Liberal Democracy: Validity and Method Factors in Cross-National Measures
Author(s): Kenneth Bollen
Source: *American Journal of Political Science*, Vol. 37, No. 4 (Nov., 1993), pp. 1207-1230
Published by: Midwest Political Science Association
Stable URL: http://www.jstor.org/stable/2111550
Accessed: 20-10-2015 10:33 UTC

*Liberal Democracy: Validity and Method Factors in Cross-National Measures**

Kenneth Bollen, *University of North Carolina at Chapel Hill*

This paper examines the definition and measurement of liberal democracy. Specifically, my purposes are (1) to propose a working definition of liberal democracy; (2) to outline a theory of "method factors" in subjective measures of liberal democracy; (3) to provide the first estimates of the proportion of variance due to systematic error, validity, and random error in commonly used measures; (4) to replicate these results across several years; and (5) to estimate the degree of liberal democracy in more than 150 countries. All but one measure contain systematic error, and in some cases the bias component is large. Furthermore, a new liberal democracy index has a .96 squared correlation with the liberal democracy latent variable and has negligible correlations with the method factors that are present in the individual indicators. The results suggest that the current practice of treating unadjusted democracy indicators as error free can be misleading.

Not since the early 1960s have so many countries experimented with liberal democratic political systems. Movements toward such structures are evident in Eastern Europe as well as in many parts of the developing world. This worldwide trend has renewed interest in monitoring and measuring liberal democracy. For instance, the U.S. Department of State publishes annual reports on international human rights, including democratic rights, to assist Congress on foreign assistance decisions. U.S. AID has started a Democratic Pluralism Initiative and the United Nations' Development Program has begun to investigate ways of measuring political freedoms and electoral rights (Redfern 1990; UNDP 1992).

Scholars debate whether a democratic government helps or hinders Third World industrialization, while others propose that democratic governments are incapable of reducing population growth in developing societies (e.g., Lenski and Lenski 1987; Heilbroner 1980). Some investigators

*I thank Peter Bearman, Gerhard Lenski, Philip O'Connell, Michael Lewis-Beck, the referees, and the attendees of colloquia at Indiana University and Ohio State University for their comments on this paper. Arthur Banks generously provided data that was necessary to this analysis, and he, Raymond Gastil, and Leonard Sussman provided helpful information on their ratings. Walt Davis, Glenn Deane, Hyo Joung Kim, and Scott Barber provided exceptional research assistance. I gratefully acknowledge the support of the Sociology Program of the National Science Foundation (SES-8908361, SES-9121564), the University of North Carolina's University Research Council, and the Institute for Research in Social Science.

American Journal of Political Science, Vol. 37, No. 4, November 1993, Pp. 1207–1230

see liberal democracies as less likely to initiate wars (see Merritt and Zinnes 1991). Resolution of these hypotheses presupposes that liberal democracy can be accurately measured. The available measures usually suffer from restrictions in coverage or have unknown accuracy. But for lack of alternatives, many apply these measures ignoring their limitations. Neglecting measurement errors can have serious consequences, as is illustrated in the long-standing debate on the linkages between liberal democracy and income inequality (Cutright 1967; Jackman 1974; Hewitt 1977; Bollen and Jackman 1985, 1989; Muller 1988; Simpson 1990).

This paper has several purposes: (1) to propose working definitions of liberal democracy and its two primary dimensions—political liberties and democratic rule; (2) to outline a theory of systematic error (or "method factors") in subjective measures of liberal democracy; (3) to provide the first estimates of the proportion of variance due to method factors, validity, and random error in commonly used measures; (4) to replicate these results across several years; and (5) to estimate the degree of liberal democracy in more than 150 countries.

Meanings and Measures of Democracy

The term "democracy" originates from the Greek words of *demos* ("the people") and *kratein* ("to rule"). Much of the contemporary world agrees that democracy or "people's rule" is a desirable goal. Controversy arises in defining the methods to bring about power to the people in less abstract terms. For most of the twentieth century, the dominant competition in democratic governments has been between "people's democracies" and "liberal democracies."

Liberal democracy is the focus of this research. My choice is governed by several considerations. First, liberal democracy has great salience in the contemporary world. It is not only present in many of the world's most powerful and influential nations (e.g., U.S., U.K., Japan, Germany), but it is a political system that is being tried by many countries in the former Second and Third World. Though people's democracies are still important, their international presence has declined. Second, most of the hypotheses on the causes and consequences of democracy refer to liberal democracy rather than other forms of democracy. Therefore, an examination of liberal democracy is more relevant to contemporary research. Third, by restricting oneself to liberal democracy without claiming it to be the only "true form" of democracy, we bypass some of the debate on the true meaning of democracy.

I define liberal democracy as the extent to which a political system allows political liberties and democratic rule (Bollen 1980, 1986, 1990). Political liberties exist to the extent that the people of a country have the

freedom to express a variety of political opinions in any media and the freedom to form or to participate in any political group. Democratic rule (or political rights) exists to the extent that the national government is accountable to the general population, and each individual is entitled to participate in the government directly or through representatives.

Some definitions of liberal democracy or similar concepts emphasize the democratic rule component. Schumpeter (1950, 269) states that democracy "is that institutional arrangement for arriving at political decisions in which the individuals acquire the power to decide by means of a competitive struggle for the people's vote." Similarly, Moore (1966, 414), Lipset (1981, 27), Huntington (1984, 195), and others (e.g., Vanhanen, 1990, 11) concentrate on the democratic rule dimension.

In other definitions, the traits attributed to liberal democracy fall under both dimensions. Lenski (1966, 319) identifies universal adult suffrage, an aspect of democratic rule, and the freedom of organized political opposition, a reflection of political liberties, as basic to democracy. Therborn (1977, 4) lists popular representation, universal and equal suffrage, and freedoms of speech, assembly, organization, and the press as crucial variables of democracy. Casanova (1970, 180) and Dahl (1971; see Bollen 1990, 6–7) provide two other examples of definitions that include both dimensions.

Participation and competition are additional terms that occasionally appear in definitions. Like liberal democracy, these terms have many connotations. Only some are part of the definition that I use. Restrictions on franchise or political participation are part of the democratic rule dimension. However, when it comes to measurement, the extent of the franchise, the prime barometer of inclusiveness, is less useful (see Coppedge and Reinicke 1990, 51–52) because a universal franchise is widespread, even in dictatorships or people's democracies where the elections are rigged, or little choice is offered.

Another connotation of participation is voter turnout. To the degree that voter turnout is due to restrictions on the franchise, turnout indirectly reflects democratic rule. Nevertheless, we cannot say that higher voter turnout means higher levels of liberal democracy. High voter turnout is a desired symbol of legitimation in most political systems. That is, voter turnout reflects a lot more than just the degree of liberal democracy. Voter turnout has been legally mandated in some countries (e.g., Australia, Belgium, and Italy). The former people's democracies of Eastern Europe had some of the highest voter turnout figures in the world, even though the elections allowed little choice, and political liberties were largely absent. Indeed, the correlations between voter turnout statistics and other indicators of liberal democracy are low and sometimes

even negative (see Bollen 1980, 373–74). Thus, voter turnout may have value in studying participatory democracies or in distinguishing political systems at high levels of liberal democracy (e.g., Jackman 1987), but turnout is conceptually distinct from liberal democracy.

