

# Measuring Liberal Democratic Performance: an Empirical and Conceptual Critique

Joe Foweraker and Roman Krznaric

*Department of Government, University of Essex*

Liberal democratic performance is understood as the delivery of liberal democratic values, and not as regime longevity or government efficacy. Measuring it is a matter of how far liberal democratic governments achieve in practice the values they endorse in principle.

It is recognized that the performance of liberal democratic governments varies widely. But extant attempts to measure this variation suffer problems of reliability and validity, and the object of measurement is often unclear.

By defining the range of liberal democratic values we demonstrate that performance is multi-dimensional and that trade-offs across different values can create distinct performance profiles. The narrow gauge of the extant measures – usually of just one or two values – is often disguised by single scales that masquerade as summary performance indicators.

---

## Comparative Liberal Democratic Performance

Liberal democratic government may be defined in a minimal and procedural fashion as a political system where multiple political parties compete for control of the government through relatively free and fair elections. But, beyond this minimum benchmark, it is recognized that the liberal democratic performance of such political systems varies widely.

This variation inspired Dahl's description of 'really existing' liberal democratic governments as 'polyarchies', at a time when there were just thirty-five or so such political systems, most of them in rich and industrialized nations of the Western Hemisphere (Dahl, 1971). Today this number has grown to some one hundred and twenty.<sup>1</sup> But, although claims to liberal democracy – including an emphasis on individual rights and the rule of law – now serve as an almost universal principle of political legitimacy across the globe,<sup>2</sup> the real variation in liberal democratic performance remains, and may even have increased.

Liberal democratic performance concerns the practices of liberal democratic governments. It does not have to do with competing claims to democratic governance, such as 'people's democracy' or associational democracy, still less with democracy writ large. It is accepted that some minimum level of democratic performance must be achieved for a system of government to be defined as a liberal democracy (a familiar problem of degree and kind), but it is the variation in the practices of governments that matters.

Liberal democratic performance is understood in different ways, and this tends to make comparisons more difficult. It is therefore helpful to distinguish the three

principal interpretations of this performance as, first, regime endurance or longevity; second, government efficacy; and third, the delivery of liberal democratic values, or how far liberal democratic governments achieve in practice the values to which they subscribe in principle.

Studies of regime endurance examine the 'survival rates' of different types of regime in different economic or cultural settings (Przeworski, Alvarez, Cheibub and Limongi, 1996, p. 39). Studies of government efficacy look at macroeconomic management, or social policy and welfare provision, in order to assess 'government performance' rather than democratic performance per se.<sup>3</sup> A focus on liberal democratic performance, in contrast, will exclude the values that may provide proper measures of the efficacy of *any* system of government (for example national security, social welfare, protection of the environment, even legitimacy and system support), in favour of values that are intrinsic to liberal democratic government. Liberal democratic performance then 'refers to the degree to which a system meets such democratic norms as representativeness, accountability, equality and participation' (Lijphart, 1993, p. 149).

There is now an impressive range of 'democratic indicators' that have come to constitute barometers of comparative democratic performance in academic, public policy and business circles. Indeed, the accumulated effort to measure such performance may be seen to comprise a distinct field within political science. But the measures remain problematic, and the problems are of two main types. On the one hand, difficulties with the empirical construction of the measures lead to doubts about their reliability and validity. On the other, a recurrent lack of conceptual clarity means that the object of measurement is often opaque.<sup>4</sup>

This essay responds to these problems, first, by providing an empirical critique of the capacity of the measures to gauge liberal democratic performance; and, second, by pursuing a conceptual critique that may clarify what precisely is being measured. To do so, the argument addresses two kinds of measures. On the one hand, it considers measures that are (more or less) explicitly concerned with democratic performance, such as those employed in Dahl's *Polyarchy*. On the other, it considers measures developed through the comparative study of institutions and rights that may (and, we argue, should) be integrated into the study of democratic performance.<sup>5</sup>

The empirical critique is mainly methodological, and self-explanatory. The conceptual critique is built on a prescriptive account of the core values of liberal democratic government, and founded axiomatically on the two key principles of liberty and equality as upheld through the rule of law and the sovereignty of the people.<sup>6</sup> It is argued that these principles are achieved in practice through the operation of eight values that combine the *individual* experience of democracy (rule of law) with the *institutional* efficacy of democratic government (sovereignty of the people). The first dimension contains the legal values of civil rights, property rights, political rights, and minority rights. These rights and the rule of law are important guarantees of individual freedoms and protections, and so help to deliver the substance of democracy to the citizenry at large. The second dimension contains the institutional values of accountability, representation, constraint and participation. These are the values that protect the rule of law by making government accountable to the people.

The primary purpose of defending this range of values is to demonstrate that liberal democratic performance is not uni-dimensional, but can be measured across the range, so creating 'performance profiles' and admitting the likelihood of trade-offs across different values.<sup>7</sup> It will be seen that many measures of liberal democratic performance tend to focus on just one or two of these values, and consequently can only provide a rather partial picture of this performance.<sup>8</sup> It will also be seen that this narrow gauge is frequently disguised by single scales that masquerade as summary indicators of the performance of liberal democratic governments. In addition, the range of values may improve the purchase of many existing measures by bringing them into a common compass. In this sense, the argument sets out to organize the field of performance measures, and so contribute to realize their potential. Any claim to originality cannot reside in its parts, but only in its overall composition.

### **The Empirical Critique**

The inquiry begins by reviewing the empirical scope and design of the extant measures of liberal democratic performance. In particular, it compares these measures according to five empirical categories, namely time frames, case selection and geographical focus, data sources, weightings, and measurement level. It is suggested that particular care must be taken with measures that are highly aggregated with unexplained weightings, or based on little information or on some forms of survey data, or constructed subjectively from the broad categories of ordinal scales.

At the same time it is important to assess the reliability and validity of the measures, and to distinguish carefully between the two. Reliability is simply 'the extent to which measurements are consistent when repeated by the same observer, or by different observers using the same instrument' (Barsh, 1993, p. 94). Validity, very differently, concerns the degree to which the measures reflect or 'capture' liberal democratic values (or, by extension, the institutional and legal conditions that maintain and reproduce those values).<sup>9</sup> Values such as representation or constraint cannot be observed directly, and their measurement can never be exact. But the empirical design of each measure must influence just how close it gets to the value being measured. This emphasis on reliability and validity reflects an overall concern with good measurement practice.

The first objective is to compare the empirical scope (time frames, case selection and geographical focus) and empirical design (data sources, weightings, measurement levels) of the forty-five existing studies that provide explicit or implicit measures of democratic performance. These studies are listed in Table 1. Previous critiques of the scope and design of performance measures have tended either to be more particular or more polemical. Thus, Moore (1995) focused exclusively on the democracy measures of Arat and Vanhanen, while Goldstein (1992) delivered a trenchant and highly skeptical account of certain measures of civil and political rights. This review sets out to be as comprehensive, and therefore as inclusive as possible.

#### ***Time Frames***

More than half of the forty-five studies simply provide cross-national measures at one point in time, and many of them do so in order to explore the causes of

**Table 1: Comparative Democratic Performance: Geographical Scope and Temporal Range**

Author	Time frame	No. cases	Geog. focus
Arat (1991)	Annual time-series (1948–1982)	152	Broad range
Banks (1997)	Annual time-series (1815/1950–1997)	115	Broad range
Bollen (1980)	Cross-national (1960, 1965)	c.120	Broad range
Bond <i>et al.</i> (1997)	Annual time-series (1984–1994)	4	LDCs
Borner <i>et al.</i> (1995)	Cross-national (1981–1985, 1986–1990, 1981–1990)	28	LDCs
Cingranelli and Richards (1998)	Time series (1981, 84, 87,90, 93, 96)	79	Broad Range
Cingranelli and Wright (1986)	Cross-national (1980)	152	Broad range
Coppedge and Reinicke (1991)	Cross-National (1985)	170	Broad range
Cutright (1963)	Cumulative (1940–1960)	70	Broad range but excludes Africa
Dahl (1971)	Cross-national (circa 1969)	114	Broad range
Deininger and Squire (1996)	Time series (c.1950–1990)	108	Broad Range (few Africa)
Elklit (1994)	Cross-national (1990–1993 average)	5	Bulgaria, Kenya, Latvia, Mongolia, Nepal
Ersson (1998)	Time series (1957–1997, election years)	18	Western Europe
Foweraker and Landman (1997)	Annual time-series (1958–1990)	4	Spain, Chile, Brazil, Mexico
Franklin (1996)	Cross-national (averages 1960–1995)	37	New and old democracies
Freedom House (1997a)	Annual time-series (1973–1997)	191	Broad range
Freedom House Economic Freedom (1996)	Cross-national (1995)	82	Broad range
Freedom House Press Freedom (1997b)	Cross-national (1997)	187	Broad range
Gasiorowski (1996)	Annual time-series (independence/ 1945–1992)	97	LDCs
Hadenius (1994)	Cumulative (varies)	31	LDCs
Hadenius (1992)	Cross-national (1988)	132	LDCs

Table 1: continued

Author	Time frame	No. cases	Geog. focus
Haxton and Gurr (1996)	Mixture of every five years/ two years/ year from 1945	112 (with 268 groups at risk)	Broad range
Humana (1992)	Time-series (1983, 1986, 1991)	104	Broad range
International IDEA (1997)	Time-series (1945–1997, election years)	163	Broad range (excl. one-party states)
Inter-Parliamentary Union (1995)	Time series (1945–1995, election years)	186	Broad range
Jackman (1973)	Cross-national (data from 1964, 1966, 1972)	60	Broad range (excludes Communist states)
Jagers and Gurr (1995)	Annual time-series (1800–1994)	157	Broad Range
Johnson <i>et al.</i> (1998)	Annual time-series (1995–1998)	154	Broad range
Knack and Keefer (1995)	Cross-national (1972 and 1982)	100+	Broad range
Lijphart (1994a)	Mixture (1945–1990)	27	Established democracies
Lijphart (1994b)	Cross-national (averages for different periods)	18	Established democracies
Lijphart (1984)	Cross-national (generally averages for 1945–1980)	21	Established democracies
Lipset (1959)	Cross-national (c.1950)	48	Europe, English- speaking democracies, Latin America
Mauro (1995)	Cross-national (averages of 1980–1983)	68	Broad range
Poe and Tate (1994)	Annual time-series (1980–1987)	153	Broad range
Political Risk Services (1998)	Annual time-series (1982–1995)	129	Broad range
Pougerami (1988)	Cross-national (averages of 1983–1985)	92	Broad range

