 1 to 7 (where 1 is the most democratic) on two dimensions: political rights (PR) and civil liberties (CL). The FHI is an average of PR and CL. Both PR and CL are indices constructed from large subsets of indicators, which are formulated as "check questions." Freedom House scores its indices (primarily) on the basis of 25 such questions, each of which has several sub-questions. Ten of the check questions relate to political rights, and fifteen to civil liberties. The FHI is, to a large degree, based on the presence or absence of different institutions, but it also seeks to account for the actual functioning of these institutions (Freedom House, 2005: 775). Therefore, this operationalization is presumably valid for substantive democracy definitions.

II. The Effective Democracy Index. If one adheres to a substantive definition of democracy, one important task is to look for *systematic* components affecting democracy other than formal institutions and rights. Power dispersion and the functioning of institutions in transforming popular demands to collective decisions are hard to observe and evaluate directly. However, Inglehart and Welzel make a strong case for one component affecting the actual functioning of institutions: the absence of corruption and the practice of the rule of law more generally. The idea is that, in order to make them effective, "civil and political rights also require honest, law-abiding elites. Corruption, broadly understood, reflects nepotism, favouritism, and other illegal mechanisms used by elites to

circumvent the rule of law and use their power for their own benefit, depriving ordinary people of their legal rights" (Inglehart and Welzel, 2005: 192). Inglehart and Welzel therefore distinguish between formal democracy and effective democracy, where formal institutions and rights are combined with the relative absence of corruption. The authors operationalize effective democracy by multiplying the FHI by a measure of corruption. The authors have used different corruption indices, with the CPI used in Welzel et al. (2003) and the CCI in the study from 2005. After reversing and normalizing the FHI to a measure ranging from 0 to 100 (A), and normalizing the corruption score to a measure ranging from 0 to 1 (B), the EDI is given as A*B. The EDI is intended as a proxy for "how much freedom people actually have rather than how much freedom they have on paper" (Inglehart and Welzel, 2005: 196).

3. Evaluating the Effective Democracy Index

3.1 Benefits of the Effective Democracy Index

Inglehart and Welzel have described the EDI and also defended it against criticism (Inglehart and Welzel, 2005; Welzel and Inglehart, 2006). They provide good discussions on the benefits of this measure, and I will therefore only provide a quick survey. As mentioned, the EDI is explicitly constructed to tap the actual functioning of institutions and individual rights, as well as their form. If one defines democracy substantively, the explicit focus on how institutions function is a great benefit for the operationalization of the definition.

Equating free elections and freedom of speech with political equality and popular control over collective decision-making in its strictest sense is problematic. There are several examples of what is referred to by some as "democratic deficits" (Levinson, 2007: 260) in OECD countries. Italy is often accused of clientelism (in the south) combined with a not fully independent media (Beetham et al., 2002). Japanese politics has been labeled elite driven and oligarchic (Johnson, 1995). American politics is perceived to be heavily influenced by organized business interests and other narrow lobby groups (Olson, 1982). However, the most obvious discrepancies between "formal" and "effective" democracy, to use Inglehart and Welzel's terms, are in the developing world. Clientelism is a particular political trait that seems to strip formal democratic institutions of their purpose. Médard (1996) claims that African politics is characterized by "neo-patrimonialism": formal state institutions exist, but these are not the arenas where the actual political action happens. Politics is, rather, managed through vertical, personalized ties in a "patron–client" fashion. Chabal and Daloz (1999) claim that these structures and practices make democratic institutions function badly in Africa. Another example is Philippine politics. Politics in the country is, according to Sidel (1999), characterized by "bossism," and the abuse of electoral practices is widespread. Landowners, in particular, have used means not recognized as true to democratic ideals to maintain electoral office. Generally, when elites manipulate democratic practices, neglect the rule of law, and undermine ordinary citizens' rights by corrupt practices, formal democratic institutions and rights are rendered ineffective. These examples indicate the benefit of incorporating "quality considerations," such as lack of corruption, into operationalizations of democracy.

There is another benefit of the EDI when compared to, for example, the FHI. The FHI, and other indicators like Polity, have been accused of differentiating too little between regime types, and especially between the most democratic regimes. Several empirical cases are assigned the highest score on the FHI. The EDI produces more finely graded scores, and it diversifies more at the top end of the scale. Inglehart and Welzel explicitly recognize this comparative benefit of the EDI: "If we took the formal democracy measure at face value, we would conclude that Latvia and Slovakia are

just as democratic as Britain and Germany. But in reality they have relatively corrupt elites who devalue the constitutional rights their people theoretically possess" (2005: 195).

3.2 Five Validity and Reliability Problems with the Effective Democracy Index

The EDI has been heavily criticized by Axel Hadenius and Jan Teorell (2005) on several counts. One criticism is that Inglehart and Welzel are wrong in multiplying two different concepts into one measure, claiming that "the mistake, thus, is to confuse basic and quality criteria" (2005: 90). Welzel and Inglehart (2006) answer by pointing out that both lack of formal institutions and rights *and* deficiencies in the actual practice related to the protection of guaranteed rights and liberties can deprive people of their actual rights and liberties. I would go even further than Inglehart and Welzel, who claim that the combination of "two different things" can produce a very sensible measure: both formal institutions and corruption relate to the same underlying concept, namely democracy defined substantively. Therefore, I am in principle sympathetic to the approach taken by Inglehart and Welzel, which is an attempt to incorporate a quality dimension when measuring democracy properly conceptualized. However, Hadenius and Teorell (2005) produce other criticisms that are relevant to the validity of the EDI. I will summarize some of these criticisms and elaborate upon them. I will also present some novel criticisms that cast grave doubts on the reliability and validity of the measure.

I. The Freedom House Index as a Partially "Quality-Oriented" Index. One of the most controversial claims made by Inglehart and Welzel is that the FHI measures *formally* guaranteed political rights and civil liberties. This assumption stands in stark contrast to the explicit intention of Freedom House to measure "the real-world rights and freedoms enjoyed by individuals" (2005: 775). Hadenius and Teorell criticize Inglehart and Welzel for assuming that the FHI only limits itself to formal institutions and rights, since it also incorporates "actual upholding of such rights" (2005: 90). Welzel and Inglehart (2006) in their explicit response agree that the FHI incorporates whether rights and liberties are free from *obvious* violations. Nevertheless, they still claim that the FHI is insensitive to more *subtle* violations. Table 1 lists some of the FHI check questions and sub-questions that capture quality aspects (Freedom House, 2005: 780–2).