The controversy in *measuring* liberal democracy parallels the debates about the meaning of the construct. Extensive reviews and critiques of these measures are available (e.g., May 1973; Bollen 1980, 1986; Vanhanen 1990; Inkeles 1991), so I shall not repeat those discussions here. Rather, my focus is largely on subjective indicators.

One advantage of subjective measures is that they can gauge key traits of liberal democracy that escape detection by objective measures. Freedom of expression, freedom of association, and fairness of national elections are examples. In addition, judges are capable of incorporating many factors within a country when making their assessments. For instance, repressive practices often are not objectively recorded but may be widely known. These would be missed by objective indicators, yet taken into account in expert judgments.

The main disadvantage of subjective measures is the random and systematic measurement error the rater may introduce. These errors could render the subjective measures useless, if of sufficient magnitude. On the other hand, if the measurement errors are small, researchers are less likely to be misled by subjective indicators. The paper will explore the degree of random error and systematic bias in such measures.

Prior Research on Biases

Two ways in which measures of liberal democracy are invalidated are (1) the selection of indicators that do not match the working (or theoretical) definition of liberal democracy and (2) the selection of indicators that have high face validity but contain systematic error. Examples of the first problem are measures of democracy that include indicators of political stability (e.g., Muller 1988), voter turnout (e.g., Stack 1979), and party composition of legislative bodies (e.g., Vanhanen 1990). Though often related to liberal democracy, such indicators measure constructs that are conceptually distinct from liberal democracy. The previous section mentions some of the problems. Detailed discussions are available elsewhere (May 1973; Jackman 1974; Bollen 1980, 1986, 1990; Bollen and Jackman 1985, 1989). I focus on the systematic error in subjective ratings. I refer to this type of error as "method factor," since it is due to the methods that a judge uses to form the rating.

Two scholars who have produced excellent worldwide and annual coverage of measures of political liberties and democratic rule are Arthur

Banks (1979) and Raymond Gastil (1988).[1] The wide usage of these variables is indicated by the frequent citations to their data (e.g., 35 citations to Gastil's *Freedom in the World* and eight citations to Banks's volumes in the *Social Science Citation Index Annual,* 1990). Recent empirical analyses using their data include Boswell and Dixon (1990), Starr (1991), Gonick and Rosh (1988), and Arat (1988).

The evidence of systematic error or method factors in their indicators is largely anecdotal and virtually all critiques have implicated Gastil's work. The most common claim is that his indicators are conservatively biased at least partially because of the survey's sponsorship by Freedom House. For instance, Hartman and Hsiao (1988, 797) argue that "American client states are coded as 'partly free' regardless of the amount of terror produced by their regimes. . . . Gastil's coding equates the private sector and the market with freedom." Similarly, Scoble and Wiseberg (1981, 160) and Nagle (1985, 95) suggest a conservative bias. A few critics have perceived a *liberal* bias in the ratings (Bozeman 1979, 218; Feen 1985, 240), and at least one author has suggested that the ratings were relatively unbiased (McCamant 1981, 132). None of the critics has performed a systematic analysis of the ratings across all countries. However, there are enough critics and enough anecdotal evidence to demand a more thorough search for systematic error.

Is there any quantitative empirical evidence of biases? Most studies of systematic error in cross-national data have investigated event data such as protest demonstrations, assassinations, or riots. Typically, such research finds that reports of political events are influenced by government or self-censorship, the limited resources of domestic and international media, whether an event is spectacular or not, the intensity of events, and by regional biases in reporting (Taylor and Hudson 1972, Appendix 1; Doran, Pendley, and Antunes 1973; Jackman and Boyd 1979; Taylor and Jodice 1983, vol. 2, chap. 6; Jabine and Claude 1992). The implications of this research for the subjective ratings of liberal democracy are ambiguous. One consequence is that the judges who consult only a few references are likely to read of fewer events in regions that are poorly covered by the general Western media, such as Africa. But

[1]After the analysis for this paper was completed, Polity II, a new data set by Gurr, Jaggers, and Moore (1990), was released. They focus on authority patterns in countries. Some of their indicators are similar to the Banks measures and are related to the liberal democracy construct defined here. The indicators are available for fewer countries than the Banks, Gastil, and Sussman indicators. Also, since the measures were not devised to measure liberal democracy, some of the categories would need to be rescored before treating them as indicators of democracy.

in what way will this affect their subjective ratings? Omissions include liberalizations as well as political restrictions imposed in a country. Also, there is not a one-to-one relation between the count of certain events and the degree of liberal democracy. North Korea had fewer government sanctions imposed than the United Kingdom during much of the post–World War II period, according to Taylor and Hudson (1972), but this did not make North Korea more of a liberal democracy, nor were the judges misled into rating it as having greater political liberties and democratic rule.

Fewer quantitative empirical analyses of systematic measurement error in liberal democracy indicators exist. Bollen (1980) and Bollen and Grandjean (1981) find correlated errors of measurement for democracy indicators that have a common origin in Banks (1979). Hill and Hurley (1981) find correlational evidence of a method trait for democracy indicators for Latin America. Other articles have raised concern about the reliability of indicators of liberal democracy in Africa (von der Mehden and Hill 1980; Morrison, Mitchell, and Paden 1989) or from expert-generated data more generally (Hopple and Kuhlman 1981), but these have not focused on systematic error. A final limitation of previous research is the lack of a theory of how systematic error or a "method factor" enters ratings. The next section addresses this issue.

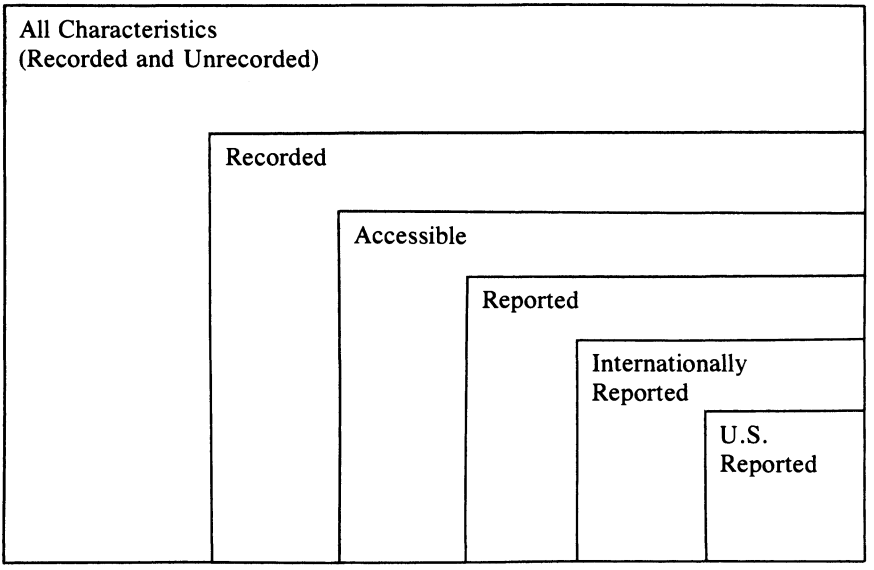
Method Factors in Ratings

The characteristics of the judges can affect subjective ratings of political liberties and democratic rule.[2] For instance, the political orientation of a judge, the relation of the rated country to the judge's home country, the interests of the agency that is funding the ratings, or other political, social, economic, and personal factors could affect a judge's ratings. Most of the time, these effects are subtle. Right-wing judges, for instance, may judge right-wing countries more favorably than socialist countries, and left-wing judges may do the opposite even when identical information is available to all. Even if judges were totally objective, a second factor could create method factors in their ratings. This is the incomplete information that is available to a judge.

