

University of Georgia

From the Selected Works of Cas Mudde

June, 2010

Introduction: Rational Data Choice

Cas Mudde, *DePauw University*
Andreas Schedler



Available at: https://works.bepress.com/cas_mudde/37/

Political Research Quarterly

<http://prq.sagepub.com/>

Introduction: Rational Data Choice

Cas Mudde and Andreas Schedler

Political Research Quarterly 2010 63: 410

DOI: 10.1177/1065912910371286

The online version of this article can be found at:

<http://prq.sagepub.com/content/63/2/410>

Published by:



<http://www.sagepublications.com>

On behalf of:

The University of Utah



Western Political Science Association



Additional services and information for *Political Research Quarterly* can be found at:

Email Alerts: <http://prq.sagepub.com/cgi/alerts>

Subscriptions: <http://prq.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://prq.sagepub.com/content/63/2/410.refs.html>

Introduction: Rational Data Choice

Cas Mudde

DePauw University & University of Antwerp, Antwerpen, Belgium

Andreas Schedler

Centro de Investigación y Docencia Económicas (CIDE), Mexico City, Mexico

We live in an era of unprecedented abundance of cross-national political data. The study of comparative politics disposes of more quantitative information than anytime before, and year by year the number of cross-national datasets keeps expanding. Even if we still lack data for innumerable research questions, many areas of inquiry host not just one, but several competing datasets. We thus face the luxury, and necessity, of choice—a choice, however, for which we are not well equipped as a discipline.

Since data production in comparative politics essentially proceeds in a private, decentralized, and unregulated fashion (see Schedler 2012), users of cross-national political data face structural problems of information about data supply and data quality. To make rational, that is, informed and justified choices among competing datasets, they must overcome these informational uncertainties. Which is easier said than done. Over the past years, practitioners of comparative politics have been paying increasing attention to issues of cross-national measurement.¹ Yet the “uncritical use of problematic data sets” (Herrera and Kapur 2007, 372) continues to be widespread, and so is the uncritical choice among competing datasets. Irrational, that is, blind and nonjustified data choices are disciplinary failures as much as they are individual ones. Comparative political science has not yet developed the requisite infrastructure, norms, and practices that would allow scholars to choose among available data in reflexive and reasoned ways, without previously having to make huge investments in the acquisition of basic consumer information.

Selecting data consciously, rather than faithfully, involves four seemingly simple tasks whose realization still imposes high information costs on data users in many fields of comparative research: (1) surveying the supply of relevant data, (2) assessing the quality of available data, (3) estimating the inferential implications of alternative datasets, and (4) choosing appropriate data for empirical research.

Task 1: Surveying Data Supply

Buyers in consumer markets do not need perfect market information. They can rely on brands, habits, or emotional appeals. Critical data users must not. Before choosing their quantitative inputs, they need to inform themselves broadly about the existing supply of data. What is out there? Which datasets exist in a given field of research? What do they pretend to measure? Do they offer original or composite data? Factual measures, subjective data, or expert judgments? Which are their sources? Which are the units of analysis? How many countries and which period do they cover? Who creates the data and who funds their creation? Are the data and their documentation accessible on the Internet? In which format? Is access free?

Given the sheer number of potentially relevant datasets as well as their constant growth, finding orientation within the complex, evolving landscapes of available cross-national data represents a major challenge in many fields of research. To give just two examples: In an extensive thematic data review, Todd Landman and Julia Häusermann (2003, i) identified more than 170 “seminal efforts to measure democracy, human rights, and good governance.” Similarly, in an overview over cross-national data on armed conflicts, Kristine Eck (2005, 3) inventoried almost 60 datasets, even while limiting herself to only “the

Editors' Note: First versions of the articles included in this mini-symposium were presented at the workshop “The Numbers We Use, the World We See: Evaluating Cross-National Datasets in Comparative Politics,” European Consortium of Political Research (ECPR), Joint Sessions of Workshops, Rennes, France, April 11-16, 2008. We thank the participants for their invaluable inputs to the project. In addition, we are most grateful to *PRQ* editors Amy Mazur and Cornell Clayton for their constructive criticism and critical encouragement. The Author corresponds to Cas Mudde, Department of Political Science, University of Antwerp, Sint-Jacobstraat 2 2000 Antwerpen, Belgium; Email: cas.mudde@ua.ac.be

most prominent” ones. Swimming in apparent data wealth, we run the risk of drowning in numbers.