**Table 1: continued**

Author	Time frame	No. cases	Geog. focus
Powell (1982)	Cross-national (averages of 1958–1976 and averages for each decade)	29	Established democracies
Purdue University (1998)	Annual time series (1980–1994)	147	Developing only
Taylor (1991)	Annual time-series (1948–1992)	9	Australia, India, Switz., Kenya, Albania, Syria, China, Algeria, Argentina
Transparency International (1997)	Annual time-series (1995–1997, plus averages of 1980–1985, 1988–1992)	52	Broad range
Van belle (1997)	Annual time-series (1948–1995)	c.163	Broad range
Vanhanen (1997)	Decennial and annual time-series (1850–1993)	172	Broad range
Walmsley (1996)	Annual time-series (1980–1995)	16	East/Central Europe

*Hereafter, unless specified, references in the text to the authors in Table 1 correspond to the work indicated in Table 1.*

democracy, or the relationship between development and democracy in particular. In the latter studies, democracy may act either as an independent or a dependent variable (for example Dahl, Jackman, Pougerami, Hadenius). But it is only time-series indicators that can reflect increases in the number of liberal democracies, or variations in democratic performance over time, usually by providing annual scores for the countries under review.<sup>10</sup> Freedom House (Comparative Survey of Freedom) provides such scores from 1973 to the present, Arat from 1948 to 1982, and Poe and Tate from 1980 to 1987; while Banks, and Jagers and Gurr, attempt an ambitious annual coverage from the 19th century onwards. However, some time-series measures do not provide annual scores, and this may make them less sensitive. Vanhanen rates countries on a decennial basis before 1980, Haxton and Gurr mix annual, biannual and other intervals, and Humana's World Human Rights Guide is only available for 1983, 1986, and 1992.

The institutional analysis of Lijphart and Powell is more difficult to classify, since it tends to present measures of party system and constitutional design as averages of annual scores over time. Typically, Lijphart's measure of Minimum Winning

Cabinets is derived from the proportion of time each of his twenty-one countries had minimum-winning cabinets over the years 1945–1980 (Lijphart, 1984, p. 212). Similarly, Powell's (1982, p. 19) measure of Stability is the average durability of governments or presidents for the period 1967–1976. Such institutional measures may vary little over time, or may not be meaningful at a single point in time, and so there is little or no advantage in providing annual scores.

### *Case Selection and Geographical Focus*

The majority of measures cover a large number of nations, often more than one hundred (see table), and aspire to general statements about democratic performance across the globe. These large-N studies are often criticized for using inaccurate or inadequate data.<sup>11</sup> It may be inaccurate insofar as it relies on expert opinion, and inadequate insofar as it is based on constitutional principle rather than actual political practice.<sup>12</sup> Similar criticisms may be leveled at some of the more geographically restricted studies. Hadenius, Gasiorowski, and Borner *et al.* focus on the developing world, while Lijphart (1984), in contrast, concentrates on the institutional design of twenty-one countries of the developed world. Other restrictions tend to operate in an *ad hoc* fashion. Cutright removes African countries from his analysis of National Political Development. Lipset compares the developed western democracies solely to Latin America. Jackman excludes communist states. Haxton and Gurr include only those countries with 'minorities at risk'.

Variable coverage limits the opportunities for comprehensive comparisons. Press freedom can be compared across most countries using Van Belle's indicators, but Lijphartian measures of representation are far more restricted. Hence, a variety of measures could be used to construct a complex 'performance profile' for the Netherlands in 1996, but it would be impossible to do the same for Belarus. Consequently, broad comparisons of democratic performance have perforce to focus on some liberal democratic values and not others. If these measures happen to vary little or not at all in the developed world, the comparisons may lead to the mistaken conclusion that democracy is pristine and unproblematic in this world, and only beset with problems elsewhere. Lipset tended to import this assumption into his comparison of Latin America and the developed West by setting a lower 'democratic threshold' for the Latin American cases.

### *Data Sources*

Data sources directly affect the scope and indirectly influence the design of performance measures. The main sources are remarkably few. Most of the 'events' measures, as well as some institutional measures, are taken from data-sets constructed by Banks and Textor (1963), Taylor and Hudson (1972), Taylor and Jodice (1983), and Banks. Studies that typically use these sources include Arat, Bollen, Dahl, Jackman, and Lijphart. Many of the electoral and institutional measures, including those of Arat, Coppedge and Reinicke, Gasiorowski, Hadenius, Lijphart, Powell, and Vanhanen, rely on a number of standard reference works such as Europa World Yearbook, Keesings Contemporary Archives, the Statesman's Year Book and the International Almanac of Electoral History – although they may

use specialist materials to supplement them. Newspapers are another common source, with Banks remaining heavily reliant upon the New York Times, and Bond *et al.* similarly dependent on Reuters. In the case of Freedom House the main sources are never obvious.

It is therefore clear that most measures use 'western' reference materials and news sources. These sources may be the most complete, and certainly the most 'available.' But their almost exclusive use does introduce a systematic bias into performance measures. Thus, it cannot be claimed that the New York Times provides comprehensive and consistent foreign news coverage, when Reuters covers ten-times as much foreign news (Bond, Craig Jenkins, Taylor and Schock, 1997, p. 567), and countries and issues go in and out of fashion.<sup>13</sup> In the case of times-series measures over the long term (Banks, Jagers and Gurr), it seems unlikely that equally reliable data is available to support the coding of measures in the 1890s and 1990s. None of this invalidates the extant attempts to produce measures of comparative democratic performance, but it does mean that the measures should be viewed with a certain caution.

### **Weighting**

The weighting of variables that enter the construction of aggregate measures of democratic performance will closely affect the final performance score; and, insofar as the weighting emphasizes some democratic values at the expense of others, it may also affect the validity of this score. Yet the weightings within aggregate measures are rarely justified, and often appear to be decided for reasons of statistical presentation, or for no reason at all.<sup>14</sup> In this regard, an equal weighting of variables is itself a decision that requires justification: Freedom House appears to assume that it is 'natural' to average the scores of its Political Rights and Civil Liberties scales to create its global map of Free, Partly Free and Not Free countries in the Comparative Survey of Freedom.<sup>15</sup> It seems that weightings, whether explicit or implicit, often remain unjustified because they are so difficult to justify. Justification may entail judgements about the relative importance of democratic values.

Measurement techniques may also influence weighting, more or less directly. Variables with a greater range of measurement will tend to trump those with a lesser range. In *Polity III's* Index of Democracy the Executive Constraint variable has double the range of Competitiveness of Executive Recruitment, and emerges as the single main determinant of the aggregate democracy and autocracy scores (Gleditsch and War, 1997, p. 380). There is therefore a danger that 'the available empirical material will come to dictate the relative significance of the various democratic components'.<sup>16</sup> Furthermore, statistical methods may have similar effects. Jackman's use of natural log values for his Political Deaths variable places more weight on lower scores than higher ones (Jackman, 1991, 178). Vanhanen's decision to multiply the two equally-weighted variables of Participation and Competition to produce his aggregate Index of Democratization tends to mark down those countries with rather different scores on the two variables in favour of those with more equal scores. In all these cases the critical moment is that of aggregation, and it may be preferable to leave measures as separate as the research design will allow.<sup>17</sup>



### *Measurement level*

The two main ways of constructing democratic performance measures may be characterized as the 'subjective' and the 'objective.' The great majority of measures are subjective, entailing the coding of countries into the categories of a scale, and yielding ordinal (or categorical) level indicators. The coding is undertaken by an individual or group (Freedom House, Humana) onto scales that usually have between two and five categories. Banks scores Legislative Effectiveness from nought to two, Cingranelli and Wright rate Extensiveness of Due Process from one to four, and Poe and Tate place human rights records in an index of one to five. Exceptionally, ordinal scales may have as many as ten categories. Objective measures, on the other hand, apparently eschew individual or group judgement, since there should be no room for debate about measures like the Percentage of Minority Cabinets or Electoral Turnout. Institutional and electoral measures of this kind may have a large range (for example Number of Effective Electoral Parties), or be expressed in percentage terms (for example Frequency of Parliamentary Majorities).<sup>18</sup> And events data that counts the incidence of specific events like riots and demonstrations (Banks) may have an unlimited range. The 'objective' measures tend to yield interval or ratio level indicators.

Commentaries on the subjective measures have identified problems of measurement range, intersubjectivity, and plain political bias. It is argued that the restricted range of scales with just three or four categories often proves insensitive to real variation and so introduces measurement errors. Thus, Coppedge and Reinicke's three-point scale is not nuanced enough to capture the wide variation in the practice of Free and Fair Elections (Bollen, 1991, p. 13).<sup>19</sup> It may be legitimate for Przeworski *et al.* to use the simple dichotomy of democracy and autocracy, since their focus is the relationship between development and democracy, and a bimodal distribution will differentiate richer and poorer countries. But more sensitive scales are needed in order to differentiate the democratic performance of 'first world' democracies, or they will all receive the same top score (a comforting but unrealistic result). In fact, this is exactly what occurs even with the Freedom House seven-point Political Rights index in the Comparative Survey of Freedom, leading Diamond to argue that it lacks sensitivity at both ends of the scale.<sup>20</sup>

Problems of intersubjectivity refer to the difficulties that coders may have in establishing the meaning of key terms such as 'free and fair elections,' or 'civilian control of the military' (let alone the 'real power' of legislators or the 'real representation' of minority interests), and in applying them to particular political circumstances. Reliability tests have demonstrated that the degrees of discrepancy between coders can be disconcertingly high.<sup>21</sup> Logically, the same problems of intersubjectivity may occur between coders and the 'consumers' who try to make sense of categories like 'unlimited authority' and 'executive parity or subordination' in *Polity III's* Executive Constraint scale (Jagers and Gurr). In this instance, consumers cannot always be clear what is meant by the qualitative but narrow differences between different degrees of Constraint.<sup>22</sup> In general, the relative vagueness of coding criteria often makes it difficult either to compare different subjective measures of the same phenomenon, or to replicate measures in a different time or context.<sup>23</sup>

The problem of political bias may take a general or a particular form. General biases may include those of Freedom House towards right-wing regimes, the US State Department against left-wing governments (for example the Soviet Union or North Korea), and Amnesty International in favour of them. Examples of more particular, but avowed, biases include those of Freedom House towards Poland, South Africa and Yugoslavia, and of Arat against Turkey and Israel.<sup>24</sup> Other examples of apparent political bias seem to derive from ignorance rather than deliberate distortion, such as Banks' mistaken coding of the governments of Guatemala and El Salvador in the 1970s as civilian and not military or civilian-military.<sup>25</sup>