Figure 1 helps illustrate the role of information restrictions in creating systematic error. In the ideal situation, we would have a nation's record on all characteristics relevant to political liberties and democratic rule. Overt and covert restrictions on the media would be known. Obstacles to freedom of assembly or the organization of political groups would

[2]This section is largely based on and summarizes a theory of bias presented in Bollen (1986).

Figure 1. Filtering of Information on Political Liberties and Democratic Rule



be detailed. Similarly, we would have total knowledge of incidences of voting fraud; prevention of voter registration or voting; the extent of the franchise, bribes, and voting corruption; and unequal weighting of votes.

The largest rectangle in Figure 1 represents this ideal condition. The second largest rectangle stands for the subset of characteristics that are recorded. Military, police, or other government agencies might have information on the denial of protest permits, the censorship of media, the attacks on opposition political groups, all of which would be relevant to assessing political liberties and democratic rule. Given the sensitivity of this information, only part, if any, of it would be accessible to the public or researchers. The accessible information makes up the third box in Figure 1. Accessible does not necessarily mean that all this information would be reported. Reported information includes that reported by domestic sources as well as information that is discovered locally but only reported beyond the country's borders. This category includes all information that is reported, even if it is only described by foreign sources. The next-to-the-smallest box in Figure 1 represents that subset of all reported information that is only reported internationally. The final, smallest box is the subset of the internationally reported information that is reported in the United States.

Each box in Figure 1 represents a subset of the information that is available in the larger box that precedes it. Determining the nature of the selection process by which information in a larger box filters to the smaller box is the key to understanding the representativeness of the information at each stage. In the ideal case, each box would contain a random sample of the information that was available in the next larger box that encloses it. Then even the smallest box that contains the U.S. available information would be a random sample of the biggest box of all characteristics, and on average the information should be representative. However, this seems unrealistic.

Those countries that are the most developed are likely to have the greatest capabilities to record events. Or a country with a highly developed state has the capacity to monitor more information. In addition, the sensitivity of the government to reports of rights violations could affect both the recording of such events and the accessibility of the records. Furthermore, the internationally and U.S.-reported information is likely to be affected by a country's size, strategic importance, distance from the reporting agencies, and relevance to powerful U.S. groups. Taken together, these factors provide the potential for bias to enter the information on which judges rely in developing their ratings.

A third cause of systematic error is the method of constructing the ratings or scales. By this I refer to such things as whether a panel or single judge is employed, the type of scaling used, the time during a year to which the ratings refer, whether the ratings are done retrospectively, or the manner in which the ratings are recorded.

In sum, three factors contribute to a method factor: (1) characteristics of the judges, (2) the information available to the judges, and (3) characteristics of the scaling process. Their combined action creates systematic error that I call a "method factor." Measures originating with the same judge are likely to reflect this method factor in varying degrees, and with multiple measures from the same judge, an estimate of the method factor becomes feasible.

This general theoretical discussion has implications for the widely used data that I analyze in this paper. Specifically, I analyze variables from Banks (1971, 1979), Gastil (1985), and Sussman (1980, 1981, 1982). Consider the first factor that can create method factors: characteristics of the judge. Anecdotal evidence suggests that Gastil's political beliefs may shade his ratings. Little similar discussion has occurred for Banks or for Sussman. However, Sussman, like Gastil, is affiliated with Freedom House, and both exchanged some information when developing their ratings. Thus, it would not be surprising to find a method factor similar to Gastil's.

The second factor, the information for the ratings, is somewhat similar across judges. The number of references that Banks (1979) checks is vast, but the majority of them originate in the United States or Western Europe. This would place Banks's information sources primarily in the two smallest boxes in Figure 1. Similarly, Gastil (1978, 8–9) obtains information from numerous sources. Some appear to be locally reported, but most are published in the United States and Europe. Sussman (1982) also uses many sources in developing his ratings, including some firsthand accounts of local situations. The techniques of constructing the scales are the same for each indicator coming from the same data source. Taken together, this leads to the prediction that method factors will be present in indicators from the same data source.

Data

Based on the earlier theoretical definition of democratic rule, I selected the following measures from Banks (1971, 1979): openness of the nominating process (segment 19, field 4), selection process for effective executive (segment 21, field 6), selection process for legislative body (segment 22, field 5), and the effectiveness of the legislative body (segment 22, field 4). I multiply the last two variables to form a new variable that measures the effectiveness of elected legislatures (see Bollen 1980, 376). This prevents "puppet" legislatures that are elected from scoring as high as elected legislatures that are effective, on the assumption that the latter type indicates greater democratic rule than the former. One measure of political liberties reflects the freedom of group opposition (segment 19, field 6).

Gastil's (1986, 1988) indicator of political rights appears to track democratic rule. His civil liberties rating measures political and other related liberties (e.g., freedom of religion, freedom of emigration) and thus is somewhat broader than the political liberties definition given above.

Sussman (1980, 1981, 1982) provides two indicators of political liberties. Specifically, he reports ratings of freedom of the print media and freedom of the broadcast media. Sussman's data are less known than Gastil's and Banks's, partly because they are more recent data, and unlike the others, they are not available in an electronic form. I include Sussman's media freedom variables for several reasons. One is their high face validity. Another reason is that Sussman, like Gastil, was affiliated with Freedom House. This suggests that a positive correlation between the method factors for Sussman and Gastil should be found. Including these measures enables a more rigorous examination of the method factors and whether they behave as predicted.

In sum, I have four measures of political liberties: freedom of broad-

cast media (X_1), freedom of print media (X_2), civil liberties (X_3), and freedom of group opposition (X_4)—and four measures of democratic rule: political rights (X_5), competitiveness of nomination process (X_6), chief executive elected (X_7), and effectiveness/elective legislative body (X_8). I linearly transformed each variable to range from zero to 10 with 10 being the highest degree of liberal democracy. This provides the same range for each variable without forcing them to have the same standard deviation and mean. All variables are available for 1979 ($N = 153$), 1980 ($N = 153$), and 1981 ($N = 154$).[3] I analyze the covariance matrix (see Jöreskog and Sörbom 1989, 47–49) of the observed variables.[4]