Table 1. FHI Questions That Capture Democratic Quality Aspects

Index	Item	Question
Political rights	B3	"Are the people's political choices free from domination by the military, foreign powers, totalitarian parties, religious hierarchies, economic oligarchies, or any other powerful group?"
	C2	"Is the government free from pervasive corruption?"
	C2 sub-question	"What was the latest Transparency International Corruption Perceptions Index score for this country?"
	C3	"Is the government accountable to the electorate between elections, and does it operate with openness and transparency?"
Civil liberties	D1	"Are there free and independent media and other forms of cultural expression?"
	D1 sub-question	"Is self-censorship among journalists common, especially when reporting on politically sensitive issues, including corruption or the activities of senior officials?"
	F5	"Does the rule of law prevail in civil and criminal matters? Are police under direct civilian control?"

There are also examples of quality-oriented questions in the FHI other than those listed in Table 1. The belief that these "qualitative" components do not affect the overall grading of the FHI seems unwarranted. The quality-oriented questions make up a decent proportion of the FHI, and questions such as C2, D1, and F5 seem likely to tap subtle, as well as obvious, violations of rights. If one still has doubts, the country studies gathered in the Freedom in the World reports show that several countries score lower than their formal institutions and rights would suggest. Jamaica, for example, scored 2.5 in 2005 and was penalized mainly because of political violence and corruption (Freedom House, 2005). India was actually averaging 4.0 from 1993 through 1995, showing that formal democracies are penalized for dismally functioning institutions and rights.

II. Additive and Multiplicative Measures and Implications for Rank. One of the concerns of Hadenius and Teorell is that certain quality criteria can compensate for the lack of "basic" democratic criteria in the EDI. Their fear is that relatively non-corrupt authoritarian regimes will get higher scores on the EDI than corrupt democratic regimes, thereby reversing the "correct" rank between countries (2005: 89). Welzel and Inglehart (2006) reply that this also could have been the case if one were using an additive measure. However, Inglehart and Welzel point to the multiplicative nature of their measure. They claim that "high scores in effectiveness cannot compensate for low scores in democracy," since effectiveness is only a weighting factor of a given level of democracy. The mathematical intuition is simple. Zero multiplied by something is zero, and a small number multiplied by one is the small number. However, even if we take the "multiplicative weighting argument" at face value, there arises a methodological problem for the EDI because of the way the FHI is constructed. As we saw above, a decent number of the FHI's check questions are "quality" oriented. Now denote FHI as A and CPI as B. These indices are *intended* by Inglehart and Welzel to respectively measure real-world "formal institutions and rights," X, and "quality or actual functioning," Y. X and Y have lower bounds of 0 and upper bounds of 1. In the ideal case where A perfectly reflects X and B reflects Y, A*B perfectly captures X*Y. When the degree of formally guaranteed rights and democratic institutions becomes "very small":[1]

$$X \rightarrow 0 \Rightarrow A \rightarrow 0 \Rightarrow A*B \rightarrow 0$$

This ideal case is precisely Inglehart and Welzel's argument formalized. However, the ideal case is not a valid description of the actual measurement process. It was shown above that the FHI includes not only elements related to X, but also elements related to Y. Let us now assume that an arbitrary ¼ of the FHI-points come from "quality" considerations, which I believe is not an overestimation. This implies that $A = (\frac{3}{4}X + \frac{1}{4}Y)$. Take the hypothetical case of a regime lacking formal institutions such as free elections and formal political and civil rights, but where there exists rule of law and no corruption, no political violence, etc.; that is, $X \rightarrow 0$ and $Y \rightarrow 1$:

$$X \rightarrow 0 \ \& \ Y \rightarrow 1 \Rightarrow A \rightarrow \tfrac{1}{4} \ \& \ B \rightarrow 1 \Rightarrow A*B \rightarrow \tfrac{1}{4}$$

In this case, the EDI resembles an additive measure, since a regime lacking formally guaranteed rights and institutions, but curbing corruption and keeping firmly to its authoritarian laws, scores ¼ on the EDI, while it should have scored 0 ($X*Y \rightarrow 0$). The ¼ score on the EDI for the non-corrupt dictatorship is higher than the EDI score received by a totally corrupt formal democracy. For such a country, $X \rightarrow 1$ and $Y \rightarrow 0$, which implies that $A \rightarrow \frac{3}{4}$, $B \rightarrow 0$, and $A*B \rightarrow 0$.

More generally, if z is the share of "quality-based" points in FHI, and $z > 0$, a regime lacking all traits of "formal democracy" will be assigned a value on the EDI of $(z*Y)*Y = z*Y^2$. This is not equal to zero unless Y is zero. Therefore, Inglehart and Welzel's argument that the EDI avoids the

criticism of Hadenius and Teorell (2005) because of its multiplicative nature is invalid. It is theoretically possible that non-corrupt dictatorships can receive a higher rank on the EDI than corrupt "formal democracies," once one takes into account the real nature of the FHI.

III. Corruption as a Narrow Measure of Democratic Quality. Inglehart and Welzel are relatively clear on the point that corruption is a proxy for the wider phenomenon of lacking rule of law. Two questions can be raised. First, is corruption a valid and reliable measure for lacking rule of law? Second, are there factors, other than lacking rule of law, that can render formally guaranteed rights ineffective? The plausible answers to both of these questions (*no* and *yes*) point to the insufficiency of using corruption scores as a measure for "quality" or "effectiveness" of democracy. With regard to the first question, it is important to notice that the CPI used in Welzel et al. (2003), which is constructed on the basis of several sub-indices that measure perceptions of corruption in a country, is mainly based on perceptions of businesspeople and country specialists rather than ordinary citizens (Hadenius and Teorell, 2005: 90).[2] The CPI largely measures perceptions of economic corruption, and a country keeping its trade with large-scale international businesses relatively free from corruption could, therefore, gain a fair CPI score, even if it was corrupt in dealing with its citizens. The CCI used in Inglehart and Welzel (2005) tries to capture political corruption, and therefore represents a validity improvement. Nevertheless, several of the sub-indices in the CCI also rely on perceptions and experiences of firms in the marketplace (Kaufman et al., 2007: 75).

When it comes to the second question, there are several ways through which formally guaranteed rules and rights can be rendered inefficient. Political leaders who desire continued power have creatively used a variety of means to restrict popular participation in the political process. These means can be relatively subtle, such as implicit threats of sanctions (loss of position or benefits), and make for example formal freedom of speech useless. In Singapore, outspoken academics have lost their faculty positions (Sikorski, 1996). Violence and terror are other ways of restricting the actual exercise of formal rights. Political violence related to elections has been very common in Africa. Lindberg (2006: 48) compiled data on 232 elections in Africa, and only 22 percent of these elections were coded as elections without any form of political violence. Corruption is thus only one way to render formal rights for the populace ineffective. Diamond and Morlino recognize at least "eight dimensions on which democracies vary in quality" (2005: xii). These are: rule of law, participation, competition, vertical accountability, horizontal accountability, respect for political and civil freedoms, political equality, and responsiveness. The question then becomes why one should weigh corruption so heavily in the EDI.