Objective measures, on the other hand, are designed to avoid the evaluative problems inherent in the subjective measures by counting simple 'facts,' such as a percentage of votes or a number of parties. In the view of many researchers, such as Vanhanen, this methodological choice can and does resolve or diminish the problems of subjective measurement. But the 'facts' can prove intractable, since their definition will often depend on subjective judgement. The sophisticated measures of the number of parties in a political system, such as Rae's Index of Fractionalization or Laakso and Taagepera's Effective Number of Parties, cannot 'decide' if divided parties should count as one or two, or closely allied parties as two or one; or if the number of 'relevant' parties should really depend on their 'blackmail' or 'coalitional' potential (Sartori, 1997, p. 24). Similarly, it is unclear if parties should be defined as 'extremist' on the basis of their ideology or of the role they play in the political system. It may even be unclear whether a fact as stark as a 'political execution' should be counted, depending on whether the death occurs in government custody or not.<sup>26</sup> In these examples, and many others, the objective measures remain imbued with subjectivity.<sup>27</sup>

The counting itself can also be a problem, especially for those scholars who construct indicators using events data (for example Powell, Arat). Since the events themselves cannot be counted, recourse is had to secondary sources such as newspapers, that may or may not record many or few of the events in question. This problem is acute in the field of human rights, since censorship can restrict or eliminate the information, and governments are not keen to record or admit their own repressive acts. For this reason Amnesty International has always refused to compare the human rights records of different regimes, since it is 'impossible to establish a reliable and consistent basis for comparison'.<sup>28</sup> Thus, the evidence is hard to get, and, in the case of the most notorious crimes against humanity, often arrives a long time after the 'fact'. Analogous problems may affect the recording of social protest events. Furthermore, it cannot be safely assumed that an absence of repressive 'events' implies an absence of repression, since the threat of repression may be enough to impose political quiescence.

It is therefore clear that objective and subjective measures share problems of both definition and evidence, and that the difference between them is sometimes more one of degree than of kind. And they are both obliged to impose apparently arbitrary thresholds for the presence of democracy. Thus, countries are 'free' if their aggregate Civil Liberties and Political Rights score is between one and two point five (in the Freedom House Comparative Survey of Freedom); or 'democratic' if they score more than five, with at least thirty percent competition and fifteen

percent participation (Vanhanen), or score seven or above on Polity III's eleven-point scale (Jaggers and Gurr). But their common problems do not make it easier to compare them. In particular, it is difficult to compare ordinal and interval level measures of performance, and yet more difficult to compare ordinal measures derived from subjective coding with interval measures based on counting events or 'facts.' Yet, so long as the measures are kept separate the problem is simply one of heterogeneity. But when they are subsumed within an aggregate score, the result may be incoherence.<sup>29</sup>

### *Testing for Validity*

The most common claim for the validity of democratic performance measures is based on high statistical correlations with other such measures.<sup>30</sup> Arat constructs a table of correlations between her Index of Democraticness and the measures of Dahl, Jackman, and Bollen, amongst others. Jaggers and Gurr (1995, pp. 473–4) find that the *Polity III* Democracy score correlates at around 0.90 with those of Freedom House, Vanhanen, Arat, Coppedge and Reinicke, Bollen and Gasiorowski, allowing them to conclude that their measures 'accurately represent democracy'. Yet highly correlated measures may all contain the same errors, all 'share similar biases',<sup>31</sup> or all be determined by outside influences that may render their close association spurious. The correlations cannot therefore guarantee that the different measures are all quantifying the same underlying value or concept, since 'we are not sure what the underlying concept is', and hence cannot prove their validity (Fedderke and Klitgaard, p. 7).

The statistical basis of these high correlations can also be suspect since they generally assume that the data are normally distributed. But most of the datasets are not so distributed, often exhibiting significant skewness. Bollen notes this potential problem, and Fedderke and Klitgaard argue that non-normal distributions can either inflate or deflate correlation results. Furthermore, most validity tests use interval level data techniques like Pearson product-moment correlations on data that are mainly ordinal.<sup>32</sup> But the vagueness of the criteria that support the coding of most ordinal variables precludes the precision of interval measures, so that it is impossible to verify whether a move from one to two (ineffective to partially effective) on Banks's three-point scale of Legislative Effectiveness, for example, is exactly equivalent to a move from two to three (partially effective to effective).

Finally there is some evidence that comparisons of performance measures are manipulated in order to boost correlation results. Time frames are varied to suit particular purposes. Thus, Gasiorowski (1996, p. 478) uses Jaggers and Gurr data from 1800 to 1986, whereas Jaggers and Gurr (1995, p. 475) use Gasiorowski's data from 1946 to 1992, with the result that their correlation coefficients are different. Variables are chosen for their particular properties. Thus, Jaggers and Gurr use their Democracy minus Autocracy rather than their Democracy score for their comparisons, since the former has a more bimodal distribution, and can therefore boost the results by spreading the observations along the continuum (Jaggers and Gurr, 1995, 473). Even without manipulation, it is likely that the global tests obscure regional variations, since the high degree of congruence among performance measures for the established western democracies tends to inflate the overall

correlation coefficients. Consequently, regional comparisons might do more to challenge the 'validity' of the measures. Thus, Lane and Ersson suggest that Freedom House (Comparative Survey of Freedom) and Jagers and Gurr differ markedly in their classification of African countries, and to a lesser extent of socialist countries, while Bollen argues that the Freedom House scores are relatively higher for Latin America and that the Banks scores are similarly skewed in favour of Eastern and Central Europe (Lane and Ersson, 1997, p. 94; Bollen, 1993, 1221–3).

## The conceptual critique

The empirical critique of the measures marks out their geographical and temporal range, as well as revealing their structural and operational features. But the critique of their 'validity tests,' in particular, suggests that their validity may only be judged according to the liberal democratic *values* that inform them. Identifying these values is not always easy. The bold claims about measuring liberal democratic performance are rarely matched by proper attention to the particular liberal democratic values that the measures must be designed to reflect.<sup>33</sup> It therefore makes sense to make these values explicit, and to assess the measures according to the core values of liberal democratic government. The aim is to gauge the normative range of the measures, or their effective coverage of distinct aspects of performance, as well as their normative depth, or their capacity to 'capture' liberal democratic values.

## The conceptual framework

The premise of this inquiry is that there is a broad consensus on the foundational principles of liberal democracy, and substantial agreement on the institutional and legal means for achieving them. The intellectual grounds for the consensus have been created by long traditions of liberal democratic theory, and the modern development of this theory begins in seventeenth-century England, and in the encounter and conversation between liberal and democratic thinking. In subsequent centuries these hitherto distinct traditions moved progressively closer together in the common context of the defense of private property and market relations. It is far too simple to suggest, however, that the theory subsequently gave rise to liberal democratic practices.<sup>34</sup>

The classic statement of liberal principles is found in Locke's *Second Treatise*, where he argues that legitimate government must reproduce the conditions of 'perfect freedom ... and also of equality' found in the state of nature. But Locke insisted on linking his idea of (negative) liberty to property, so that the purpose and justification of political rule is to protect property in the broad sense of 'lives, liberties and estates'. Despite his defense of freedom and reason, Locke was no democrat, believing that liberty and equality could be upheld under monarchy or mixed government, so long as the rule of law is respected. The constitutional protection of individual liberty and equality under the rule of law has remained fundamental to liberal theory ever since (Locke, 1924, pp. 118, 127, 180, 183).<sup>35</sup>

Yet, a democratic impulse did exist in seventeenth-century England, and not just in the politics of the Levellers (Wootton, 1992). This impulse was developed by a

number of republican writers, including Nedham, Harrington and Milton, who were unconvinced that the rule of law was sufficient to protect individual liberty. Locke had argued that the rule of law and limited government was enough to underpin (negative) liberty. But the republicans, or neo-Romans in Skinner's term, remained sensitive to 'the danger of being forcibly or coercively deprived by (their) government of (their) life, liberty and estates'. Hence, it was inescapable that 'the government of a free state should ideally be such as to enable each citizen to exercise an equal right of participation in the making of laws' (Skinner, 1998, pp. 30, 69–70). Thus, the principle of self-rule, or the sovereignty of the people, was enshrined as a necessary condition for liberty and equality. By making government accountable to the people, it guaranteed it would uphold the law, so supplying the essential democratic link to liberal democracy.

As intellectual history, these bare statements are plainly inadequate. Their more modest intent is to provide some warrant for the claim that liberal democracy is rooted historically in the principles of liberty and equality, and secured by the rule of law and the self-rule of the citizenry. Furthermore, as liberal democratic thinking developed, there emerged a strong degree of consensus over the particular values that are required to achieve and reproduce these principles. A brief justification of these values, both institutional (accountability, constraint, representation, participation) and legal (civil rights, property rights, political rights, minority rights) will prepare the ground for the subsequent inquiry into the normative range and depth of the measures of liberal democratic performance.

### *Accountability*

It might be maintained that even John Locke was concerned with accountability, insofar as government only rules by consent of the people. But it is something of an 'alpha and omega' accountability, since consent is required at the moment of the original contract or is finally withdrawn if the government encroaches on 'the liberties and properties of subjects' (Locke, 1924, 192). Locke was not thinking of the consent reproduced through recurrent elections of representatives or direct participation. The republican tradition carried a stronger sense of accountability, which bore fruit in Thomas Paine's advocacy of 'government by election' and in the Declaration of the Rights of Man that 'every community has a right to demand of all its agents an account of their conduct'. This was the sense developed by Bentham, James Mill and Madison into what Macpherson and Held refer to as the 'protective' model of liberal democracy.<sup>36</sup> Domination by others can be prevented by 'accountable institutions' such as 'the secret ballot, regular voting and competition between political representatives' (Held, 1996, p. 88). Modern political theorists have accepted that the accountability delivered through electoral processes and other means is a basic requirement of liberal democracy (Dahl, 1989).