A path diagram of the initial measurement model is in Figure 2. The diagram conforms to the standard conventions for structural equation models with latent variables. Latent variables are enclosed in ovals or circles; observed variables are in boxes; and measurement errors (δ's) are not enclosed. The general form of the equations relating each of the eight indicators to the latent variables and random measurement errors is

$$x_i = \lambda_{ij} L_j + \lambda_{ik} L_k + \delta_i,$$

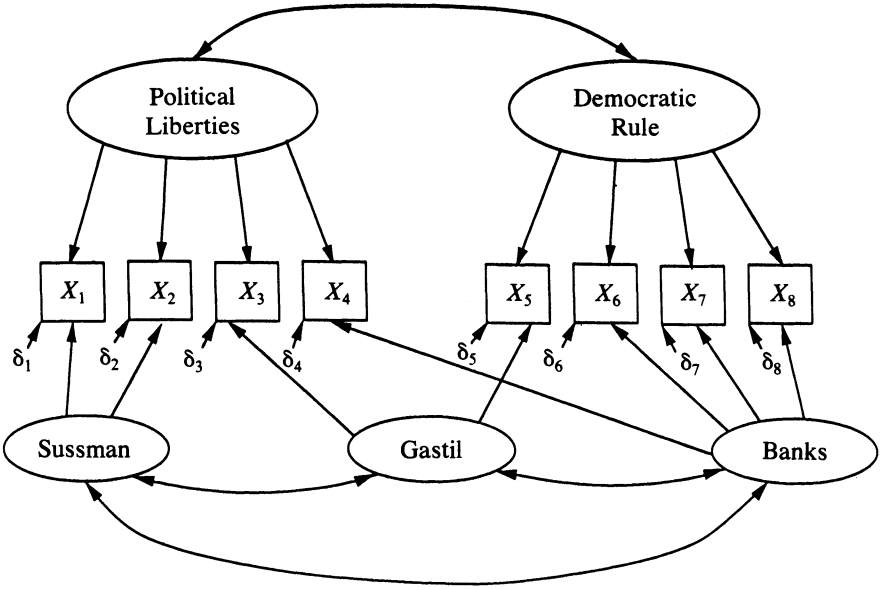
where $i = 1, 2, \ldots, 8$, $j = 1, 2$, and $k = 3, 4, 5$. The X_i represents the observed measure; L_j is the latent variable (or factor) for either political liberties or democratic rule; L_k is the latent variable for the method factor due to either Sussman, Gastil, or Banks; λ_{ij} and λ_{ik} are the parameters giving the effects of L_j and L_k on X_i; and δ_i is the random measurement error. I assume that the random error, δ_i, is uncorrelated with L_j and L_k and that $E(\delta_i)$ is zero. The "valid" component of X_i is due to L_j's influence on X_i. The systematic error or method factor results from L_k's effect on X_i and δ_i captures the random measurement error remaining in X_i.

The curved two-headed arrows in Figure 2 show variables that I hypothesize to be associated. Thus, political liberties and democratic rule are correlated as are the three method factors. The Sussman and Gastil latent variables are correlated, since both judges were affiliated with Freedom House, and their ratings were unlikely to be independent. I am

[3]The Gastil ratings refer to a country's status as of 1 January of the year. I code these ratings as referring to the immediately prior year. So, for example, I treat Gastil's 1 January 1981 ratings as ratings for 1980, since the information for the ratings come largely from 1980 and not the events in 1981.

[4]The lower-half of the 1980 covariance matrix for X_1 to X_8 is 17.65, 15.93, 19.06, 11.51, 12.47, 10.39, 14.27, 15.03, 11.60, 18.27, 12.87, 14.04, 10.95, 13.99, 13.12, 9.40, 10.40, 7.79, 11.82, 10.10, 13.91, 3.76, 4.59, 2.99, 5.49, 4.50, 7.93, 13.14, 10.59, 11.43, 9.03, 12.81, 11.16, 11.76, 5.89, and 12.82.

Figure 2. Path Diagram of Model for Eight Indicators of Political Liberties and Democratic Rule



less sure about whether Banks's method factor is associated with the other method factors, but I begin with the specification that allows this.[5]

This model cannot be estimated by traditional factor analysis procedures because it involves constraining some factor loadings and some correlations between factors to zero (see Long 1983, 11–15). Instead, I use Confirmatory Factor Analysis (CFA), a special case of the LISREL model (Jöreskog and Sörbom 1989). CFA allows estimation and testing of the hypothesized model in Figure 2. The underlying hypothesis of CFA models is $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$, where $\boldsymbol{\Sigma}$ is the population covariance matrix of the observed variables; $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is the covariance matrix that is implied by a specific model; and θ is a vector that contains the free parameters (e.g., factor loadings, factor and error variances) of a model. Once a model is specified, it implies that the covariances and variances of the observed variables will be specific functions of the model parameters. The model structure is tested by comparing the estimated implied covariance matrix

[5]Standard rules of identification (see Bollen 1989b, 238–51) did not apply to this model, so I used empirical tests of local identification. In no case did I find evidence of underidentification.

$[\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})]$ to the sample covariance matrix ($\mathbf{S}$). Support for the model comes from a relatively close match of $\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})$ to $\mathbf{S}$.

Before reporting the results, it is worthwhile to point out that in the typical multiple regression model, researchers assume that their democracy measures contain no random or systematic measurement error.[6] In other words, analyses to date treat these indicators as perfect or near-perfect measures. This assumption is justified if we find zero error variances and zero method factors when we estimate the model in Figure 2.

Results

The initial model (M_1) from Figure 2 was estimated using the 1980 data and Confirmatory Factor Analysis with the maximum likelihood fitting function (Long 1983).[7] The overall fit of the initial model is summarized in Table 1. The null hypothesis of the chi-square test is that the model structure perfectly predicts the covariances and variances of the observed variables. The nonsignificant chi-square estimate of 9.2 with 8 df ($p = .33$) provides strong support for the model. Bentler and Bonett's (1980) normed fit index, Δ_1, has a range from zero to one with one indicating perfect fit. The means of the sampling distributions of it tend to be lower in smaller samples. The index Δ_2 adjusts for this sample size effect and the degrees of freedom of the model (Bollen 1989b). A value of one indicates the best fit. Finally, Jöreskog and Sörbom's (1989) GFI and AGFI have a maximum of one, but like Δ_1 are downwardly biased in small samples.[8] All the measures indicate excellent fit.

Some individual parameter estimates were not significant or were marginally significant, so I respecified a trimmed model (M_2) that constrained the Banks method factor to be uncorrelated with the Sussman and Gastil method factors and set to zero the influence from Banks method factor to X_4 and the error variances for X_5 and X_6.[9] The estimates of overall fit of the trimmed model (M_2) are in Table 1. The fit of M_2 is excellent. Since M_2 is nested in M_1, a likelihood ratio test comparison of

[6]If random error is isolated to the dependent variable, then the OLS estimator of regression coefficients is not biased. However, systematic error in the dependent variable or random or systematic error in an independent variable can lead OLS to be a biased estimator.

[7]The ML fitting function is derived by assuming no excessive multivariate kurtosis for the distribution of the observed variables. Recent work suggests asymptotic robustness conditions that lead the ML fitting function test statistics to be valid even with excessive kurtosis. See Bollen (1989b, chap. 9) and Satorra (1990) for discussion.

[8]For further details on these measures of fit, see Bollen (1989b, 256–81).

[9]Identification of this model was established by algebraic means. Each free parameter can be written as a function of the variances and covariances of the observed variables.

Table 1. Overall Fit of Confirmatory Factor Analysis Models for 1980 (N = 153)

Model	χ^2	df	p-value	Δ_1	Δ_2	GFI	AGFI
Initial model M_1	9.2	8	.33	.99	1.00	.99	.93
Trimmed model M_2	14.2	13	.36	.99	1.00	.98	.94
$M_2 - M_1$	5.0	5	>.25	.00	.00	−.01	+.01
No method factors M_3	217	19	<.001	.86	.88	.73	.48

fit is possible by taking the difference in chi-squares for the two models. The corresponding row of Table 1 shows that the fit of the more restrictive trimmed model (M_2) is not significantly worse than the original model (M_1), thereby providing support for the more parsimonious trimmed model.