If corruption functions as an unbiased estimator for "quality," this would be a problem only because of increased uncertainty in the estimates of the EDI.[3] However, corruption is systematically linked to other variables. First of all, corruption is for several reasons highly correlated with income level. The Weberian insight that poorly paid officials might expand their earnings by illegal means is one reason (Rauch and Evans, 2000). Corruption in the economic sphere also creates badly functioning economies. Corruption in the economic sphere is therefore both a cause and an effect of a low-income economy (Mauro, 1995). Giving perceived corruption a large influence on a democratic quality measure is therefore bound to secure the measure a high correlation with income and, therefore, also other variables correlated with income such as education or "emancipative value-indices." The correlation coefficient between income measured as GDP per capita in 2000 (PPP-adjusted) and the CPI in 1999 was .87. Causal analysis relating effective democracy to income, *or variables that are highly correlated with income*, could therefore tend to produce stronger results than are actually warranted, especially if there are other threats to democratic

quality than corruption in richer countries. The 1999 EDI based on the CPI had a correlation of .87 with GDP per capita in 2000, which is far higher than the .62 correlation between FHI in 1999 and GDP per capita.[4]

IV. Subjective Scoring: Validity and Reliability Issues. Suppose for a minute that FHI, A, measures the real degree of formal democracy, X, plus a random measurement error, ε. CPI, B, measures "functioning of institutions," Y, and an error term, γ. This gives:

$$A*B = (X+\varepsilon)*(Y+\gamma) = X*Y + X*\gamma + Y*\varepsilon + \varepsilon*\gamma$$

Finding estimates for the variance of a product of variables is not a straightforward task (Goodman, 1962), even if the variables are independently distributed and measures are unbiased. The multiplicative formulation of the EDI has the implication that the size of the total variance, and also the error variance, is relatively large when compared with any of its components.[5] Moreover, the assumptions of independence and unbiasedness probably do not hold when it comes to FHI and CPI. The indices are strongly positively correlated. There are also reasons to believe that specific *biases* in the indices are positively correlated, and alas the error terms are correlated as well. One obvious reason is that the CPI is an integrated part of an FHI check question. These two indices are also constructed as relatively subjective measures. As Bollen (1993) points out, judges can be biased in their scoring of such indices, and they might be systematically influenced by certain other variables. In this case the indices' error terms are not independent, which could lead to further problems.

More specifically, economic variables, like GDP level and growth, might influence judgments systematically. Might Singapore and Botswana, for example, be judged as more democratic and less corrupt than they actually are, because of their shining economic performances? This particular bias is referred to as a "poor is bad" bias, and it has received attention in the literature on corruption indices (Søreide, 2006: 7). If the error terms of the CPI and FHI are systematically correlated, this will exaggerate the size of the error term given by ε*γ. The new error term will probably not be normally distributed, and the reliability of the measure will suffer. Also, in the particular case of the "poor is bad" bias, poorer countries will be punished excessively on the EDI, since the two systematically biased error terms are multiplied.

When it comes to the CPI, Transparency International argues that the strong correlation between the CPI's different sub-indices is evidence that its validity is good (Lambsdorff, 2001). However, this reasoning requires that corruption perceptions in the different sub-indices are formed independently, which is questionable (Thompson and Shah, 2005: 14–18). A correlated bias could stem from judges reading other judges' reports or because of some common prior variable affecting judges systematically. There is also a validity problem related to the rating of corruption and freedom of the press in a country. If judges' *perceptions* of corruption are influenced by what they are able to access through the media, democracies might be rated systematically as more corrupt than is actually warranted when compared to authoritarian regimes. A free and functioning media might contribute to highlighting the need to combat corruption in a country, and the revelation of large scandals might actually drive CPI numbers higher than if these scandals were not discovered. There is solid empirical evidence for an effect on corruption perceptions as a result of scandals revealed by a free press (Thompson and Shah, 2005: 20–1). Since freedom of the media is an integral aspect of a well-functioning democracy, Belgium and Italy might suffer on indices such as the CPI relative to Singapore and China *because of their democratic qualities*. This effect is detrimental to the validity of the CPI as a democratic quality component. Couched in terms of the

equation above, the expected value of X*γ is not equal to zero. Even if E(ε) and E(γ) are both equal to zero, E(A*B) would not be E(X*Y).

V. Ordinal Measures and Multiplication. A variable has an ordinal measurement level if the numbers assigned to its values only contain information about the ranking of values, and no information about the distances between values. A cardinal measure, on the other hand, contains information about the absolute distance between values. In the ordinal case, if two units, P and Q, are assigned respectively 1 and 2 on a variable X, this only implies that Q has "more" of X than P has. If X has a cardinal measurement level, Q has twice as much of X as P. *The FHI and CPI are only variables at an ordinal measurement level.* If a country has half the corruption score or democracy score of another, this does not imply that the country is half as corrupt or half as democratic as the other in a strict metaphysical sense. There is no obvious yardstick that can decide what twice as democratic or corrupt actually means. One mathematical trait of cardinal measures is that they can be multiplied, and the multiplied measure still retains a meaningful interpretation. The product of two *ordinal* measures cannot be given a meaningful interpretation in the strictest sense, since ordinal scores only contain information about rank on the respective dimensions. This point has some serious de facto implications for the validity of the EDI, especially for the issue of interpretation of relative distance between units on the new variable, but also on the arbitrariness of rank between units. The stringent assumptions underlying the possibility of meaningful multiplication indicate that more subjective evaluative methods have to be applied in order to judge whether the EDI validly captures reality, or if it is systematically misrepresenting the actual degree of democracy.

3.3 Empirical Considerations on the Effective Democracy Index's Validity and Reliability

The above discussion should have made relatively clear that there are no God-given ways of measuring "substantive" or "effective" democracy. There are contingencies related to the choice of scoring principles, scaling, and included indicators. This should, however, not give rise to unconstrained intuitionism or discredit any attempt to classify. Democratic theory and empirical generalizations on the historical functioning of certain institutions can restrict the set of possible classificatory principles for effective democracy. I will draw on an analogy from John Rawls' (1999: 18–19) treatment of ethical principles versus ethical intuitions from single cases in ethics, when elaborating upon the role of principles in the measurement of democracy and their relation to the evaluation of specific cases. Rawls suggests that the mutual revisions of case interpretation and general principles should go on until we reach a "reflective equilibrium," where principles and case interpretation are in harmony. We saw above that the EDI could be questioned strongly on validity grounds. Therefore, according to Rawlsian logic, the more case-based interpretation of the degree of democracy becomes an important factor in settling whether the measure actually works as a good yardstick for democracy, or whether we should revise it. As we will see, the *rank* of some cases and the *distances* between cases, when they are interpreted linearly, point to deficiencies in the EDI.