### *Constraint*

Constraint is usually understood to require the separation and balance of government powers, namely legislature, executive and judiciary, to prevent 'the same men and the same body' from enacting laws, executing public resolutions and

trying the causes of individuals (Held, 1996, p. 85, *après* Montesquieu). But the concern with constraint did not necessarily entail a commitment to democracy. Montesquieu himself advocated a separation of powers as an aristocratic bulwark against the will of the majority; and other eighteenth-century liberal theorists argued for mixed government to constrain monarchical prerogatives (Maier, 1992, p. 133). Nonetheless, the principle of constraint has been accepted as a core value of liberal democracy, partly through its early embodiment in the constitution of the United States. Modern formulations adhere closely to the classic statements, with Diamond (1997, p. 9) noting that that 'executive power is constrained, constitutionally and in fact, by the autonomous power of other government institutions (such as an independent judiciary, parliament, and other mechanisms of horizontal accountability)'.

### ***Representation***

The idea of representation is not confined to liberal democratic thought. But it was the advent of representative government in the eighteenth-century, characterized by Dahl as the 'second democratic transformation' (the first being the formation of democratic city-states), that allowed liberal democracy to evolve into a system of government for the modern age (Dahl, 1989). There was no early theory of representative government, or certainly not one that enshrined the principle of election rather than appointment. It was practice that promoted the later commonplace of liberal democratic thought that representative government is elected government. J. S. Mill concurred that the system of representation 'was the grand discovery of modern times.' But the idea of representation did not resolve whether persons or interests should be represented. While the civic republicans were clear that it was persons, J. S. Mill argued that the government must be directed by an elected assembly with 'an identity of interest with the community' (Birch, 1964, pp. 16, 46–7). Such an aspiration seems unpromising, being beset by both practical and philosophical problems, and it has not proved to be a viable basis for representative government. But the idea of representing interests has been reflected in the very different thinking of pluralists, neocorporatists and consociationalists.

### ***Participation***

It might be argued that direct participation was the core value of the ancient Athenian model of city-democracy. With the passage to representative government and the advent of mass society, participation began to seem either difficult or dangerous. Either participation became so attenuated that there were doubts whether democracy was now attainable, or mass participation threatened democracy with authoritarianism, and ultimately fascism (Pateman, 1970, p. 2).<sup>37</sup> And scepticism about 'too much' participation is still current among theorists of elite or deliberative democracy. Yet Mill saw participation as essential to the moral improvement of democratic citizens, and judged political institutions according to their promotion of participation and 'the great community' (Mill, 1910, pp. 277–9). This 'developmental' or 'civilizing' model of liberal democracy remained strong early in this century (Macpherson, 1997, p. 69), but recent justifications of (electoral) participation have been grounded in political equality. Low electoral turnout

'is systematically biased against less well-to-do citizens', and low levels of participation make elected officials less responsive (Lijphart, 1997, p. 1).

### ***Civil Rights***

Early liberal thinking was constantly concerned to protect personal liberty. The rule of law was required to prevent state interference with the autonomous individuals of civil society. The classical statements of Hobbes and Locke referred to the absence of such interference as liberty, and the subsequent Declarations of the Rights of Man embodied the same objective of protecting citizens from arbitrary power.<sup>38</sup> J. S. Mill's *On Liberty* sits squarely in this tradition, arguing that the only justification for interfering with the liberty of the individual is 'to prevent harm to others' (Mill, 1991, pp. 16–17). Modern theory refers to such liberty as 'negative liberty,' since 'I am normally said to be free to the degree to which no man or body of men interferes with my activity' (Berlin, 1969, p. 122). The cluster of laws that today protects such negative liberty is characterized as civil rights, and enshrined in the UN Universal Declaration of Human Rights and the International Covenant on Civil and Political Rights. Equality before the law, due process, and *habeas corpus* are all civil rights that are intended to protect the individual from such interference and abuse as arbitrary detention, torture, and 'disappearance,' while freedom of religion and marriage are bulwarks of personal autonomy.

### ***Property Rights***

Liberal thinkers see the protection of private property as a key requirement of negative liberty. Indeed, democracy itself was only acceptable to the likes of Bentham and James Mill once they were assured that it would not threaten property rights or the free market. Eventually, most liberal theorists came to believe that an extension of the franchise – if constrained by property qualifications – would placate social protest and provide the political stability for the free market economy.<sup>39</sup> Contemporarily, property rights are explicitly defended in Article 17 of the UN Declaration of Human Rights, and polemicists from Friedman to Fukuyama consistently assume that property rights are fundamental to individual freedom. These beliefs appear to be confirmed by the spate of 'dual transitions' from authoritarian to democratic regimes, and from state-controlled to free-market economies in the 1980s and 1990s.

### ***Political Rights***

In liberal democratic thought it is axiomatic that civil rights and property rights are necessary but not sufficient to protect individual liberties (and equality before the law), since the rule of law must itself be protected by making government accountable to its citizens, who must therefore enjoy political rights. By the early 19th century theorists like Bentham and James Mill had come to see that citizens were best protected by free and fair elections, with secret ballot and freedom of the press (Macpherson, 1977, p. 34). Yet the franchise remained restricted by gender, race and property qualifications well into the present century, with the restrictions defended by liberal thinkers such as J. S. Mill. In subsequent years liberal theory

was gradually purged of first its race and then its gender biases (Pateman, 1989, pp. 3–4), and there is now a strong liberal consensus that all adults should enjoy full political rights. According to the Universal Declaration of Human Rights, these must include freedom of speech, association, assembly, movement, information and press, as well as universal suffrage and the secret ballot. These rights compose one of the most common ‘checklists’ of liberal democratic performance among political scientists, from Dahl’s *Polyarchy* to the present.

### **Minority Rights**

Although liberalism and individualism grew up together, they are distinct doctrines (Raz, 1986, pp. 17–18). Nearly all liberals believe that liberty and equality require the provision of civil rights, property rights and (individual) political rights. Some liberals also believe that liberty and especially equality require ‘group-differentiated’ rights, or minority rights (Kymlicka, 1995, p. 45). This is justified by the argument that liberalism assumes collective goods. Rights may have been defended in the name of individual freedom, but ‘would not have served their avowed purpose’ without these goods.<sup>40</sup> Kymlicka can then claim a liberal pedigree for minority rights, asserting ‘that individual freedom is tied in some important way to membership of one’s national group, and that group-specific rights can promote equality between the minority and the majority’. In his view, modern liberal thought has defended these rights, although contemporary liberals are ‘surprisingly silent’ about them.<sup>41</sup> The silence stemmed from the theoretical fusion of liberalism and individualism that led to the conclusion that civil rights and political rights were sufficient to protect national and ethnic minorities. But minority rights are now increasingly important to debates about the quality of democracy, and so should be included in measures of liberal democratic performance (Linz, 1994, p. 23).<sup>42</sup>

### **The Normative Range and Depth of the Measures**

These liberal democratic values may now serve to gauge the scope or conceptual reach of existing performance measures. A review of these measures immediately reveals that many reflect just two values, namely accountability and political rights. The early boundaries were set by Dahl’s seminal work on *Polyarchy*, with its emphasis on public contestation and participation, and with participation understood as the (political) *right* to participate, rather than the degree of real participation. Since democracy is primarily concerned with ‘the continuing responsiveness of the government to the preferences of its citizens’ (Dahl, 1971, pp. 1–2), it mainly requires freedom of association, expression, political opposition and information, free and fair elections, the franchise, and institutions for making policies depend on votes. Dahl then ranks one hundred and fourteen countries according to the greater presence or absence of these elements of ‘polyarchy’ in 1969, but the ranking ignores civil and minority rights, constraint and even real participation.<sup>43</sup>

It is apparent that Dahl’s assumptions (and, implicitly, those of Downs, Lipset and eventually Schumpeter) have exerted considerable influence over subsequent studies of liberal democratic performance. Bollen (1993, p. 1209) again focuses on accountability (popular sovereignty expressed through free and fair elections, and



open selection of executive and legislature) and political rights (press freedom, freedom of opposition, and the absence of government sanctions), and maintains this focus in subsequent work. Coppedge and Reinicke set out to replicate Dahl's study for 170 countries in 1985 with their four main variables of free and fair elections, freedom of organization, freedom of expression, and availability of alternative sources of information.<sup>44</sup> Hadenius (1992) ranks one hundred and thirty-two countries in 1988 according to indicators of universal suffrage, meaningful elections, organizational freedom, freedom of opinion, and political opposition. Similarly, Gasiorowski's Political Regime Change Dataset (1996, p. 471) defines democracies by meaningful and extensive competition, no barriers to participation, and enough political rights to guarantee both things, while Przeworski *et al.* (1996, p. 39) again invoke Dahl's criteria to justify their democratic cases.

The focus on accountability and political rights means that these measures have a (relatively) narrow normative range. In addition, whatever their range they tend to be aggregated into single performance scales, as if democratic performance were one-dimensional. All the above studies construct summary measures of this kind, as do most others.<sup>45</sup> Although such scales are an advance on a dichotomous classification of democracies and non-democracies (for example Lipset), they remain insensitive to specific aspects of liberal democratic performance. In particular, 'a single index obscures any pattern of interaction among the component variables, reducing rather than increasing the information available' (Barsh, 1993, p. 102), and disallowing any perception that regimes may be 'differently democratic'.<sup>46</sup> The advantage of a multidimensional approach, in contrast, is that the separate measures may be used to construct performance 'profiles' across a range of values that may trade-off against each other (as demonstrated by Powell). In *Polyarchy* Dahl sets out to measure performance across the two dimensions of contestation and participation, but ends up by constructing a single polyarchy scale.<sup>47</sup> In recent years Dahl appears to disavow the possibility of constructing such a scale.<sup>48</sup>

A number of studies have extended the scope of Dahl's *Polyarchy* by including measures of constraint and civil rights. The Polity III data-set is designed to respond first to 'the presence of institutions and procedures through which citizens can express effective preferences,' but it also sets out to measure 'the existence of institutionalized constraints on the exercise of executive power'.<sup>49</sup> Analogously, the Freedom House Comparative Survey of Freedom employs the usual indicators of accountability and political rights, but also addresses civil rights in the forms of freedom from terror, imprisonment, exile and torture; freedom of movement, marriage and religion; and equality before the law (Gastil, 1987, p. 4).<sup>50</sup> And Arat moves beyond Participation, Inclusiveness of the Process, and Competitiveness to include Civil Liberties as one of her four components of 'democraticness'.