Are the method factors necessary? I estimated another model that eliminated all three method factors, but kept the two substantive variables and the random errors of measurement. This no-method factor model (M_3) had a poor fit as indicated by the chi-square estimate of 217 with 19 df ($p < .001$) and the large drop in the remaining fit indices. These results provide strong support for the presence of method factors in these measures. All the results were replicated for 1979 and 1981.[10]

The trimmed model estimates allow a partitioning of the variance of the indicators into components due to validity, method factor, and random measurement error. The variance of each indicator is

$$\text{VAR}(X_i) = \lambda_{ij}^2 \text{VAR}(L_j) + \lambda_{ik}^2 \text{VAR}(L_k) + \text{VAR}(\delta_i),$$

where VAR represents the population variance of the random variable in parentheses; λ_{ij} and λ_{ik} are the parameters of the effects of L_j and L_k on x_i; L_j is the latent variable for either political liberties or democratic rule; and L_k is the latent variable for method factor error due to either Sussman, Gastil, or Banks. Table 2 represents this breakdown as a percentage of the VAR (X_i) for all eight 1980 measures. (The results for 1979 and 1981 were essentially the same.) The variables with the highest validity are Gastil's political rights variable (X_5) and Banks's freedom of group

[10]For 1979, the fit statistics were chi-square estimate = 19.2, df = 13, p = 0.12, Δ_1 = .99, Δ_2 = 1.00, GFI = .97, and AGFI = .92. The figures were similar for 1981 except that the chi-square estimate was higher and the p-value was lower (p = .01).

Table 2. Variance in Indicators Due to Validity, Method Factor, and Random Measurement Error (1980 Data)

	Percent of Total Variance of Indicator Due to:		
Variable	Validity %	Method Factor Error %	Random Measurement Error %
X_1	68	14	18
X_2	71	22	7
X_3	78	16	6
X_4	92	0	8
X_5	93	7	0
X_6	62	38	0
X_7	14	22	64
X_8	79	9	12

Where:
X_1 = freedom of broadcast media (Sussman)
X_2 = freedom of print media (Sussman)
X_3 = civil liberties (Gastil)
X_4 = freedom of group opposition (Banks)
X_5 = political rights (Gastil)
X_6 = competitiveness of nomination process (Banks)
X_7 = chief executive elected (Banks)
X_8 = effectiveness of legislative body (Banks)

opposition variable (X_4), which have over 90% of their variance due to the substantive variables. Banks's chief executive election variable (X_7) has by far the lowest validity. One reason is that this indicator considers only whether the chief executive of a country is elected, without taking account of the fairness of the election, franchise restrictions, and other electoral characteristics thought to be important in liberal democracies.

The main hypothesis that method factors are present receives strong support with all but one variable having 7% or more of its variance accounted for by this systematic error. In the worst case, Banks's competitiveness of the nomination process (X_6) has nearly 40% of its variance due to method factor. Furthermore, all indicators but Gastil's political rights (X_5) and Banks's competitiveness of the nomination process (X_6) have detectable amounts of random measurement error.

Analysis of Method Factors

To further understand the nature of the method factors, I used the regression-based method (Lawley and Maxwell 1971) of estimating values

for these latent variables in 1980.[11] The formula is

$$\hat{L} = \hat{\Sigma}_{LL}\hat{\Lambda}'\hat{\Sigma}_{xx}^{-1}x,$$

where $\hat{L}$ is the vector of estimated latent variables, and the four symbols to the right of the equal sign are the estimated covariance matrix of the latent variables, the estimated factor loading matrix, the estimated covariance matrix of the indicators, and the vector of observed variables (deviated from their means). I then standardized these scores by dividing by their sample standard deviations. The standardized estimates show the number of standard deviations that the method factor estimates are from zero. A positive number indicates a tendency for a judge to overrate a country, and a negative value indicates underrating. In examining the individual country scores, I found that the method factors for Sussman and Gastil were similar. This is consistent with the high correlation between these method factors (about .8 in the latent variable analysis) and is not too surprising, since both judges were affiliated with Freedom House. There is some tendency for countries with conservative regimes to have positive method factors. For instance, the standardized Gastil and Sussman method factor estimates for Honduras are 3.4 and 3.8. For Panama, the estimates are 2.3 and 2.9, respectively. But there are exceptions such as South Africa that has negative standardized estimates of -1.9 and $-.8$. Banks's method factor estimates are somewhat distinct from Gastil's and Sussman's. His estimates appear to favor Marxist-Leninist or socialist countries. Romania and Yugoslavia, for example, have standardized estimates of 3.2 and 2.5.

Regional patterns also seemed present in the Gastil, Sussman, and Banks method factor estimates. Table 3 provides regression results that address this issue. The standardized factor score estimates for Gastil, Sussman, and Banks for 1979 to 1981 are regressed on eight regional dummy variables: Western Industrial, Eastern Europe, Northern Africa and Middle East, Sub-Saharan Africa, Central and South America, Caribbean, Asia, and Oceania. The regional classifications of countries are listed in the appendix. A separate dummy variable is included for each region, and the intercept is omitted from each equation. As a result, the coefficient estimates indicate the deviation in method factor from the grand mean of zero for each method factor and each year. In addition,

[11]The estimated correlations between the latent variables and the factor score estimates are 0.79 for the Gastil method factor, 0.82 for the Sussman method factor, and 0.98 for the Banks method factor. Part of the reason for the extremely high correlation for the Banks method factor is that several indicators were available from Banks, whereas only two indicators each came from Gastil and Sussman.

Table 3. Coefficient Estimates for Standardized Method Factors for Gastil, Sussman, and Banks (1979 to 1981) Regressed on Regional Dummy Variables

	Method Factor								
	Gastil			Sussman			Banks		
Region	1979	1980	1981	1979	1980	1981	1979	1980	1981
Western Industrial	0.40** (2.04)	0.35* (1.91)	0.34 (1.75)	0.18 (0.94)	0.15 (0.85)	0.23 (1.20)	−0.12 (−0.62)	−0.19 (−0.94)	−0.17 (−0.85)
Eastern Europe	−0.69** (−2.20)	−0.56* (−1.91)	−0.65** (−2.09)	−0.65** (−2.12)	−0.47 (−1.63)	−0.58* (−1.90)	1.07** (3.31)	1.14** (3.54)	0.98** (3.00)
Northern Africa & Middle East	0.16 (0.81)	0.16 (0.87)	0.13 (0.64)	0.02 (0.09)	−0.06 (−0.32)	−0.08 (−0.43)	−0.32 (−1.52)	−0.33 (−1.62)	−0.22 (−1.05)
Sub-Saharan Africa	−0.34** (−2.37)	−0.45** (−3.35)	−0.31** (−2.19)	−0.31** (−2.20)	−0.35** (−2.65)	−0.29** (−2.12)	0.00 (0.03)	0.06 (0.39)	−0.04 (−0.28)
Central & South America	0.68** (3.16)	0.87** (4.26)	0.66** (3.16)	1.02** (4.87)	1.22** (6.12)	0.96** (4.72)	−0.20 (−0.87)	−0.21 (−0.94)	−0.18 (−0.81)
Caribbean	0.15 (0.51)	−0.04 (−0.15)	−0.12 (−0.46)	0.11 (0.38)	−0.12 (−0.46)	−0.13 (−0.50)	−0.10 (−0.33)	−0.11 (−0.39)	−0.02 (−0.07)
Asia	−0.32 (−1.59)	−0.36* (−1.89)	−0.34 (−1.65)	−0.28 (−1.43)	−0.37** (−1.97)	−0.32 (−1.61)	0.19 (0.90)	0.11 (0.55)	0.17 (0.80)
Oceania	0.13 (0.27)	1.17** (2.64)	0.99** (2.12)	0.01 (0.03)	0.83* (1.91)	0.88* (1.93)	0.16 (0.32)	0.43 (0.90)	0.50 (1.02)
N	153	153	154	153	153	154	153	153	154
R^2	0.16	0.25	0.17	0.20	0.28	0.21	0.10	0.11	0.08

*$p \leq .06$ two-tailed test; **$p \leq .05$ two-tailed test.

since the method factor estimates are standardized, these coefficients indicate the regional effect in standard deviation units. For instance, on average we expect a country in Central and South America to be about .87 standard deviation above zero for the 1980 Gastil method factor score. Positive coefficients indicate a tendency for ratings that are influenced by the method factor to be too high, and negative coefficients suggest ratings that are too low.