Concerns about rank between cases were voiced by Hadenius and Teorell (2005), and I will therefore focus mainly on the distance issue. For example, regression analysis requires cardinal interpretation, and both the rank of and distance between units will influence regression results. I will look at empirical characteristics of the EDI systematically. Table 2 shows EDI scores for some selected

Table 2. EDI and FHI (Reversed and Normalized) in Selected Countries (1999)

Country	EDI	FHI
Denmark	100	100
Norway	89	100
Germany	73	92
Spain	61	92
Belgium	49	92
Mauritius	45	92
Italy	43	92
South Korea	32	83
Latvia	31	92
Singapore	30	33
Argentina	23	75
Jordan	22	50
Ghana	22	67
India	22	75
Ecuador	18	75
Malaysia	17	33
Paraguay	12	58
Zimbabwe	10	25
Russia	10	42
Tanzania	10	50
Nigeria	9	58
Belarus	6	17
China	3	8

countries in 1999 when CPI is used as corruption indicator, and Figure 1 is a scatter-plot with FHI in 1999 along the X-axis and EDI based on CPI in 1999 along the Y-axis.[6]

I. Rank. Singapore was put forward by Hadenius and Teorell as an example of a regime that receives too high a democracy score on the EDI. The degree to which the elite in the People's Action Party (PAP) has controlled political decision and policy-making, as well as curbed civil liberties, makes the regime deserve the badge "relatively authoritarian," or at least "semi-authoritarian." Singapore holds elections, but they are arranged in such a way that the opposition has typically won one or two seats in the 80-plus seat parliament. This result is also due to other aspects of political life such as the lack of civil liberties and the arrest of opposition leaders (Knutsen, 2006: 404–8). However, the fact that there exist elections and property rights, and there is little corruption, secures Singapore approximately 30 percent of the maximal FHI score. This score is almost duplicated in the EDI because of the success of the PAP in stamping out economic corruption, which is the type of corruption the CPI largely measures the perception of. However, as the discussion on ordinal measures showed, having 30 percent of the maximal EDI is not necessarily the same as having 30 percent of the maximal FHI.

Singapore is ranked below India and Argentina on the FHI, but above these two countries on the EDI. Even if democratic rights are rendered less effective by clientelism in India (Wilkinson, 2007) and corruption in Argentina (Freedom House, 2005), it seems difficult to claim, based on a case-by-case understanding of the politics in these countries, that the actual exercise of political

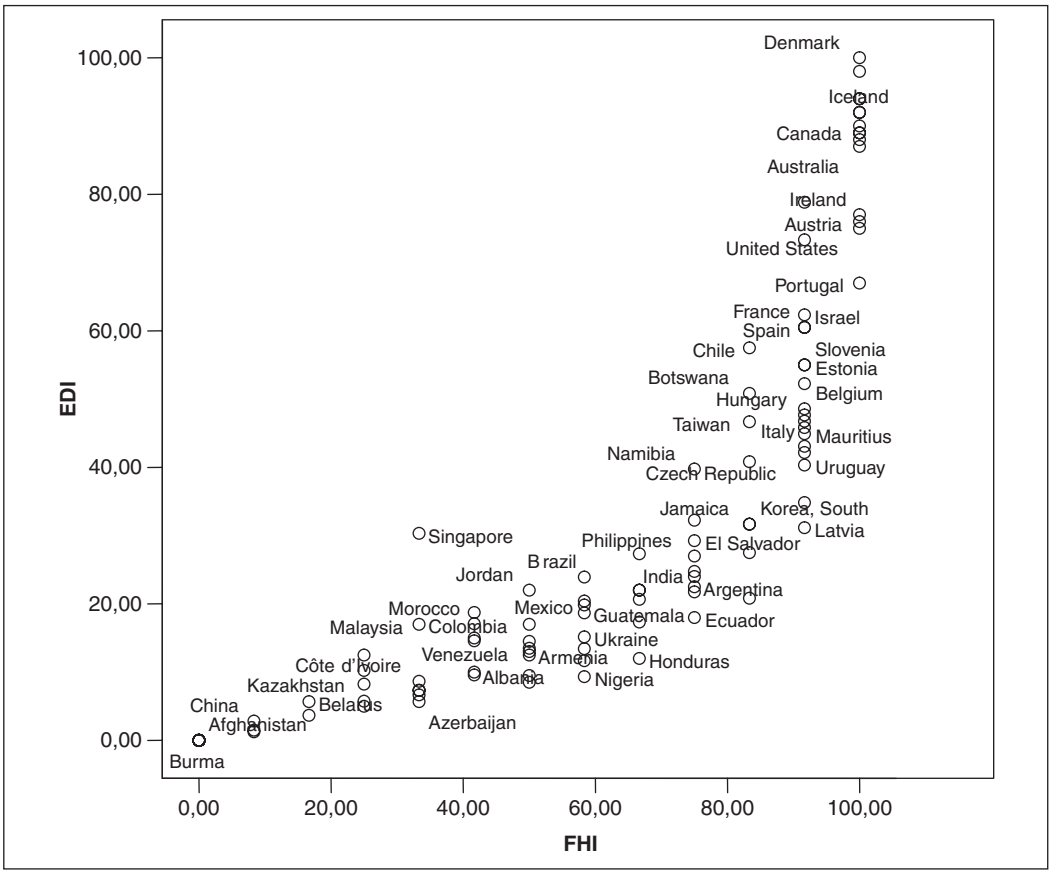

Figure 1. FHI (Reversed and Normalized) and EDI (1999)

and civil rights is greater in Singapore than in Argentina and India. Countries such as India and Argentina are punished "twice" because they are first penalized for dysfunctional democratic practices in the FHI, and thereafter punished by multiplying this score with a low CPI, which is intended to capture dysfunctional democratic practices. Jordan is another country that surpasses India on the EDI, even if it has a lower FHI score. Indeed, the number of rank reversals in 1999 is not negligible. The evidence can be read from Figure 1: if country F is to the right of country G on the X-axis but below F on the Y-axis, the two countries have different ranks on the FHI and EDI. The rank-reversal problem could have been even more problematic, were it not for the fact that there is a positive correlation between formal democratic institutions and lack of corruption (Fjelde and Hegre, 2008). However, young democracies (e.g. Benin, Ghana, and Ecuador) often have high levels of corruption (Rock, 2009). The rank-reversal problem is therefore generally found when comparing these countries with more authoritarian counterparts that have low (Singapore) or modest levels of corruption (e.g. Malaysia, Tunisia, and Jordan).