Different studies conceive of civil and political rights in distinct ways, and their analytical status varies as a consequence. For Gasiorowski civil rights are only important if they contribute to political rights and accountability. For Pougarami civil and political rights are a direct measures of liberal democratic performance. For Poe and Tate, in contrast, there is a rigorous distinction between 'human rights to personal integrity' and their democracy measures, since they seek to test the relationship between them.<sup>51</sup> Humana, very differently, combines institutional

values with forty different rights measures that he divides into six *sui generis* categories. Yet other studies eschew general rights measures in favour of particular measures of political rights like press freedom (Van Belle), or of civil rights like due process (Cingranelli and Wright).<sup>52</sup> The more precise focus on an issue like due process has the advantage of highlighting the importance of the judiciary in maintaining the rule of law, which is fundamental to both political rights and the mechanisms of accountability.

Yet, despite some extension of the normative range, most of these studies continue to construct single performance scales, with one or two more or less honourable exceptions. Since Polity III provides both democracy and autocracy scores across a number of values, all regimes will possess 'mixed' authority characteristics, and may reach their 'threshold' democracy score in different ways (Jagers and Gurr, 1995, p. 472). But, although the different values may be presented separately, there is still a strong disposition to add the scores together in a single performance scale. (Humana also provides raw scores for his variables but ends up by aggregating them). Scores on this eleven-point scale represent the sum of five selected performance measures. Nonetheless, Polity III, Freedom House, Arat and Pougarami do extend the range of values encompassed by the measurement of liberal democratic performance. And the explicit emphasis on rights measures of Poe and Tate, Humana, Van Belle, and Cingranelli and Wright provides a salutary corrective to exclusively 'institutional' approaches to such performance.

## **Assimilating the Institutional and Rights Measures**

This review of the existing performance measures reveals that they tend to cover accountability, constraint, political rights and civil rights, or just four of the eight core values of liberal democratic government. The other four values, namely property rights, minority rights, participation and representation, have been subject to measurement, but seldom for the explicit purpose of comparing liberal democratic performance. These measures appear, severally, in studies of international relations, the economics of corruption, electoral systems, and institutional design in general. Yet there is no reason of theory or method why they should not be assimilated to the study of liberal democratic performance, or why they cannot contribute to construct more robust performance measures by extending their normative range and empirical reach.

### ***Property Rights***

Property rights are central to liberal democratic thought and practice, so it is odd that most measures of democratic performance ignore them completely. It is true that Freedom House's indicator of Civil Liberties does include property rights (and freedom for business), but, since the raw scores of the indicator's components are unavailable or simply absent, it is impossible to discern the contribution of property rights to the aggregate score. Although Freedom House would certainly deny that it rates countries primarily by their free-market credentials, many economists use the Civil Liberties score as a measure of property rights when investigating the causes of growth and investment (Knack and Keefer, 1995, p. 208).

Yet some ordinal proxy measures of property rights have been compiled by investment managers and consultants in the form of indicators of political corruption and distortion of the marketplace. It is self-evident that political corruption can infringe property rights and distort free and equal access to the market under the rule of law. Common examples are the demand for bribes in return for government contracts, or the requirement of 'special payments' to obtain export licenses,<sup>53</sup> but virtually any illegal bureaucratic delay or any failure to comply with the law of contract can have the same effect. For this reason Political Risk Services supplies data on Government Repudiation of Contracts, Risk of Expropriation and Corruption; and the Transparency International Corruption Perception Index provides time-series corruption data for fifty-two countries, based on the average scores of six survey sources.<sup>54</sup> Knack and Keefer develop two indicators, explicitly designed as indices of property rights. One includes measures of corruption in government, of quality of bureaucracy, and of rule of law.<sup>55</sup> The other uses the Business Environmental Risk Intelligence measure of bureaucratic delays (compare Mauro, 1995, pp. 684–6).

### *Minority Rights*

Studies of democratic performance rarely refer to minority rights. The theoretical fusion of liberalism and individualism in recent decades has favoured measures that subsume minority rights into more general categories, and so are unable to distinguish degrees of liberty and equality among minority groups on the one hand, and majority populations on the other. These categories include Freedom of Organization (Coppedge and Reinicke), Inclusiveness (Arat), and Political Rights (Freedom House Comparative Survey of Freedom), where minority rights are dissolved into Political Rights just as property rights were joined to civil rights. Humana measures minority rights, but eschews any indicator of political discrimination. Lijphart measures 'minority representation and the protection of minority interests', but addresses women as a minority rather than ethnic or national minorities, and emphasises institutional attainments, such as the provision of family policy, rather than minority rights proper (Lijphart, 1994b, p. 4).

The Minorities at Risk data-set (Haxton and Gurr) is not directly concerned with liberal democratic performance. But the premise of this study of 268 minority groups in 112 countries is that, since the end of the Cold War, 'minority peoples ... have become the principal victims of gross human rights violations' (Gurr, 1993a, p. 314). Against this background, Haxton and Gurr set out to measure the cultural, economic and political rights of minorities by constructing indicators of Political Discrimination, Cultural Restrictions, Rights in Judicial Proceedings, and so forth, that reflect liberal democratic values amongst minorities. Liberal democracy provides no automatic guarantee against such abuse because it is 'susceptible to the politics of ethnocentric reaction', and indicators of infringements of minority rights could compose an effective measure of liberal democratic performance (Gurr, 1993b, pp. 161–201; Gurr, 1993a, p. 322). There is good reason, therefore, to include at least some elements of this disaggregated data-set into a more comprehensive performance profile.

## *Participation*

Prior to 1980 it was quite usual to take rates of electoral participation as a measure of democratic performance (Bollen, 1980, p. 373). But more recent studies follow Dahl in arguing that the key liberal democratic value is the right to vote, and that actual voter turnout levels are irrelevant to democratic performance. Indeed, low turnout may be a sign of citizen satisfaction and, by extension, of strong performance. Conversely, high turnout might be associated with electoral fraud, or at least 'engineered' in some way. Furthermore, compulsory voting, on the one hand, and legal or political restrictions on voter registration, on the other, might distort the comparative measures. Notwithstanding these strictures Powell has continued to use turnout as a performance indicator because 'without significant citizen involvement the democratic process falls short of its goals', and Lijphart does the same for much the same reasons. Vanhanen too incorporates national electoral participation in his Index of Democratization, while Beetham suggests that turnout should be part of an index of 'electoral democracy'.<sup>56</sup> Moreover, most electoral analysis tends to assume that liberal democracies are performing well if electoral participation is high (Franklin, 1996, p. 216). It therefore seems reasonable to take real turnout rates as a measure of democratic performance, alongside indicators of the right to vote. It might then be possible to verify whether performance is improving in terms of political rights (the right to vote), while it is declining in terms of participation (turnout), or vice-versa; or whether both measures are moving in the same direction.

## *Representation*

Most of the comparative institutional literature does not directly address liberal democratic performance, but is mainly concerned with variations in institutional design, and especially in party and electoral systems. Yet this literature can and does supply implicit performance measures, as demonstrated by Lijphart's original inquiry into the institutional characteristics of twenty-one 'fully democratic regimes'. His factor analysis of measures such as Electoral Disproportionality, Effective Number of Parties and Executive Dominance, reveals 'two diametrically opposite models of democracy,' namely consensus and majoritarian, and he has no doubt that it is the former that delivers superior performance, at least under certain conditions (Lijphart, 1984, p. 3, 95). His subsequent study of eighteen established democracies, using different measures, finds that consensus democracy is the equal of the majoritarian model in 'maintaining public order and managing the economy,' but achieves 'superior representation' (Lijphart, 1994b, p. 1).<sup>57</sup> It follows that those institutional characteristics that distinguish consensus from majoritarian democracy are *themselves* measures of liberal democratic performance, and that several of the 1984 indicators, such as Minimum Winning Cabinets, Effective Number of Parties and Electoral Disproportionality are measures of performance in representation.<sup>58</sup> In a similar fashion, Powell had demonstrated, *inter alia*, that majoritarian electoral systems provide better executive stability than more proportional systems, and that multiparty systems tend to enjoy high levels of participation and low levels of violence. In sum, the 'elements in constitutional design have a substantial impact on democratic performance' (Powell, 1982, p. 54, 72, 109).<sup>59</sup>

It is also argued, following Juan Linz, that parliamentary regimes perform better than presidential ones. Stepan and Skach (1994, p. 123) use Vanhanen's Index of Democratization to show that – after controlling for levels of development – parliamentary regimes will 'overachieve' democratically, while presidential ones will 'underachieve'. Since the scores on the Vanhanen Index are a multiple of voter turnout and the percentage of seats of all but the largest party, Stepan and Skach plainly imply that parliamentary regimes achieve better participation and representation. The nominal division between parliamentary and presidential is in fact a two-point ordinal scale of performance that may include 'mixed regimes' as a third point. Shugart and Carey respond that value trade-offs are involved in all constitutional designs, and that presidential regimes are better able to minimize the trade-offs between accountability and representation. Once such trade-offs are taken into account, presidential regimes may deliver superior liberal democratic performance, always depending on what values 'one might want to maximize' (Shugart and Carey, 1992, p. 286).

Few of these performance claims go uncontested. But there is some evidence that more proportional electoral systems provide better representation within parliamentary regimes, that majoritarian systems secure greater accountability by clarifying responsibility, and that presidential regimes can both increase constraint on the executive and enhance accountability (Lijphart, 1994b; Sartori, 1997, p. 71; Shugart and Carey, 1992). In other words, constitutional design and party systems make a difference to liberal democratic performance. There is therefore some justification for employing measures of design and of party systems as performance indicators. Design features like the degree of electoral proportionality and of the autonomy of the executive, and party systemic indicators like the percentage of seats of the largest party, may help to achieve more robust measures of representation, constraint and accountability.<sup>60</sup>

## Conclusions

The empirical critique suggests that, in an ideal world, measures of democratic performance would be designed according to some exacting specifications. They would be time-series measures with annual scores, except where the electoral or institutional content made period averages more appropriate. They would be large-N with a broad geographical scope. They would benefit from diverse data sources, including local and regional sources, but events data would be treated with caution, and survey data with suspicion. They would strive to avoid unnecessary weightings, and seek to justify them where unavoidable. They would never combine ordinal and interval level indicators, nor impose arbitrary thresholds. There would be clear coding criteria and procedures for ordinal measures to enhance intersubjectivity, and categories would be differentiated to achieve adequate sensitivity. There would be a consistent effort to define the object of interval level measures before comparing or collating them. Finally, the validity of the measures would never be tested simply by correlating them to other measures of a similar kind but by seeking to ascertain their effective purchase on the liberal democratic *value* being measured.