In brief, the Gastil method factor tends to favor countries in Central and South America, Western Industrial nations, and, to a lesser extent, countries in the Oceania region. On the other hand, the Gastil method factor tends to be lower than the mean for the Sub-Saharan Africa, Eastern Europe, and Asia regions. The R^2s for the three equations range from 0.16 to 0.25. The Sussman method factor estimates exhibit a similar regional pattern, though it is less pronounced. By far the strongest effect for the Banks method factor is the positive, roughly one standard deviation, coefficients for Eastern Europe. The other regional coefficients fail to obtain statistical significance and are much smaller. The R^2s for the three equations range from .08 to .11.

The regional patterns provide indirect evidence on the origin of the method factors. Recall that earlier it was argued that three components contribute to method factors: (1) the traits of the judge, (2) the filtering of information, and (3) the methods of scale construction. Which of these can explain the regional patterns observed? Consider the third first. It is hard to imagine how the way the scales are constructed or how the time at which they are made could explain the regional patterns that occur. In addition, this could not explain the different patterns across method factors. For instance, how could the scale construction explain the Central and South America positive effects for the Gastil and Sussman method factor, but no effect for the Banks method?

The second factor, the filtering of information, is hardly a better explanation. Banks and Gastil rely on very similar information sources to inform their ratings. Many of these sources are Western and U.S. publications. Given the similar pool of information available to these raters, it is unclear how this factor could explain the regional patterns. For instance, if information availability is roughly the same for Gastil and Banks, why does the Gastil method factor have a negative association with Eastern European countries whereas the Banks method factor has a positive effect?

By process of elimination, the traits of the judge probably have the biggest effect on the regional patterns observed. I would qualify this statement in that I suspect that there could be an interaction effect between the information available and the traits of the judge. The U.S. and

Western pools of information that are available to these judges run the gamut from left to right in political orientations. The political orientation of judges could lead them to give more credence to sources that are consistent with their predispositions. So even if very similar information is accessible, not all of this information will be given the same weight in assessing the political liberties or democratic rule in a country.

An indirect test of the role of information availability on the method factor is possible for the Gastil method factor. In the annual survey of freedom in the periodical *Freedom at Issue,* Gastil indicated those countries that had a change in their ratings of civil liberties or political rights because of new information and not necessarily because of any real changes in the countries' conditions. Assuming that the traits of the judge and the method of scale construction stayed the same, the only change is in the information available. Therefore, by studying what happens to the method factor for those cases that are rerated, we can indirectly assess the impact of more information on the method factor. In several tests for this, I found little effect of new information on the Gastil method factor.[12]

Measuring Liberal Democracy

Given the presence of systematic error in the usual indicators of liberal democracy, how should researchers include liberal democracy in their analyses? Any such effort should be guided by three goals: (1) to maximize validity, (2) to minimize method factor (systematic) error, and (3) to minimize random measurement error. There are several strategies. Structural equation models that incorporate random and nonrandom measurement errors can optimize these goals. A disadvantage of this strategy occurs if the model is much more complicated than that used here. Such modeling could involve numerous free parameters, and one could have too few countries to estimate reliably the number of free parameters. Pooling cross-sectional and time-series data for nations could increase the sample size, but unfortunately such techniques are underdeveloped in latent variable models.

[12]First, I created dummy variables to indicate whether civil liberties or political rights were rerated in a given year and entered these into regressions with the method factor estimates for the same year and the method factor estimates for the prior year as dependent variables. Second, I designed panel regressions with the method factor estimate as the dependent variable and with the prior year method factor estimate and the preceding dummy variables as the explanatory variables. Finally, I ran other panel regressions with the civil liberties or political rights ratings as the dependent variable and the year prior ratings and the dummy variable as the explanatory variables. To conserve space, I do not show these results, but I found no consistent significant effects.

A second strategy is to choose a single indicator to measure liberal democracy. Unfortunately, none of the indicators is satisfactory. Gastil's political rights indicator has the highest validity coefficient estimate of 0.93 and negligible random measurement error (see Table 2). Yet 7% of the variance in the political rights variable is due to method factor bias. Is this 7% ignorable? There are reasons to be cautious. The method factor explains 7% of the variance in the Gastil measure, and the analysis of the method factor showed that it is significantly associated with world region and other related factors. If we treat Gastil's political rights variable as a perfect measure of liberal democracy, we are omitting the method factor that affects it, and it is likely to be associated with other variables in the analysis. Conditions are even worse for most of the other measures that have either lower validity or higher systematic or random error.

Furthermore, democracy measures that have systematic error can affect theory testing. In Bollen (1980, 382–84), democracy indices that had systematic error related to political stability led to empirical support for the hypothesis that the timing of a country's development was related to political democracy, while democracy indices that excluded stability rejected the hypothesis. This occurred despite moderately high correlations between indices. Thus, there are drawbacks to the single indicator measurement of liberal democracy even if one chooses the indicator with the optimal validity.

Factor score estimates of the latent liberal democracy variables are a third alternative. Using the previous factor score estimation formula, I estimated the political liberties and democratic rule dimensions of liberal democracy. The squared correlation between the mean of the factor score estimates and the mean of the true latent variables for political liberties and democratic rule is 0.967, and it has a squared correlation of 0.01 or less with the method factors. Thus, the factor score estimate appears superior to choosing a single indicator. One disadvantage is that factor score weights will change from year to year. Another drawback is its complexity. I also formed the equally weighted sum of all eight indicators, but this had a much lower squared correlation (0.905) with the latent political liberties and democratic rule variables and higher squared correlations (about 0.04) with the method factors. However, with some experimentation, I found that the equally weighted sum of Banks's political opposition variable (X_4), Gastil's political rights (X_5) measure, and Banks's legislative effectiveness (X_8) had a squared correlation with the mean of the political liberties and democratic rule latent variables of 0.956 and squared correlations of less than 0.01 with the method factor variables. Thus, if researchers cannot construct a full structural equation

model, then the three-indicator index is a reasonable alternative that maximizes validity and minimizes systematic and random error. The appendix reports this liberal democracy index for the 1980 data.

Conclusions

This analysis provides the first quantitative evidence that method factors are present in widely used ratings of liberal democracy. The findings raise concerns about the overwhelmingly dominant practice of treating these measures as error free. With the exception of one variable, the method factors explained between 7% and 38% of the variances in the indicators. Adding the percentages of the variances due to random error leads to a range of 7% to 86% of the variances in the indicators due to invalidity.