II. Distance. The linear interpretation of distances shows even more clearly the validity problems of the EDI. Even if the EDI is not metaphysically cardinal, one has to give it a cardinal

interpretation in a linear regression. I have no problem when comparing Singapore's EDI score with Denmark's, indicating that the degree of popular control over politics in the former is approximately 30 percent of the latter (whatever this actually means). However, some of the distances when interpreted in the light of other distances are counter-intuitive and fall short of empirical evidence. Mauritius is a country where the populace has had a relatively large influence over politics, with well-functioning elections, changes of government appearing to proceed smoothly, popular working-class movements coming to power, and freedom of speech and assembly (Knutsen, 2006: 438–40). There has also existed general political responsiveness to popular demands in areas such as health care and education (Bräutigam, 1997). Even if there are claims of corruption in Mauritian society, it is an exaggeration to claim that Mauritius is closer in democracy score to Singapore than to Spain, and it is even more disturbing that the distance upward to Denmark on the EDI is almost quadruple the distance downward to Singapore. Mauritius is also closer to China and Belarus than to Denmark.

There are several other problematic examples. The difference in EDI score in 1999 was larger between Denmark and Norway than between Yeltsin's Russia and China, and the difference between Denmark and Germany was larger than the difference between Latvia and Zimbabwe. Surely, corruption in Belgium and Italy makes popular control over politics less efficient, but it is nonsense to place these countries closer to North Korea, Myanmar, and Cuba than to some Nordic countries. Either there is a serious validity problem with the EDI or the reliability is weak, causing the score of Belgium, for example, to be driven by relatively arbitrary events.

Indeed, the Belgian case indicates weak reliability of the measure. The EDI score for Belgium in 2005 was 74. According to the measure, Belgium had experienced an "effective democratization" of 25 points from 1999 to 2005, which is difficult to attribute to anything but measurement error. This is a larger estimated democratization than we could have expected in Benin, if we were to have corruption scores before and after the end of the Kèrèkou regime in 1989–90.[7] This case illustrates a validity problem with the EDI: Benin shifted from an extremely oppressive military dictatorship, illustrated by consistent scores of 6.5 and 7 on the FHI, to a relatively well-functioning democracy with multi-party politics and freedom of speech and assembly (Gisselquist, 2007: 795; Magnusson, 2005: 81). Even though Benin falls short of being a "liberal democracy," it is a "consolidated minimal democracy" in the Schumpeterian sense, according to Gisselquist (2007). Although there is still widespread political clientelism, Benin has "experienced three successful alternations of executive power" (Gisselquist, 2007: 789), and this should be recognized as a considerable democratization when comparing the current regime with the earlier military dictatorship. However, since Benin is assigned CPI scores below 3, its regime change equals the effective democratization that Belgium is supposed to have experienced in recent years. The mathematical logic behind this result can be illustrated by the EDI's partial derivatives:

$$EDI = FHI*CPI \rightarrow \partial EDI/\partial FHI = CPI*\partial FHI \text{ and } \partial EDI/\partial CPI = FHI*\partial CPI$$

These differentiations show that large increases in the FHI have small effects on change in the EDI if the CPI level is low. A modest change in the CPI can have modest effects on the EDI if the FHI level is high. These considerations show a fundamental flaw with the EDI: when a dictatorship in a corrupt country is replaced by electoral democracy, the change in the EDI is miniscule. Corrupt countries are, as mentioned, most often poor countries in the third world and democratization in these countries is, therefore, not captured properly by the index.

Table 3. Bivariate Correlations between FHI, CPI, and EDI (1999)

	FHI	CPI	EDI
FHI	1	.67	.84
CPI	.67	1	.95
EDI	.84	.95	1

It is the multiplicative formulation of the EDI that drives the result above. Inglehart and Welzel claim that corruption scores only function as a grading factor and, therefore, cannot compensate for a low FHI score (2005: 194). However, this argument does not explain away the empirical fact that the EDI has a higher correlation with the CPI than the FHI has, as Table 3 shows. With a correlation coefficient of .95, we could actually use the CPI as a decent proxy for effective democracy.

One general problem of the EDI is that it compresses many observations on relatively low values. Corrupt democracies are far closer to harshly authoritarian regimes than well-functioning democracies. Modestly corrupt and semi-democratic regimes receive scores that are undeservedly low, especially if they go together. There is indeed evidence that so-called semi-democracies are more conducive to corruption than dictatorships for some operationalizations of democracy (Rock, 2009). The reason for the compression of observations on low values on the EDI is the multiplicative operation on the two ordinal indices CPI and FHI: 0.1 squared is 0.01, which is close to 0.3 squared (0.09), but 0.8 squared (0.64) is relatively far away from 1 squared. Empirically, only 25 countries score above 50 on the index ranging from 0 to 100 in 1999. The EDI has the attractive property of differentiating well between the relatively "effective democracies," but it fails to do so among other countries. The mean score is 31.2 and the median is 21.3. Half the countries are pressed together on one-fifth of the interval. This is problematic when using the index in a linear regression, since results would tend to represent differences between the "West" and "the rest," with the few countries dispersed over the relatively large intervals at the upper end of the scale driving regression coefficients. This can be illustrated by looking at regression coefficients for the EDI when GDP per capita is the dependent variable.[8] If we leave out the 15 countries with the lowest EDI scores in 2000 from the original 166-country sample, the regression coefficient increases by 5 percent from 293 (full sample) to 307. However, if we leave out the 15 countries with the highest EDI scores, all "Western" countries, the coefficient is reduced by 21 percent from 293 to 231. Western countries with high EDI scores are thus influential observations in this analysis.

III. Empirical Estimates of Biases. We can elaborate further on the systematic biases of the EDI by performing a factor analysis on several democracy indices and investigate how the EDI systematically deviates from the constructed factor. I use the Polity index, Vanhanen's Index of Democratization, the FHI, and the EDI in the factor analysis.[9] I use the CCI instead of the CPI as a corruption indicator. Factor scores are generated for those years all indices have data: 1996, 1998, and 2000. A principal component analysis finds only one factor with an eigenvalue above 1, confirming the presupposition that the indices are tapping one underlying concept: democracy.[10] I then run a factor analysis restricted to only one factor. As Table 4 shows, the factor loading is highest for the FHI. The FHI also had high loadings in factor analyses performed by Bollen (1993) and Bollen and Paxton (2000) in which different democracy indicators were utilized. Together, these results indicate that the FHI might be a decent measure of democracy.

Table 4. Factor Analysis on Four Democracy Indicators

Indicator	Factor loading	Uniqueness
Reversed FHI	.97	.06
Polity	.91	.16
EDI	.87	.24
Vanhanen	.86	.25

The share of unique variance is much higher for the EDI than for the FHI. Is the EDI's unique variance systematically related to specific cultural and economic factors? If so, such relations can be interpreted as systematic biases in the EDI. In order to investigate the issue systematically, I first run OLS with Panel Corrected Standard Errors (PCSE), with the factor retained from the factor analysis as independent variable and the EDI as dependent.[11] The residuals from this regression are the components of the EDI scores not explained by the democracy factor. As discussed above, one potential bias in the EDI might be that poor countries are punished excessively and rich countries are given higher scores than they deserve. The positive bivariate relationship between the EDI residual and GDP per capita, shown in Figure 2, supports this hypothesis. The predicted residual increases strongly with income.