In response to the daunting costs of consumer information in the disorderly markets of cross-national political data, some public institutions have begun to develop or commission data inventories that present systematic information about a broad range of cross-national data. Two fine examples are the MacroDataGuide offered by the Norwegian Social Science Data Services (NSD), which provides structured information about dozens of cross-national political datasets (www.nsd.uib.no/macrodataguide); and the Global Programme on Capacity Development for Democratic Governance Assessments and Measurements at the Oslo Governance Center of the United Nations Development Programme (UNDP), which has published comprehensive data users’ guides on democratic governance, local governance, bureaucratic performance, corruption, and state fragility (www.undp.org/oslocentre/). These public data inventories go well beyond conventional data portals, search engines, or listings of data sources. In addition to providing structured descriptive information about datasets, they also evaluate certain aspects of data quality.²

Task 2: Evaluating Data Quality

Once prospective data users have shortlisted the data candidates that seem relevant for their research purpose, they face the challenge of evaluating their quality. Quality assessments may either emphasize procedural criteria that examine the process of data generation (transparency, reliability, and replicability) or substantive criteria that evaluate the final outcome (validity, accuracy, and precision). Procedurally sound data generation basically requires that each step from the choice of factual information to the assignment of numbers is well documented, nonidiosyncratic, and repeatable. Substantively sound data production basically requires that the numbers we cook up match the political realities they pretend to map.

Over the past years, we have witnessed the emergence of vigorous and sophisticated debates on the quality of cross-national data in various fields of research, such as the comparative study of political regimes and the quality of governance.³ In addition, more occasional and dispersed discussions on data quality have arisen in numerous fields of comparative inquiry, like ethnic politics (Chandra and Wilkinson 2008), taxation (Lieberman 2002), terrorism

(Krueger and Laitin 2004), satisfaction with democracy (Canache, Mondak, and Seligson 2001), and democratic support (Schedler and Sarsfield 2007). Nevertheless, we still do not see evaluations of data quality appear in academic journals on a regular basis, and uncertainty over the quality of the cross-national data we use continues to be pervasive in many areas of research. The academic rewards for systematic, in-depth assessments of data quality are scant (see Herrera and Kapur 2007) and the required efforts substantial.

In principle, procedures of data generation are relatively easy to evaluate. It suffices to look at the codebook and documentation of a dataset to establish the extent to which data users are able to follow the whole chain of steps from information sources to coding decisions for each data point. Procedural evaluations turn more complex, though, if we look beyond what we may call the bureaucratic quality of data development and try to assess the quality of the primary information that goes into data construction. How accurate, how complete, how balanced, how unambiguous, how homogeneous across cases, and how proximate to historical events are the evidentiary bases that sustain the construction of cross-national data?⁴

If we dispose of cross-national datasets that pretend to measure similar phenomena yet rely on different types of sources, we can reach statistical estimates of source biases. For instance, in their review of cross-national measures of liberal democracy, Kenneth Bollen and Pamela Paxton (2000) estimate the biases introduced by the uneven country coverage of U.S.-based news sources. In their discussion of cross-national measures of corruption, Angela Hawken and Gerardo Munck (2009) document the systematic differences introduced by different categories of actors who are asked for integrity assessments. However, as Kirk Bowman, Fabrice Lehoucq, and James Mahoney (2005) show in their review of cross-national datasets on political regimes, a full appreciation of the quality of information sources that feed comparative data requires close contextual knowledge. After estimating the magnitude of what they call “data-induced measurement error” for Central American political regimes since independence, the authors go through the painstaking process of mobilizing full available historical knowledge in order to correct judgmental errors that arise from “the use of inaccurate, partial, or misleading secondary sources” (p. 940).⁵

If evaluations of procedural quality sometimes turn into demanding exercises, substantive evaluations of data quality almost invariably do, if taken seriously.

Even the most elemental demand of measurement validity is often difficult to establish in practice. While we all recognize “the need for conceptual clarity” (Landman and Häusermann 2003, 35), numerous substantive areas of cross-national data generation suffer from considerable conceptual and operational opacity. To meet the “central challenge” of developing “the best fit between the indicator(s) and the particular concept under investigation” (Lieberman 2002, 95), scholars often have to clear their way through “a startling diversity of definitions and measures” (Schneider 2003, 32). To complicate matters, where datasets contain aggregate indicators, it may be impossible to evaluate the validity and reliability of aggregation procedures if access to primary data is constrained (see Hawken and Munck 2009). Even single datasets may involve considerable conceptual uncertainties, as demonstrated by the in-depth analysis of Polity data conducted by Kristian Gleditsch and Michael Ward (1997). Furthermore, conceptual clarity does not ensure conceptual validity. Entire fields of quantitative research, such as the comparative study of ethnic politics, have evolved on the basis of cross-national measures that fail to correspond to the theoretical concepts that animates this research (see Laitin and Posner 2001; Chandra and Wilkinson 2008).