I hypothesized that method factors arose from traits of the judge, information filtering, and methods of scale construction. The indirect tests I performed indicated that the traits of the judge, perhaps in conjunction with the information that the judge gives the most credence to, can best explain the regional patterns of the bias that were documented. It should be possible to devise better tests to help understand how these method factors develop and perhaps how they can be minimized or eliminated.

My final point is to stress that I am not questioning the efforts to measure democratic political structures or questioning the motives of those developing measures. On the contrary, researchers such as Banks, Gastil, and Sussman have performed a great service for the social sciences by developing time-series measures of democracy. Until very recently, most national and international agencies shied away from such efforts because of the politically sensitive nature of the topic. My goal is to improve the validity of human rights measures and to encourage the collection and formation of more and better data. Hopefully, the evidence provided here can assist in these efforts.

Manuscript submitted 10 August 1991
Final manuscript received 8 February 1993

APPENDIX
Regional Classifications and 1980 Liberal Democracy Estimates $[(X_4 + X_5 + X_8)/.3]$

Below are the regional classification of countries and the equally weighted sum of three indicators to measure liberal democracy. The index is scored to range from zero to

100 and is rounded to the nearest integer. Countries at the maximum or minimum of the scale still fall considerably short of an absolute liberal democracy or an absolute dictatorship. Small differences in scores can be due to measurement error and should not be overinterpreted.

Western Industrial:
Australia (100), Austria (100), Belgium (100), Canada (100), Denmark (100), Finland (94), France (100), Germany (West) (89), Greece (94), Iceland (100), Ireland (100), Italy (100), Luxembourg (100), Malta (94), Netherlands (100), New Zealand (100), Norway (100), Portugal (94), Spain (83), Sweden (100), Switzerland (100), United Kingdom (100), United States (100).

Eastern Europe:
Albania (11), Bulgaria (11), Czechoslovakia (11), Germany (East) (11), Hungary (17), Poland (17), Romania (11), Soviet Union (17), Yugoslavia (28).

Northern Africa and Middle East:
Algeria (17), Bahrain (11), Cyprus (56), Egypt (33), Iran (33), Iraq (17), Israel (94), Jordan (6), Kuwait (6), Lebanon (50), Libya (17), Morocco (61), Oman (6), Qatar (17), Saudi Arabia (6), Sudan (22), Syria (33), Tunisia (28), Turkey (11), United Arab Emirates (17), Yemen (North) (11), Yemen (South) (17).

Sub-Saharan Africa:
Angola (11), Benin (11), Botswana (83), Burkina Faso (6), Burundi (0), Cameroon (17), Cape Verde (17), Central African Republic (0), Chad (6), Congo (11), Equatorial Guinea (0), Ethiopia (0), Gabon (28), Gambia (72), Ghana (83), Guinea (11), Guinea-Bissau (6), Ivory Coast (17), Kenya (33), Lesotho (39), Liberia (6), Madagascar (28), Malawi (17), Mali (0), Mauritania (0), Mauritius (83), Mozambique (11), Niger (0), Nigeria (83), Sao Tome (17), Senegal (39), Seychelles (17), Sierra Leone (33), Somalia (11), South Africa (56), Swaziland (22), Tanzania (28), Togo (11), Transkei (33), Uganda (44), Zaire (17), Zambia (22), Zimbabwe (67).

Central and South America:
Argentina (6), Bolivia (0), Belize,[a] Brazil (39), Chile (6), Colombia (72), Costa Rica (100), Ecuador (83), El Salvador (6), Guatemala (44), Guyana (72), Honduras (17), Mexico (56), Nicaragua (17), Panama (28), Paraguay (33), Peru (72), Suriname (0), Uruguay (17), Venezuela (89).

Caribbean:
Antigua,[a] Bahamas (100), Barbados (100), Cuba (17), Dominica (94), Dominican Republic (72), Grenada (11), Haiti (17), Jamaica (94), St. Lucia (94), St. Vincent (94), Trinidad (83).

Asia:
Afghanistan (0), Bangladesh (44), Bhutan (22), Burma (11), China (17), India (94), Indonesia (33), Japan (100), Kampuchea (0), Korea (North) (11), Korea (South) (33), Laos (6), Malaysia (67), Maldives (22), Mongolia (11), Nepal (44), Pakistan (0), Philippines (33), Singapore (44), Sri Lanka (83), Thailand (67), Vietnam (11).

Oceania:
Fiji (94), Nauru (61), Papua New Guinea (83), Tonga (22).

Note: [a] Data not available in 1980.

REFERENCES

Arat, Zehra F. 1988. "Democracy and Economic Development: Modernization Theory Revisited." *Comparative Politics* 21:21–36.

Banks, Arthur S. 1971. *Cross-Polity Time-Series Data*. Cambridge: MIT Press.

———. 1979. *Cross-National Time-Series Data Archive User's Manual*. Binghamton: State University of New York at Binghamton.

Bentler, Peter H., and Douglas G. Bonett. 1980. "Significance Tests and Goodness-of-Fit in the Analysis of Covariance Structures." *Psychological Bulletin* 88:588–600.

Bollen, Kenneth A. 1980. "Issues in the Comparative Measurement of Political Democracy." *American Sociological Review* 45:370–90.

———. 1986. "Political Rights and Political Liberties in Nations: An Evaluation of Human Rights Measures, 1950 to 1984." *Human Rights Quarterly* 8:567–91.

———. 1989a. "A New Incremental Fit Index for General Structural Equation Models." *Sociological Methods and Research* 17:303–16.

———. 1989b. *Structural Equations with Latent Variables*. New York: Wiley.

———. 1990. "Political Democracy: Conceptual and Measurement Traps." *Studies in Comparative International Development* 25:7–24.

Bollen, Kenneth A., and Burke Grandjean. 1981. "The Dimensions of Democracy: Further Issues in the Measurement and Effects of Political Democracy." *American Sociological Review* 46:651–59.

Bollen, Kenneth A., and Robert W. Jackman. 1985. "Political Democracy and the Size Distribution of Income." *American Sociological Review* 50:438–57.

———. 1989. "Democracy, Stability, and Dichotomies." *American Sociological Review* 54:612–21.

Boswell, Terry, and William J. Dixon. 1990. "Dependency and Rebellion: A Cross-National Analysis." *American Sociological Review* 55:540–59.

Bozemann, Adda. 1979. "Review of *Freedom in the World: Political Rights and Civil Liberties, 1978*." *Orbis* 23:213–20.

Casanova, Pablo González. 1970. *Democracy in Mexico*. New York: Oxford University Press.

Coppedge, Michael, and Wolfgang H. Reinicke. 1990. "Measuring Polyarchy." *Studies in Comparative International Development* 25:51–72.

Cutright, Phillips. 1967. "Inequality: A Cross-National Analysis." *American Sociological Review* 32:562–78.

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

Doran, Charles F., Robert E. Pendley, and George E. Antunes. 1973. "A Test of Cross-National Event Reliability." *International Studies Quarterly* 17:175–203.

Feen, Richard H. 1985. "Review of *Freedom in the World: Political and Civil Liberties, 1984–1985* by Raymond Gastil and *Country Reports on Human Rights Practices for 1984* by U.S. Department of State." *Washington Quarterly* 8:237–41.