We can investigate the potential biases of the EDI more systematically by running OLS with PCSE with the EDI residual as dependent variable. GDP per capita, a dummy for "Western"

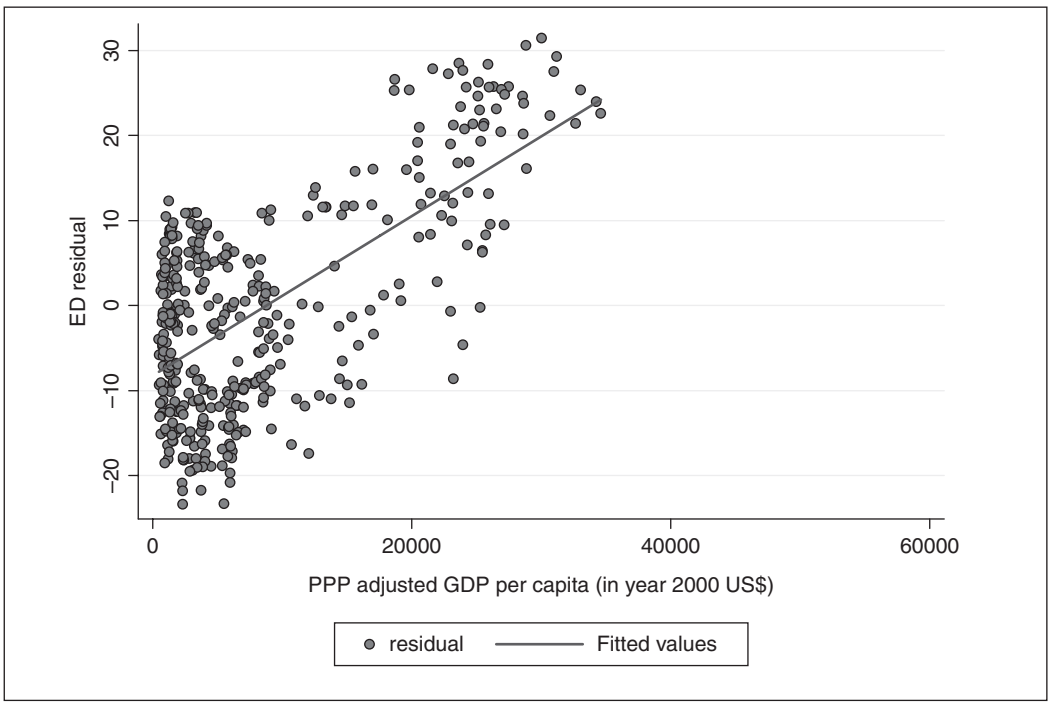

Figure 2. GDP per Capita and EDI Residuals; Global Sample (1996, 1998, and 2000)

countries (USA, Canada, Europe west of the old iron curtain, New Zealand, and Australia), and a dummy for countries with Protestant or Anglican Christianity as plurality religion enter as independent variables. The analysis includes 402 country-years from 141 countries. All three coefficients are positive and significant at the 1 percent level. These three variables, therefore, go together with significant biases in the EDI. The estimated effect on the residual of being a Western country is 9.0 and the estimated effect of Protestantism is 3.5. Moreover, an increase in GDP per capita of US$1000 (PPP-adjusted, measured in year 2000 US dollars) increases the residual by 0.5. Therefore, the predicted difference in EDI residuals between a non-Protestant, non-Western country with a GDP per capita of US$2000 and a Protestant, Western country with a GDP per capita of US$32,000 is 27.5, which is more than a quarter of the EDI's range. If these estimates are valid, the biases in the EDI are both systematic and substantial, and the biases will most likely affect analyses using the EDI as a dependent or independent variable.

The results above have been treated as indications of biases in the EDI. The measure used as a background for establishing these biases was a democracy factor constructed by running a factor analysis on four different indices. Since the EDI is one of those indices, the factor will have some resemblance to the EDI by construction. Therefore, the biases might be even stronger than these results indicate. The obvious counter-argument put forward by proponents of the EDI is that the other indices of democracy are skewed, since they do not sufficiently incorporate democratic-quality components. If these components are overrepresented in richer, Protestant, Western countries, the claimed biases could to a certain degree reflect the high degree of democratic quality. However, the arguments presented above indicated serious biases in the EDI, so that this section can be viewed as a statistical exercise to establish crude estimates of the sizes of these biases.

4. Conclusion

When estimating a political regime's democratic characteristics empirically, the need to go beyond the counting and classification of formal institutions and constitutional articles seems obvious to proponents of substantive-democracy definitions. The measure of effective democracy proposed by Inglehart and Welzel takes the issue of the actual functioning of political institutions seriously. However, there are some grave flaws with the measure, indicating that the search for a measure of substantive democracy has not ended. The relatively unsuccessful attempt by Inglehart and Welzel to operationalize a substantive-democracy concept reflects a more general problem identified by Goertz (2006: 2), who claims there is a chasm between "the qualitative scholar's concern for substantively valid concepts and the quantitative scholar's interest in good numerical measures." There seems to be a trade-off between the reliable measurement of narrower, institutionally based definitions of democracy, on the one hand, and the problematic measurement of broader, substantive-democracy definitions on the other. The FHI might represent a decent middle ground, but there are problems with the FHI as well (see, for example, Munck and Verkuilen, 2002). Can we do better than using the FHI in quantitative studies of democracy's causes and effects?

One very interesting alternative is the so-called democracy audit (Beetham and Democratic Audit UK, 1997), which is argued to be appropriate in contexts where there are grave deficiencies in the functioning of formal democratic institutions, for example in Africa (Baker, 1999). Instead of relying on the a priori multiplicative assumption of formal democracy interacting with corruption, this approach is more nuanced, asking a large amount of check questions explicitly related to the actual functioning of rights, liberties, and institutions. The drawback of such audits is the amount of work needed to construct scores and the difficulty in securing comparability across

countries. In practice, therefore, only a few countries have had democracy audits, and no straightforward comparable scores have been generated from such work.[12]

If one wants to stick to parsimonious frameworks like Inglehart and Welzel's, one could improve upon the EDI by making some adjustments. If one believes that the FHI incorporates quality components but nevertheless underestimates the importance of these components for substantive democracy, one could give the FHI a stronger weighting in the scoring procedure than do Inglehart and Welzel (2005), for example by using a Cobb-Douglass formula such as $EDI = FHI^{\alpha} * CCI^{1-\alpha}$, where $0.5<\alpha<1$. This index would also be homogeneous of degree one rather than of degree two, like the original EDI, meaning that a doubling of both the FHI and CCI would lead to a doubling of the new measure but a quadrupling of the original EDI. This would mitigate the problem of having only a few units with high scores, which troubled the original EDI. This would of course move the index away from a clean and intuitively appealing formula, although a multiplicative formulation will in any case give arbitrary results, since we are not dealing with cardinal measures. This suggested formula would give rankings and distances that make more intuitive sense than the originally proposed EDI: India would, for example, be somewhat further away from China, and Italy and Spain would be closer to other OECD countries. The new measure might not represent a Rawlsian reflective equilibrium, but we would probably be closer to one after the revision of the EDI.