The conceptual critique demonstrates that most measures of liberal democratic performance address just one or two liberal democratic values. It follows that any

comprehensive attempt to measure such performance should extend its normative range. The values deployed in this inquiry can both accommodate the available performance measures and assimilate certain institutional and rights variables into these measures.<sup>61</sup> The range of values encourages the presentation of separate measures and the construction of performance profiles that will often be structured by possible trade-offs across values. Such trade-offs may include those between representation and accountability, participation and minority rights, and property rights and civil rights. The presentation of multidimensional measures will inevitably be more complicated than the construction of single scales. But they should deliver a more rounded and robust measure of overall performance.

The way to hell is paved with good intentions. Many measures of liberal democratic performance set out to measure 'liberal democracy,' or a serviceable proxy for the same, such as 'free and fair elections.' Moreover, they nearly always tend to combine component scores, whatever they may be, into a single, aggregate scale of 'freedom' or 'liberal democracy,' as if liberal democratic performance were unidimensional. Yet the clear lessons of this critique of the measures are that it is nigh impossible to establish the validity of any measure of liberal democracy *tout court*, and that the move to aggregate scores always confuses more than it clarifies. It follows that studies of liberal democratic performance should always specify the liberal democratic value they seek to measure, and always keep their component scores as separate as conceptual clarity will permit.

(Accepted: 8 July 1999)

### About the Authors

**Joe Foweraker**, Department of Government, University of Essex, Wivenhoe Park, Colchester CO4 3SQ; email: [joef@essex.ac.uk](mailto:joef@essex.ac.uk)

**Roman Krznnaric**, Department of Government, University of Essex, Wivenhoe Park, Colchester CO4 3SQ; email: [rkrznnaric@aol.com](mailto:rkrznnaric@aol.com)

### Notes

We wish to acknowledge the support of the Economic and Social Research Council for research into *Comparative Democratic Performance: Institutional Efficacy and Individual Rights*.

- 1 Diamond (1997, p. 22) takes Freedom House data to show that the number of democracies increased from 39 in 1974 to 118 in 1996.
- 2 The 'remarkable consensus concerning the legitimacy of liberal democracy' served as the premise of Fukuyama's thesis on 'the end of history' (Fukuyama, 1992, p. xi).
- 3 Lijphart's measure of 'family policy,' for example, includes maternity and parental leave, and the flexibility of retirement schemes (Lijphart, 1994b, p. 4).
- 4 At the same time, the measures are assailed by cultural relativism, or the claim that democracy means different things to different peoples. This claim is examined in detail in Foweraker and Landman (1997, ch. 1, esp. pp. 19–21).
- 5 In our view this is a perfectly proper methodological move, and a necessary one. But it is important to keep the terms of the discussion clear. Much of the institutional literature will refer to the attributes or properties of democratic government *per se*, and so cannot be imported uncritically into an analysis based on the degree to which specific liberal democratic values are realized in practice.
- 6 We chose not to derive these values from citizen preferences. This methodological choice is justified by our very incomplete knowledge of such preferences across the globe, and by our lack of confidence in the comparative application of the existing survey data on these preferences. Where such data is

- applied it often addresses the relationship between citizen preferences and specific policy outputs, not democratic performance, e.g. Huber and Powell (1994, pp. 291–326).
- 7 Powell (1982) was one of the first to examine the potential trade-offs across democratic values within a comparative approach to liberal democratic government.
  - 8 By extension, they provide only a rather constricted view of democracy's 'third wave'.
  - 9 Thus, it is important to ask, for example, whether voter turn-out is an adequate measure of political participation, or whether the 'effective number of parliamentary parties' is a proper measure of representation.
  - 10 Like most cross-national measures, some time-series measures are designed for causal analysis (e.g. Arat and Vanhanen).
  - 11 It is recognized that there are large-N studies in a single country or with few country cases, 'N' simply referring to the number of observations. This essay focuses on studies where the large-N is created by many country cases simply because it concerns the *comparative* performance of all contemporary liberal democratic governments.
  - 12 A small group of people coding one hundred countries or more is unlikely to have in-depth knowledge of every country.
  - 13 See the data on the variable news coverage of political murder victims in different countries in the *New York Times*, *Time*, *Newsweek* and CBS in Herman and Chomsky (1988, p. 40–1).
  - 14 Thus, in presenting her measure of Civil Liberties (Government Coerciveness) Arat (1991, p. 26n40) explains that 'because of the large magnitude, the coerciveness component was divided by six before it was subtracted from the addition of the other components of democracy'. And Humana is accused of 'creating formulas which express one type of repression in terms of another' without 'any theoretical justification' (Goldstein, 1992, p. 50).
  - 15 Furthermore, Freedom House claims that there are multiple components that inform these indices, but (until very recently) have not provided raw scores for the various components. Thus, Civil Liberties are rated from 1 to 7, based on the coding of press freedom, free business or cooperatives, free religion, and a number of other components. Yet the weighting and sensitivity of these components remain obscure.
  - 16 'There is a risk, as seems to have occurred in several earlier studies (see, e.g. Arat 1985 pp. 49f. and Dahl 1971, Appendix A), of allowing the relative distribution of points to result from the number of criteria and degree of refinement of the scale, which in turn may depend on a varied quality of the data.' (Hadenius, 1992, p. 38n5).
  - 17 Lijphart (1984) is exemplary in this regard, choosing to present his eight separate indicators of consensus democracy without any attempt to aggregate them.
  - 18 Lijphart, 1994a, p. 96.
  - 19 It seems unlikely that a three-point ordinal scale can adequately reflect variations in – severally – intimidation of voters by state security forces, fraudulent ballot counts, unequal pre-election media access, prohibition of parties, discriminatory registration rules and procedures, and bribery of voters by candidates.
  - 20 For example, the following countries received perfect scores on the Freedom House Political Rights indicator for every year from 1980 to 1987: United States, Canada, United Kingdom, Ireland, Netherlands, Belgium, Luxembourg, France, Austria, Sweden, Norway, Denmark, Iceland, Australia and New Zealand. (Diamond, 1997, p. 17).
  - 21 Poe and Tate (1994, p. 867) note that their two coders, using a scale of one to five, disagreed on around fifteen percent of the cases, particularly with regard to prisoners of conscience. In Van Belle's (1997, p. 408n3) press freedom scale of one to four, coders found it very difficult to distinguish between categories one and two, and three and four.
  - 22 It was to emphasize such difficulties that Goldstein (1992, p. 50) asked, 'What does it mean to say that there is 18 times more freedom to organize a political party in one country than another?'
  - 23 In general, there is little or no information on the coding procedures themselves in Banks (1997), Freedom House (1997a) or most other studies (with one or two valiant exceptions like Haxton and Gurr, 1996). In particular, nothing may be said about how many coders are used, whether they change for different countries or time periods, what definitions are used, what is the relative importance of different variables in a single index, what are the main data sources, and so on.
  - 24 Bollen, 1993, pp. 1221–3; Poe and Tate, 1994, p. 864. According to Goldstein (1992, p. 48n46), who is quoting Gastil, Freedom House upgraded the rankings of these countries because the upgrades would 'better serve the educational purposes of the survey' and encourage 'the forces of freedom'. But it is difficult to imagine that similar considerations convinced Jagers and Gurr to award South

- Africa a 'perfect' score for Executive Recruitment Competition during the 1970s and 1980s. Arat (1991, p. 24) used data from Banks to construct a measure of Participation, but re-coded the cases of Turkey and Israel, about which she appeared to have detailed knowledge, while leaving the remaining one hundred and fifty cases in her sample unchanged.
- 25 Their chief executives were military, and the elections were fraudulent and subject to widespread state violence. Banks also fails to register the 1954 coup in Guatemala, which overthrew an elected regime and ushered in over 30 years of military rule.
- 26 The Taylor and Jodice *World Handbook of Political and Social Indicators* records no political executions in El Salvador between 1948 and 1977, allegedly for this reason (Goldstein, 1992, p. 48).
- 27 See Feyerabend's discussion of 'natural interpretations' for a theoretical reading of this observation (1993, pp. 57–60).
- 28 Amnesty International quoted in Stohl, Carleton, Lopez and Samuels (1986, p. 594).
- 29 For example, Arat's Index of 'Democraticness' = [(Participation x (1 + Inclusiveness)) + Competitiveness] – Coerciveness, where Participation, Inclusiveness, and Competitiveness are ordinal measures, but Coerciveness is an interval level measure derived from events data.
- 30 This claim is usually made alongside an equally strong but inconsistent claim that the new measure is superior in significant respects to previous measures (Arat, 1991, pp. 26–7).
- 31 Jagers and Gurr, 1995, p. 476. With regard to rights measures in particular, Barsh (1993, pp. 94, 106) points out that 'all these indices are based on roughly the same secondary sources, use similar (albeit imprecise) definitions of "rights," and employ the same kinds of panels of Western judges for scoring...[so that]...consistency in such parallel measures is an artifact of their interrelationship'. See also Bollen (1980, p. 382).
- 32 The skewness of Vanhanen's data (from Poe and Tate) is 1.149 (a skew to the right). Bollen, 1980, p. 385 n267; Fedderke and Klitgaard, p. 5; Sirowy and Inkeles, 1991, p. 140; and Barsh, 1993, p. 93. Only a few scholars, such as Gasiorowski, and Fedderke and Klitgaard, use Spearman rank-order coefficients, which are more appropriate for ordinal measures.
- 33 Hadenius (1992, p. 5) argues that most studies of democratic performance take 'the underlying criteria [of democracy] more or less for granted and instead concentrate on explaining which empirical measures and methods of enquiry will be used'. The Freedom House measures are not founded on a theoretical approach to democracy, but on the availability of information (Gastil, 1991, p. 26). But, if pragmatism is pursued at the cost of conceptual coherence, the measures can neither be compared nor contribute cumulatively to a fuller understanding of democratic performance.
- 34 Thus, in seventeenth-century England it was the execution of the King and the civil war that raised new questions about the relationship between government and people. And the growth of Protestant thought (including the doctrine of the priesthood of all believers) and the translation of the Bible into English introduced novel ideas about liberty (Hill, 1996, pp. 180, 183, 230, 245–50).
- 35 Prior to Locke there were competing uses of the concept of liberty that were unrelated to property (including that of Hobbes), but after the 1688 Revolution his view was widely accepted 'among those who mattered' (Hill, 1996, p. 261). See also Gray, 1986, pp. 11–13, 75; Held, 1996, p. 81.
- 36 Paine, 1937, p. 80, 110; Macpherson, 1977, Held, 1996.
- 37 In his early writings, Dahl argued that '... in the United States the lower one's socioeconomic class, the more authoritarian one's predispositions ... (so that) we cannot assume that an increase in political activity is always associated with an increase in polyarchy' (Dahl, 1956, p. 88, Appendix E).
- 38 Skinner, 1998, p. 10; Paine, 1937, p. 80; Robertson and Merrills, 1996, p. 6.
- 39 Gray, 1986, p. 62; Macpherson, 1962, p. 264; Macpherson, 1977, p. 10; Held, 1996, p. 95.
- 40 'Thus, while religious freedom was usually conceived of in terms of the interest of individuals, that interest and the ability to serve it rested in practice on the secure existence of a public good: the existence of religious communities within which people pursued the freedom that the right granted them' (Raz, 1986, p. 251).
- 41 Kymlicka, 1995, p. 49, 52. Yet, since group-specific rights could limit or damage the personal autonomy so dear to liberals, Kymlicka has to distinguish between the 'internal restrictions' and the 'external protections' of the group. The former can limit personal autonomy by policing internal dissent, and so are illiberal and unwelcome. But the latter can secure the collective conditions for liberty and equality by balancing the needs of the minority group and the majority society (p. 41) Such 'protections' might include boundary changes to promote fair representation, or the positive right to subsidized minority-language schooling.
- 42 Diamond (1997, p. 9) includes minority rights as one of his ten components of democratic performance.