Task 3: Choosing Data

After evaluating the quality of available data, scholars may reach one of three broad conclusions: (a) they may single out one measure that seems superior to all others; (b) they may conclude that no measure is adequate for their research purpose; or (c) they may identify a bundle of measures that seem roughly equivalent in qualitative terms.

(a) If one high-quality dataset clearly outperforms its competitors, we are in the comfortable position of following standard advice and choosing the measure that is procedurally sound and substantively “most appropriate to our theoretical purpose” (King, Keohane, and Verba 1994, 153).

(b) If all measures under examination seem to be either procedurally flawed or substantively inadequate, we face a more complicated set of alternatives. In the absence of valid observations, we may renounce the pretension of large-*N* research, amend existing measures, or develop new data. Alternatively, we may follow the widespread maxim according to which observational uncertainty is preferable to blindness (“bad data are better than no data”) and choose

second-worst measures to conduct our statistical research. If we use data of low or uncertain quality, though, we have to accept the inevitable cost that we will be generating inferences of low or uncertain quality.

(c) More often than not, assessments of data quality do not yield unambiguous recommendations. They neither permit to single out best measures, nor to simply discard weak measures, nor to arbitrate among similar measures (see also Hawken and Munck 2009, 22). It is in such situations of methodological and conceptual indeterminacy that scholars confront the challenge of data selection. Balancing pragmatic, procedural, and substantive reasons, they may either select single measures from the menu of data options (data choice); they may form composite indicators by aggregating existing measures (data aggregation); or they may work with multiple measures (data comparison). Whatever path they choose, however, they need to assess the empirical implications of their choice.

Task 4: Assessing Empirical Implications

The numbers we choose are likely to affect the world we see; they are likely to influence our descriptive as well as our causal inferences (see also Munck and Snyder 2004). If multiple data are available for a given research purpose, and if quality evaluations cannot establish the unambiguous superiority of one set of data over the others, researchers are required to assess the substantive implications of their strategies of data usage. It is standard practice in quantitative comparative politics to test the robustness of empirical results to variations in statistical model specifications. As a matter of fact, the causal estimations provided by cross-national research are often highly sensitive both to the selection of explanatory variables and the choice of statistical techniques of data analysis—which is frustrating to the goal of accumulative generation of knowledge (see Kittel 2006; Rivera 2010; Wilson and Butler 2007). By contrast, systematic efforts to examine the sensitivity of our empirical claims to the selection of cross-national data are still rare in the comparative study of politics.

Systematic efforts to assess the *sensitivity of descriptive inferences* to data selection have been presented by Bowman, Lehoucq, and Mahoney (2005) and Bogaards (2007) with respect to measures of democracy; Hawken and Munck (2009) with respect to measures of corruption; Gandhi (2008) with respect

to election data; Benoit and Laver (2007) with respect to data on programmatic positions of political parties; Facello (2008) with respect to political satisfaction; and Montero et al. (2008) with respect to religiosity. Systematic efforts to assess the *sensitivity of causal inferences* to data selection have been presented by Ríos-Figueroa and Staton (2008), with respect to data on judicial independence; Burkhart (2010), with respect to income inequality; and Elkins (2000), Casper and Tufis (2003), and Cheibub, Gandhi, and Vreeland (forthcoming), with respect to measures of democracy.

These are laudable examples of meta-analyses of data that allow us to ground our scientific inferences in honest “estimates of uncertainty” (King, Keohane, and Verba 1994, 32). They are exceptional, rather than common practice, though. In general, comparative “researchers still do little to check the robustness of their results using more than one data source” (Hawken and Munck 2009, 3). Even if we know in the abstract that the numbers we use shape the world we see, we often lack the concrete knowledge of how the specific measures we select affect the empirical inferences we draw.