Another improvement would come from recognizing that it is not only corruption that affects the quality and actual practice of democratic rights and institutions. Instead of letting CPI or CCI be the single quality component, one could introduce a weighted index, where corruption could be accompanied by other variables like "responsiveness of elites," "political power distribution," and other "rule of law" indicators. Even if less parsimonious, such a measure would clearly exhibit superior validity in capturing the actual functioning of democracy in a country. If researchers want to stick with a multiplicative formula, a suggestion could then be to construct an EDI with n quality components according to the formula: $EDI = FHI^{\alpha} * (b_1/n + ... + b_n/n)^{1-\alpha}$. Here, b_i represents quality component i, where $i \in \{1,2,...,n\} i$.

As Munck and Verkuilen (2002) point out, there are deficiencies with all available democracy indicators. However, the minimalist measures often have logically well-founded operational rules, whereas the more maximalist measures, such as the FHI, rely on problematic aggregation rules. The EDI is beset with even more severe methodological problems than the indices surveyed in Munck and Verkuilen. The adjusted EDI proposed in this conclusion would probably mitigate some of these problems, but it would still face problems stemming from the multiplication of ordinal measures. How should we analyze the causes and effects of democracy defined substantively when minimalist measures systematically leave out important aspects of democracy, and maximalist measures have problems related to the aggregation of aspects and the reliability of measurement? First, we should be aware of the specific methodological problems related to different operationalizations, and how these problems might bias results. Second, since there are no perfect measures of democracy, we should use different indicators in our studies to ensure that our results are not driven by measurement error or other indicator-specific attributes. Another solution to the problem of indicator-specific attributes could be to construct democracy factors from factor analysis on the most appropriate measures of democracy.

Acknowledgements

Thanks to Christian Welzel, Kalle Moene, Bjørn Høyland, Bjørn Erik Rasch, Oddbjørn Kntusen, Helge Hveem, Håvard Hegre, Kristian Helland-Hansen and Asmund Rygh for valuable comments and suggestions.

Notes

1. Note that → means that the measure goes toward a value, and ➔ means "implies."
2. Transparency International's arguments for not using citizens' corruption perceptions in the CPI are that citizens lack comparative experiences from other countries and that citizens in different countries might have different moral standards (Lambsdorff, 2001: 6–7). However, the CPI is intended to capture deprivation of *citizens*' formal rights in the EDI. Transparency International's Global Corruption Barometer, which evaluates "how and where ordinary people feel corruption's impact" (Transparency International, 2007: 2), for example in judiciaries, legislatures, and political parties, might therefore be more valid as an EDI component.
3. Even unsystematic errors in independent variables can generate biases in a regression framework. In a bivariate model, such bias is known as the "attenuation bias," and it draws coefficients toward zero. The bias is more complex in multivariate models (Greene, 2003: 83–6).
4. If we use the CCI instead of the CPI, the results are approximately similar. The EDI and CCI also have higher correlations with other structural variables such as school enrollment ratios and urbanization than the FHI. For example, in 2000 the correlation between the tertiary school enrollment ratio and the EDI was .70, whereas the correlation with the FHI was .57. All samples used in the article are global.
5. See Goodman (1962) for the specific expressions and approximations for variances of products.
6. The countries in Table 2 are not representative of the global sample, but have been intentionally selected to illustrate rank reversals and problematic distances between countries on the EDI. Most of the countries are referred to explicitly in the text.
7. The first CPI score for Benin was registered in 2004, when the country scored 3.2. From 2005 to 2007 the country scored between 2.5 and 2.9.
8. The CCI is used as a corruption measure. The GDP data are from the *World Development Indicators*.
9. For a description of the indices, see Marshall and Jaggers (2002), Vanhanen (2000), and Freedom House (2005). In short, the 21-point Polity index scores countries additively on the basis of five dimensions: openness and competitiveness of executive recruitment, regulation and competitiveness of political participation, and constraints on the chief executive. The dimension "constraints on the chief executive" is weighed most heavily in the scoring. The Vanhanen index is a multiplicative index based on two dimensions, contestation and participation. The first dimension is operationalized by subtracting the percentage of votes going to the largest party from 100. The second dimension measures the percentage of the population that votes in elections.
10. The first factor had an eigenvalue of 3.45 and the second of .28 (n=437).
11. All OLS with PCSE analyses in this article adjust only for autocorrelation within panels.
12. See, for example, Beetham et al. (2002).

References

Arrow, Kenneth (1951). *Social Choice and Individual Values*. New York: Wiley.

Baker, Bruce (1999). "The Quality of Democracy in the Developing World: Why and How It Should Be Measured." Paper for 27th ECPR Joint Sessions of Workshops, Mannheim, March 26–31 1999.

Beetham, David (1999). *Democracy and Human Rights*. Cambridge: Polity Press.

Beetham, David and Democratic Audit UK (1997). "Conducting a Democratic Audit." Manuscript, University of Leeds, UK.

Beetham, David, Bracking, Sarah, Kearton, Iain, Vittal, Nalani, and Weir, Stuart, eds (2002). *The State of Democracy: Democracy Assessments in Eight Nations around the World*. The Hague: Kluwer Law International.

Bollen, Kenneth (1993). "Liberal Democracy: Validity and Method Factors in Cross-National Measures," *American Journal of Political Science* 37(4): 1207–30.

Bollen, Kenneth and Paxton, Pamela (2000). "Subjective Measures of Liberal Democracy," *Comparative Political Studies* 33(1): 58–86.

Bräutigam, Deborah (1997). "Institutions, Economic Reform and Democratic Consolidation in Mauritius," *Comparative Politics* 30(1): 45–62.

Chabal, Patrick and Daloz, Jean-Pascal (1999). *Africa Works: Disorder as Political Instrument.* Bloomington: Indiana University Press.

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

Diamond, Larry (1999). *Developing Democracy: Toward Consolidation.* Baltimore, MD: Johns Hopkins University Press.

Diamond, Larry and Morlino, Leonardo (2005). "Introduction," in Larry Diamond and Leonardo Morlino (eds), *Assessing the Quality of Democracy.* Baltimore, MD: Johns Hopkins University Press.