- 43 There are virtues to this 'pluralist elitist' or 'equilibrium' model of democracy, not least its ease of comparative application. But, although Dahl readily admits that 'no large system in the real world is fully democratized', it is inescapable that a high polyarchy score might be obtained even where individuals or minorities were subject to police brutality, the executive had usurped key legislative powers, and only a handful of the population dared to turn out to vote (Dahl, 1971, 8).
- 44 They exclude the right to vote on the grounds that it does little to differentiate their cases (Coppedge and Reinicke, 1991, p. 50).
- 45 Including Arat, Freedom House, Jackman, Jagers and Gurr, Pougerami and Vanhanen.
- 46 'Since no single set of actual institutions, practices, or values embodies democracy, polities moving away from authoritarian rule can mix different components to produce different democracies. It is important to recognize that these do not define points along a single continuum of improving performance, but a matrix of potential combinations that are *differently* democratic.' (Schmitter and Karl, 1993, p. 47.)
- 47 At the very least, there is no doubt that the contestation axis is primary, with participation (or the 'percentage of adult population eligible to vote') presented as a sub-category of contestation. The contestation axis is itself constructed by adding together ten different ordinal variables into a single score, with states then ranked in thirty-one categories from 'greatest opportunity' for political opposition to 'least opportunity'. Moreover, almost the whole appendix is concerned with discussion of the contestation measures, while the participation axis only has four categories and plays a minimal role in differentiating between cases, since only nine countries out of one hundred and fourteen fall outside the categories of 'over ninety percent suffrage' and 'elections not held' (Dahl, 1971, Appendix A, pp. 232–4).
- 48 'Is it feasible to compare the relative *democracy* or "democraticness" of polyarchy in different countries? And would different countries end up differently on different indicators? Perhaps no scale is possible.' (Dahl, 1996, pp. 12–13.)
- 49 Consistent with Dahl, the expression of citizen preferences requires political competition and inclusive participation, and the political liberties to defend them. Dahl ignores constraint in *Polyarchy*, but had earlier emphasized its importance in *A Preface to Democratic Theory* (pp. 27–30). Jagers and Gurr concentrate on 'institutional' values. They did include civil rights in their definition of democracy, but not in their measures, because of lack of data (1995, p. 471).
- 50 The Freedom House distinction between Political Rights and Civil Liberties is not always clear. In particular, some indicators of Civil Liberties, such as freedom of assembly and organization, may seem to reflect political not civil rights. Property rights are categorized as Civil Liberties, but, although this may be justified historically, it is helpful to maintain the conceptual distinction between the two.
- 51 They take measures of accountability, participation, and representation from the Freedom House Political Rights Index and Vanhanen's Index of Democratization. Their research design excludes civil rights from their definition of democracy. For a methodological critique of their analysis see McCormick and Mitchell (1997, p. 511).
- 52 They code countries on five aspects of due process: pre-trial guarantees against unreasonable searches, arbitrary arrest for imprisonment, arbitrary abduction, torture, and fairness of trials.
- 53 Political corruption can only be defined in relation to the public purposes of political rule (Philp, 1997, pp. 453–4). Some instances, like private payments to elected representatives, may mainly damage accountability. Others, like bribing bureaucrats to obtain government contracts, will directly impede property rights.
- 54 The surveys tend to target businesspeople who may claim to have special knowledge. But the small samples of respondents with potentially very different conceptions of corruption can lead to a high variance in scores.
- 55 The data come from Political Risk Services' International Country Risk Guide.
- 56 Bollen, 1991, p. 4; Hadenius, 1992; Powell, 1982, p. 12; Lijphart, 1994b, p. 4. Vanhanen advocates 'objective' indicators in general, not voter turnout in particular. Thus, he excludes rights measures as insufficiently quantitative (Vanhanen, 1997, p. 34; Beetham, 1994, p. 33).
- 57 Note that he mixes measures of liberal democratic performance with economic indicators that are more properly measures of government performance.
- 58 Similarly, Executive Dominance, Unicameralism, Centralization and Constitutional Flexibility can be read as measures of performance in constraint.
- 59 Powell's three main indicators of political participation, government stability, and violence are used interchangeably as measures of political performance and democratic performance. Separate measures of civil liberties, political competition, and policy responsiveness are used exclusively for democratic performance.

60 For a detailed comparative inquiry into the institutional design, and electoral and party systems, of presidential regimes, and their implications for some aspects of democratic performance see Foweraker (1998).

61 Nonetheless, the measures will remain partial insofar as many important aspects of performance will go unmeasured. For instance, measures of accountability do not reflect the impact of clientelism or of limits to national sovereignty, just as measures of civil and minority rights ignore the rights of children and the disabled.

## References

- Arat, Z. (1991) *Democracy and Human Rights in Developing Countries*. Boulder CO and London: Lynne Rienner.
- Banks, A. S. (1997) *Cross-Polity Time-Series Data*. State University of New York at Binghamton, and previous versions in 1994 and 1979.
- Banks, A. and Textor, R. (1963) *A Cross-Polity Survey*. Cambridge MA: MIT Press.
- Barsh, R. L. (1993) 'Measuring human rights: problems of methodology and purpose', *Human Rights Quarterly*, 15, 87–121.
- Beetham, D. (1994) 'Key Principles and Indices for a Democratic Audit', in D. Beetham (ed.), *Defining and Measuring Democracy*. Sage: London.
- Berlin, I. (1969) *Four Essays on Liberty*. Oxford: Oxford University Press.
- Birch, A. H. (1964) *Representative and Responsible Government*. London: George Allen and Unwin.
- Bollen, K. (1993) 'Liberal democracy: validity and method factors in cross-national measures', *American Journal of Political Science*, 37, 1207–30.
- Bollen, K. (1991) 'Political Democracy: Conceptual and Methodological Traps', in A. Inkeles (ed.), *On Measuring Democracy: its Consequences and Concomitants*. New Brunswick: Transaction.
- Bollen, K. (1980) 'Issues in the comparative measurement of political democracy', *American Sociological Review*, 45, 370–90.
- Bond, D. Craig Jenkins, J., Taylor, C. L. and Schock, K. (1997) 'Mapping mass political conflict and civil society', *Journal of Conflict Resolution*, 41, 553–79.
- Borner, S., Brunetti, A. and Weder, B. (1995) *Political Credibility and Economic Development*. London: St Martin's Press.
- Cingranelli, D. and Richards, D. (1998) *Human Rights Dataset*, State University of New York at Binghamton.
- Cingranelli, D. and Wright, K. N. (1986) 'Measurement of cross-national variations in the extensiveness and consistency of due process', *Policy Studies Journal*, 15, 97–109.
- Coppedge, M. and Reinicke, R. (1991) 'Measuring Polyarchy', in Inkeles, A. (ed.), *On Measuring Democracy: Its Consequences and Concomitants*. New Brunswick: Transaction.
- Cutright, P. (1963) 'National political development: measurement and analysis', *American Sociological Review*, 28, 253–65.
- Dahl, R. (1996) 'The Future of Democratic Theory', Estudio/Working Paper 1996/90, Centro de Estudios Avanzados en Ciencias Sociales, Instituto Juan March de Estudios e Investigaciones.
- Dahl, R. (1989) *Democracy and its Critics*. New Haven and London: Yale University Press.
- Dahl, R. (1971), *Polyarchy: Participation and Opposition*. Chicago IL: University of Chicago Press.
- Dahl, R. (1956) *A Preface to Democratic Theory*. Chicago IL: University of Chicago Press.
- Deininger, K. and Squire, L. (1996) 'A new data set for measuring income inequality', *The World Bank Economic Review*, 10, 565–91.
- Diamond, L. (1997) 'The End of the Third Wave and the Global Future of Democracy', *Reihe Politikwissenschaft/Political Science Series No. 45*, Institute for Advanced Studies, Vienna.
- Elklit, J. (1994) 'Is the Degree of Electoral Democracy Measurable? Experiences from Bulgaria, Kenya, Latvia, Mongolia and Nepal', in D. Beetham (ed.), *Defining and Measuring Democracy*, London: Sage.
- Ersson, S. (1998) *Western European Party Systems* dataset.
- Fedderke, J. and Klitgaard, R. 'Economic Growth and Social Indicators: An Exploratory Analysis', manuscript.
- Feyerabend, P. (1993) *Against Method*. London and New York: Verso.