Gastil, Raymond D. 1978. *Freedom in the World: Political Rights and Civil Liberties, 1978*. New York: Freedom House.

———. 1985. *Freedom in the World: Political Rights and Civil Liberties, 1984–1985*. Westport, CT: Greenwood Press.

———. 1986. *Freedom in the World: Political Rights and Civil Liberties, 1985–1986*. Westport, CT: Greenwood Press.

———. 1988. *Freedom in the World: Political Rights and Civil Liberties, 1987–1988*. New York: Freedom House.

Gonick, Lev S., and Robert M. Rosh. 1988. "The Structural Constraints of the World-Economy on National Political Development." *Comparative Political Studies* 21: 171–99.

Gurr, Ted Robert, Keith Jaggers, and Will H. Moore. 1990. "The Transformation of the Western State: The Growth of Democracy, Autocracy, and State Power since 1800." *Studies in Comparative International Development* 25:73–108.

Hartman, John, and Wey Hsiao. 1988. "Inequality and Violence: Issues of Theory and Measurement in Muller." *American Sociological Review* 53:794–800.

Heilbroner, Robert L. 1980. *An Inquiry into the Human Prospect*. New York: Norton.

Hewitt, Christopher. 1977. "The Effect of Political Democracy and Social Democracy on Equality in Industrial Societies: A Cross-National Comparison." *American Sociological Review* 42:450–64.

Hill, Kim Quaile, and Patricia A. Hurley. 1981. "Convergent and Discriminant Validity Tests for Fitzgibbon-Johnson Political Scales." *Quality and Quantity* 15:433–43.

Hopple, Gerald W., and James A. Kuhlman. 1981. *Expert-Generated Data: Applications in International Affairs*. Boulder, CO: Westview Press.

Huntington, Samuel P. 1984. "Will More Countries Become Democratic?" *Political Science Quarterly* 99:193–218.

Inkeles, Alex. 1991. *On Measuring Democracy*. New Brunswick, NJ: Transaction Publishers.

Institute for Scientific Information. 1991. *Social Sciences Citation Index, 1990 Annual*. Vol. 2. Philadelphia: ISII.

Jabine, Thomas B., and Richard P. Claude. 1992. *Human Rights and Statistics*. Philadelphia: University of Pennsylvania Press.

Jackman, Robert W. 1974. "Political Democracy and Social Equality: A Comparative Analysis." *American Sociological Review* 39:29–45.

———. 1987. "Political Institutions and Voter Turnout in the Industrial Democracies." *American Political Science Review* 81:405–23.

Jackman, Robert W., and William A. Boyd. 1979. "Multiple Sources in the Collection of Data on Political Conflict." *American Journal of Political Science* 23:434–58.

Jöreskog, Karl G., and Dag Sörbom. 1989. *LISREL 7: A Guide to the Program and Applications*. Chicago: Statistical Package for the Social Sciences.

Lawley, D. N., and A. E. Maxwell. 1971. *Factor Analysis as a Statistical Method*. New York: American Elsevier.

Lenski, Gerhard. 1966. *Power and Privilege*. New York: McGraw-Hill.

Lenski, Gerhard, and Jean Lenski. 1987. *Human Societies*. New York: McGraw-Hill.

Lipset, Seymour Martin. 1981. *Political Man*. Baltimore: Johns Hopkins University Press.

London, Bruce, and Thomas D. Robinson. 1989. "The Effect of International Dependence on Income Inequality and Political Violence." *American Sociological Review* 54: 305–08.

Long, J. Scott. 1983. *Confirmatory Factor Analysis*. Beverly Hills: Sage.

McCamant, John F. 1981. "A Critique of Present Measures of 'Human Rights Development' and an Alternative." In *Global Human Rights: Public Policies, Comparative Measures, and NGO Strategies*, ed. Ved P. Nanda, James R. Scarritt, and George W. Shepherd, Jr. Boulder, CO: Westview Press.

May, John D. 1973. "Of the Conditions and Measures of Democracy." Morristown, NJ: General Learning.

Merritt, Richard L., and Dina A. Zinnes. 1991. "Democracies and Wars." In *On Measuring Democracy*, ed. Alex Inkeles. New Brunswick, NJ: Transaction Press.

Moore, Barrington, Jr. 1966. *Social Origins of Dictatorship and Democracy*. Boston: Beacon.

Morrison, Donald George, Robert Cameron Mitchell, and John Naber Paden. 1989. *Black Africa: A Comparative Handbook*. 2d ed. New York: Paragon House and Irvington Publishers.

Muller, Edward N. 1988. "Democracy, Economic Development, and Income Inequality." *American Sociological Review* 53:50–68.

Nagle, John D. 1985. *Introduction to Comparative Politics: Political Systems and Performance in Three Worlds*. Chicago: Nelson-Hall.

Redfern, Paul. 1990. "Methodology Employed in Construction of PFI." Unpublished manuscript for United Nations Development Program.

Satorra, Albert. 1990. "Robustness Issues in Structural Equation Modeling: A Review of Recent Developments." *Quality and Quantity* 24:367–86.

Schumpeter, Joseph A. 1950. *Capitalism, Socialism, and Democracy*. New York: Harper and Row.

Scoble, Harry M., and Laurie S. Wiseberg. 1981. "Problems of Comparative Research on Human Rights." In *Global Human Rights: Public Policies, Comparative Measures, and NGO Strategies*, ed. Ved P. Nanda, James R. Scarritt, and George W. Shepherd, Jr. Boulder, CO: Westview Press.

Simpson, Miles. 1990. "Political Rights and Income Inequality: A Cross-National Test." *American Sociological Review* 55:682–93.

Stack, Steven. 1979. "The Effects of Political Participation and Socialist Party Strength on the Degree of Income Inequality." *American Sociological Review* 44:168–71.

Starr, Harvey. 1991. "Democratic Dominoes: Diffusion Approaches to the Spread of Democracy in the International System." *Journal of Conflict Resolution* 35:356–81.

Sussman, Leonard R. 1980. "Freedom of the Press: Problems in Restructuring the Flow of International News." In *Freedom in the World, 1980*, ed. Raymond D. Gastil. New York: Freedom House.

———. 1981. "Freedom of the Press: A Personal Account of the Continuing Struggle." In *Freedom in the World, 1981*, ed. Raymond D. Gastil. Westport, CT: Greenwood Press.

———. 1982. "The Continuing Struggle for Freedom of Information." In *Freedom in the World, 1982*, ed. Raymond D. Gastil. Westport, CT: Greenwood Press.

Taylor, Charles L., and Michael C. Hudson. 1972. *World Handbook of Political and Social Indicators*. 2d ed. New Haven: Yale University Press.

Taylor, Charles L., and David Jodice. 1983. *World Handbook of Political and Social Indicators, III*. New Haven, CT: Yale University Press.

Therborn, Göran. 1977. "The Rule of Capital and the Rise of Democracy." *New Left Review* 103:3–41.

United Nations Development Program. 1992. *Human Development Report*. Oxford: Oxford University Press.

Vanhanen, Tatu. 1990. *The Process of Democratization*. New York: Crane Russak.

von der Mehden, Fred R., and Kim Quaile Hill. 1980. "Area Experts' Images of African Nations." *Comparative Political Studies* 12:497–510.