Fjelde, Hanne and Hegre, Håvard (2008). "Democracy Depraved: Political Corruption and Institutional Transitions." Paper for the Norwegian National Political Science Conference, Tromsø, April 28–30.

Freedom House (2005). *Freedom in the World: The Annual Survey of Political Rights and Civil Liberties*. Lanham, MD: Rowman & Littlefield.

Gisselquist, Rachel M. (2007). "Democratic Transition and Democratic Survival in Benin," *Democratization* 14(4): 789–814.

Goertz, Gary (2006). *Social Science Concepts: A User's Guide*. Princeton, NJ: Princeton University Press.

Goodman, Leo (1962). "The Variance of the Product of K Random Variables," *Journal of the American Statistical Association* 57: 54–60.

Greene, William H. (2003). *Econometric Analysis*, 5th edn. Upper Saddle River, NJ: Prentice Hall.

Hadenius, Axel and Teorell, Jan (2005). "Cultural and Economic Prerequisites of Democracy: Reassessing Recent Evidence," *Studies in Comparative International Development* 39(4): 87–106.

Inglehart, Ronald and Welzel, Christian (2005). *Modernization, Cultural Change and Democracy: The Human Development Sequence*. Cambridge: Cambridge University Press.

Johnson, Chalmers (1995). *Japan, Who Governs? The Rise of the Developmental State.* New York: W.W. Norton.

Kaufmann, Daniel, Kraay, Aart, and Mastruzzi, Massimo (2007). "Governance Matters VI: Aggregate and Individual Governance Indicators 1996–2006." World Bank Paper, Washington, DC.

Knutsen, Carl Henrik (2006). "Political Regime Types and Economic Growth: Are Democracies Better at Increasing Prosperity?" Master's thesis, Department of Political Science, University of Oslo.

Lambsdorff, Johann Graf (2001). "Background Paper to the 2001 Corruption Perceptions Index: Framework Document." Paper for Transparency International and Göttingen University.

Levinson, Sanford (2007). "How the United States Constitution Contributes to the Democratic Deficit in America," *Drake Law Review* 55 (2006–7): 859–78.

Lindberg, Staffan I. (2006). *Democracy and Elections in Africa.* Baltimore, MD: Johns Hopkins University Press.

Magnusson, Bruce A. (2005). "Democratic Legitimacy in Benin: Institutions and Identity in a Regional Context," in Leonardo A. Villalón and Peter VonDoepp (eds), *The Fate of Africa's Democratic Experiments.* Bloomington: Indiana University Press.

Marshall, M.G. and Jaggers, Keith (2002). "Polity IV Project: Dataset Users' Manual." Manual for Program Center for International Development and Conflict Management, College Park, University of Maryland.

Mauro, Paolo (1995). "Corruption and Growth," *Quarterly Journal of Economics* 110(3): 681–712.

Médard, Jean-François (1996). "Patrimonialism, Neo-patrimonialism and the Study of the Post-colonial State in Subsaharan Africa," in Henrik Secher Marcussen (ed.), *Improved Natural Resource Management: The Role of Formal Organizations and Informal Networks and Institutions*. Roskilde, Denmark: Roskilde University.

Munck, Gerardo L. and Verkuilen, Jay (2002). "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices," *Comparative Political Studies* 35(1): 5–34.

Olson, Macur (1982). *The Rise and Decline of Nations: Economic Growth, Stagflation and Social Rigidities*. New Haven, CT: Yale University Press.

Przeworski, Adam, Alvarez, Michael E., Cheibub, José Antonio, and Limongi, Fernando (2000). *Democracy and Development.* Cambridge: Cambridge University Press.

Rauch, James E. and Evans, Peter B. (2000). "Bureaucratic Structure and Bureaucratic Performance in Less Developed Countries," *Journal of Public Economics* 75(1): 49–71.

Rawls, John (1999). *A Theory of Justice*. Oxford: Oxford University Press.

Rock, Michael T. (2009). "Corruption and Democracy," *Journal of Development Studies* 45(1): 55–75.

Schumpeter, Joseph A. (1976). *Capitalism, Socialism and Democracy*. New York: Harper Perennial.

Sidel, John T. (1999). *Capital, Coercion and Crime: Bossism in the Philippines*. Stanford, CA: Stanford University Press.

Sikorski, Douglas (1996). "Effective Government in Singapore: Perspective of a Concerned American," *Asian Survey* 36(8): 818–32.

Søreide, Tina (2006). "Is It Wrong to Rank? A Critical Assessment of Corruption Indices." CMI Working Paper 2006:1, Christian Michelsen Institute, Bergen, Norway.

Thompson, Theresa and Shah, Anwar (2005). "Transparency International's Corruption Perception Index: Whose Perceptions Are They Anyway?" Discussion draft for the World Bank, Washington, DC, March 2005, URL: http://siteresources.worldbank.org/INTWBIGOVANTCOR/Resources/Transparency-InternationalCorruptionIndex.pdf.

Transparency International (2007). "Report on the Transparency International Global Corruption Barometer 2007." Policy and Research Department, Transparency International, Berlin, URL: http://www.ethics-world.org/publicsectorgovernance/PDF%20links/BAROMETER.pdf.

Vanhanen, Tatu (2000). "A New Dataset for Measuring Democracy, 1810–1998," *Journal of Peace Research* 37(2): 251–65.

Welzel, Christian (2007). "Are Levels of Democracy Affected by Mass Attitudes? Testing Attainment and Sustainment Effects on Democracy," *International Political Science Review* 28(4): 397–424.

Welzel, Christian and Inglehart, Ronald (2006). "Emancipative Values and Democracy: Response to Hadenius and Teorell," *Studies in Comparative International Development* 41(3): 74–94.

Welzel, Christian, Inglehart, Ronald, and Klingemann, Hans-Dieter (2003). "The Theory of Human Development: A Cross-cultural Analysis," *European Journal of Political Research* 42(2): 341–79.

Wilkinson, Steven I. (2007). "Explaining Changing Patterns of Party–Voter Linkages in India," in Herbert Kitschelt and Steven I. Wilkinson (eds), *Patrons, Clients and Policies: Patterns of Democratic Accountability and Political Competition.* Cambridge: Cambridge University Press.

World Bank (2008). *World Development Indicators 2008*. Washington, DC: World Bank.

Biographical Note

Carl Henrik Knutsen is PhD Fellow at the Department of Political Science, University of Oslo, and Researcher at the Centre for the Study of Civil War, PRIO (International Peace Research Institute, Oslo). His work focuses on comparative political economy, and his forthcoming PhD thesis is entitled "The Economic Effects of Democracy and Dictatorship." Address: Department of Political Science, University of Oslo, Boks 1097 Blindern, 0317 Oslo, Norway [email: c.h.knutsen@stv.uio.no].