- Foweraker, J. (1998) 'Institutional design, party systems and governability – differentiating the presidential regimes of Latin America', *British Journal of Political Science*, 28, 651–76.
- Foweraker, J. and Landman, T. (1997) *Citizenship Rights and Social Movements: A Comparative and Statistical Analysis*. New York: Oxford University Press.
- Franklin, M. N. (1996) 'Electoral Participation', in L. LeDuc, R. G. Niemi and P. Norris (eds), *Comparing Democracies: Elections and Voting in Global Perspective*. Sage: Thousand Oaks.
- Freedom House (1997a) *Comparative Survey of Freedom*. New York: Freedom House, <http://www.freedomhouse.org>.
- Freedom House (1997b), *Survey of Press Freedom*. New York: Freedom House, <http://www.freedomhouse.org>.
- Freedom House (1996) *Survey of Economic Freedom*, New York: Freedom House, <http://www.freedomhouse.org>.
- Fukuyama, F. (1992) *The End of History and the Last Man*. London: Harmondsworth.
- Gasiorowski, M. (1996) 'An overview of the political regime change dataset', *Comparative Political Studies*, 29, 469–83.
- Gastil, R. (1991) 'The comparative survey of freedom: experiences and suggestions', in A. Inkeles (ed.), *On Measuring Democracy: Its Consequences and Concomitants*. New Brunswick: Transaction.
- Gastil, R. (1987) *Freedom in the World: Political and Civil Liberties, 1986–1987*. New York: Freedom House.
- Gleditsch, K. S. and Ward, M. D. (1997) 'Double take: a reexamination of democracy and autocracy in modern politics', *Journal of Conflict Resolution*, 41, 361–83.
- Goldstein, R. (1992) 'The Limitations of Using Quantitative Data in Studying Human Rights Abuses' in Jabine, T. and Claude, R. (eds), *Human Rights and Statistics: Getting the Record Straight*. Philadelphia PA: University of Pennsylvania Press.
- Gray, J. (1986) *Liberalism*. Milton Keynes: Open University Press.
- Gurr, T. (1993a) *Minorities at Risk: A Global View of Ethnopolitical Conflicts*. Washington DC: United States Institute of Peace Press.
- Gurr, T. (1993b) 'Why minorities rebel: a global analysis of communal mobilization since 1945', *International Political Science Review*, 14, 161–201.
- Hadenius, A. (1994) 'The Duration of Democracy: Institutional Versus Socioeconomic Factors', in D. Beetham (ed.), *Defining and Measuring Democracy*. Sage: London.
- Hadenius, A. (1992) *Democracy and Development*. Cambridge, Cambridge University Press.
- Haxton, M. and Gurr, T. (1997) *Minorities at Risk Dataset*, University of Maryland (1997), <http://www.bsos.umd.edu/cidcm/mar/cbwebpg.htm>.
- Held, D. (1996) *Models of Democracy*. Cambridge: Polity Press.
- Herman, E. and Chomsky, N. (1988) *Manufacturing Consent: The Political Economy of the Mass Media*. New York: Pantheon.
- Hill, C. (1996) *Liberty against the Law: some Seventeenth Century Controversies*. London: Allen Lane, The Penguin Press.
- Huber, J. D. and Powell, G. Bingham Jr., (1994) 'Congruence between citizens and policy makers in two visions of liberal democracy', *World Politics*, 46, 291–326.
- Humana, C. (1992) *World Human Rights Guide*. New York: Oxford University Press.
- International IDEA (1997) *Voter Turnout from 1945 to 1997: a Global Report*. Stockholm: International Institute for Democracy and Electoral Assistance.
- Inter-Parliamentary Union (1995) *Women in Parliaments 1945–1995: A World Statistical Survey*. Geneva: Inter-Parliamentary Union.
- Jackman, R. (1991) 'On the Political Capacity of Nations: Institutionalization and Legitimacy', in Inkeles, A. (ed.), *On Measuring Democracy: its Consequences and Concomitants*. New Brunswick: Transaction.
- Jackman, R. (1973) 'On the relation of economic development to democratic performance', *American Journal of Political Science*, 17, 611–21.
- Jagers, K. and Gurr, T. (1995) 'Tracking democracy's third wave with the Polity III data', *Journal of Peace Research*, 32, 469–82.
- Johnson, B., Holmes, K. R. and Kirkpatrick, M. (1998) *Index of Economic Freedom 1998*. Washington and New York: Heritage Foundation and Wall Street Journal.

- Knack, S. and Keefer, P. (1995) 'Institutions and economic performance: cross-country tests using alternative institutional measures', *Economics and Politics*, 7, 207–27.
- Kymlicka, W. (1995) *Multicultural Citizenship: A Liberal Theory of Minority Rights*. Oxford: Clarendon Press.
- Lijphart, A. (1997) 'Unequal participation: democracy's unresolved dilemma', *American Political Science Review*, 91, 1–15.
- Lijphart, A. (1994a) *Electoral Systems and Party Systems: a Study of 27 Democracies, 1945–1990*. New York: Oxford University Press.
- Lijphart, A. (1994b) 'Democracies: forms, performance and constitutional engineering', *European Journal of Political Research*, 25, 1–17.
- Lijphart, A. (1993) 'Constitutional Choices for New Democracies' in L. Diamond and Plattner, M. (eds), *The Global Resurgence of Democracy*. Baltimore MD: Johns Hopkins University Press.
- Lijphart, A. (1984) *Democracies: Patterns of Majoritarian and Consensus Government in Twenty-One Countries*. New Haven and London: Yale University Press.
- Linz, J. (1994) 'Presidential or Parliamentary Democracy: Does it Make a Difference?'. in J. Linz and A. Valenzuela (eds), *The Failure of Presidential Democracy*. Baltimore MD: Johns Hopkins University Press.
- Lipset, S. M. (1959) 'Some social requisites of democracy: economic development and political legitimacy', *American Political Science Review*, 53, 69–105.
- Locke, J. (1924) *Of Civil Government: Two Treatises*, undated with an introduction by William S. Carpenter 1924, Everyman Library No. 751. London: J. M. Dent.
- Macpherson, C. B. (1977) *The Life and Times of Liberal Democracy*. Oxford: Oxford University Press.
- Macpherson, C. B. (1962) *The Political Theory of Possessive Individualism*. London: Oxford University Press.
- Maier, C. (1992) 'Democracy since the French Revolution' in J. Dunn (ed.), *Democracy: The Unfinished Journey 508BC to AD 1993*. Oxford: Oxford University Press.
- Mauro, P. (1995) 'Corruption and growth', *Quarterly Journal of Economics*, 10, 681–712.
- McCormick, J. and Mitchell, N. J. (1997) 'Human rights violations, umbrella concepts, and empirical analysis', *World Politics*, 49, 510–25.
- Mill, J. S. (1991) *On Liberty and Other Essays*. Oxford and New York: Oxford University Press.
- Mill, J. S. (1910) *Utilitarianism, Liberty and Representative Government*. London: Dent.
- Moore, M. (1995) 'Democracy and development in cross-national perspective: a look at the statistics', *Democratization*, 2, 1–19.
- Paine, T. (1937) *The Rights of Man*. London: Watts and Co.
- Pateman, C. (1989) *The Disorder of Women: Democracy, Feminism and Political Theory*. Cambridge: Polity.
- Pateman, C. (1970) *Participation and Democratic Theory*. Cambridge: Cambridge University Press.
- Philp, M. (1997) 'Defining political corruption', *Political Studies*, 45, 436–62.
- Poe, S. and Tate, C.N. (1994) 'Repression of human rights to personal integrity in the 1980s: a global analysis', *American Political Science Review*, 8, 853–72.
- Political Risk Services (1998) *International Country Risk Guide Time Series Data 1982–1995*, IRIS Centre, University of Maryland.
- Pougerami, A. (1988) 'The political economy of development: a cross-national test of development-democracy-growth hypothesis', *Public Choice*, 58, 123–41.
- Powell, G. Bingham, Jr. (1982) *Contemporary Democracies: Participation, Stability, Violence*. Cambridge MA: Harvard University Press.
- Przeworski, A., Alvarez, M., Cheibub, J. A. and Limongi, F. (1996) 'What makes democracies endure?', *Journal of Democracy*, 7, 38–51.
- Purdue University (1998) *Purdue Political Terror Scale*, Department of Political Science, Purdue University.
- Raz, J. (1986) *The Morality of Freedom*. Oxford: Clarendon Press.
- Robertson, A. H. and Merrills, J. G. (1996) *Human Rights in the World*. Manchester: Manchester University Press.
- Sartori, G. (1997) *Comparative Constitutional Engineering: an Inquiry into Structures, Incentives and Outcomes*. Basingstoke and London: MacMillan.
- Schmitter, P. and Karl, T. L. (1993) 'What Democracy is and is not', in L. Diamond and M. Plattner (eds), *The Global Resurgence of Democracy*. Baltimore MD: Johns Hopkins University Press.

- Shugart, M. S. and Carey, J. M. (1992) *Presidents and Assemblies: Constitutional Design and Electoral Dynamics*. Cambridge: Cambridge University Press.
- Sirowy, L. and Inkeles, A. (1991) 'The effects of democracy on economic growth and inequality: a review', in A. Inkeles (ed.), *On Measuring Democracy: Its Consequences and Concomitants*. New Brunswick: Transaction.
- Skinner, Q. (1998) *Liberty before Liberalism*. Cambridge: Cambridge University Press.
- Stepan, A. and Skach, C. (1994) 'Presidentialism and Parliamentarism in Comparative Perspective', in J. Linz and A. Valenzuela (eds), *The Failure of Presidential Democracy*. Baltimore MD: Johns Hopkins University Press.
- Stohl, M., Carleton, D., Lopez, G. and Samuels, S. (1986) 'State violation of human rights: issues and problems of measurement', *Human Rights Quarterly*, 8, 592–606.
- Taylor, C. L. (1991) 'Measures of Governmental Change: Indicators of Democracy from Mass Media Reporting', in A. Inkeles (ed.), *On Measuring Democracy: its Consequences and Concomitants*. New Brunswick: Transaction.
- Taylor, C. and Hudson, M. (1972) *World Handbook of Political and Social Indicators*. New Haven CT: Yale University Press.
- Taylor, C. and Jodice, D. (1983) *World Handbook of Political and Social Indicators*. New Haven CT: Yale University Press.
- Transparency International (1997) *Corruption Perception Index*. Berlin: Transparency International, Berlin, <http://www.transparency.de/documents/press-releases/1997/1997.31.7.cpi.html>.
- Van Belle, D. (1997) 'Press freedom and the democratic peace', *Journal of Peace Research*, 34, 405–14.
- Vanhanen, T. (1997) *Prospects for Democracy: a Study of 172 Countries*. London and New York: Routledge.
- Walmsley, R. (1996) *Prison Systems in Central and Eastern Europe*. Helsinki: Heuni, European Institute for Crime Prevention and Control.
- Wooton, D. (1992) 'The Levellers', in J. Dunn (ed.), *Democracy: the Unfinished Journey 508BC to AD 1993*. Oxford: Oxford University Press.