Measurement Matters

The present mini-symposium assembles a powerful group of comparative studies that take the methodological imperative to systematically examine the inferential implications of data selection seriously. Situated in different substantive areas of political research, these articles are by necessity conceptual, methodological, and empirical hybrids. To show that measurement matters, they take the required steps of “rational data choice” outlined above, including the careful conceptualization of their field of research, the critical assessment of available data, and the conduct of replication studies with alternative measures. Although they differ in the specific strategies of meta-analysis they pursue, they all offer more general methodological models of how to approach the complexities of data selection.

Employing an original dataset of articles in comparative politics, the opening article by Andreas Schedler and Cas Mudde provides an empirical assessment of both the structure of contemporary comparative research and the patterns of data usage in quantitative comparative politics. It delineates the disciplinary context that defines the reach and relevance of methodological discussions on the selection of cross-national

political data. For all their importance, these debates do not decide the fate of the entire discipline. First, although the quantitative study of comparative politics has been on the rise over the past years, the publication frequency of qualitative methods has not decreased in absolute terms. Second, the prototypical instance of quantitative comparative research is not the large-*N* cross-national study, but the statistical analysis of politics within a single country. Third, even if the “uncritical use of problematic data sets” (Herrera and Kapur 2007, 372) appears to be widespread in the discipline as a whole, comparative work of excellence goes beyond “the ongoing recycling of low-quality data and the failure to produce new data” (Herrera and Kapur 2007, 381). In articles published in the top journals, authors frequently respond to the low quality or the lack of data by either amending existing datasets or developing original data. Thus, although severely *constrained* by the availability of cross-national data, the quantitative study of comparative politics does not seem to be blindly *driven* by available data.

The piece by Jan Teorell and Catharina Lindstedt opens the series of articles that assess the empirical consequences of data choices in different fields of comparative research. Their contribution evaluates existing cross-national time-series datasets on electoral systems. After laying out the conceptual maps that guide the comparative study of electoral systems, the authors identify four relevant data sources and evaluate their substantive content, empirical coverage, conceptual validity, and procedural quality. Since these datasets are based on different research purposes and conceptual choices, they contain “surprisingly little information” that is “strictly comparable.” In the analysis of cross-national data, such problems of comparability are common. The authors are nevertheless able to identify subsets of data that pretend to measure similar properties of electoral systems. Within these areas of substantive overlap, they examine the descriptive convergence of different datasets, which they find to be “fairly satisfactory, although not perfect.” In addition, they test the equivalence of different data for causal inference by replicating two influential studies, one from an established field of comparative inquiry (on the effects of district magnitudes on the effective number of legislative parties) and one within an emergent field (on the effects of electoral rules on national levels of corruption). In their first replication analysis, they find that the choice of datasets may affect results at the margins only yet still carries the potential of producing

some “surprising and contradictory” results. In their second replication study, they show that statistical estimations are highly sensitive to the selection of empirical indicators. The choice of datasets does not affect results at the margins—but at their core.

In his article on the measurement of the rule of law, Svend-Erik Skaaning proceeds in a similar fashion. Limiting his attention to measures of actual (not just formal) levels of rule of law, he selects seven prominent cross-national measures for the purpose of comparative analysis. He critically evaluates the temporal, spatial, and substantive scope of these datasets; their conceptual clarity, consistency, and validity; and the transparency of their measurement choices, aggregation rules, and coding practices. As the author shows, the differences in the content and quality of these measures make a difference for empirical analysis. In descriptive terms, the convergence between some of these datasets is low. Some are even negatively correlated. In causal terms, statistical analyses that estimate the institutional causes of variations in the rule of law are highly sensitive to the choice of specific measures. In the comparative study of both electoral and legal systems, it seems, measurement matters.

The final two articles travel different routes to demonstrate that the choice of measures bears substantive consequences for our empirical inferences. In his contribution on the measurement of labor rights, Emmanuel Teitelbaum evaluates the most prominent cross-national dataset on labor right violations developed by David Kucera. After confirming the one-dimensional structure of the index through item response theory, the author focuses his critique on a classical problem in comparative politics: the problem of conceptual equivalence of identical measures across national contexts. He observes that current practices to measure the protection of rights through manifest violations of rights lead to inaccurate classifications of some labor-repressive countries. Making creative use of his observation of discrepant cases, Teitelbaum amends the dataset and uses his new measure to replicate Kucera’s analysis of the relationship between the guarantee of labor rights and foreign investment flows. The replication study shows that Kucera’s finding of no relationship between rights and investment are robust to a more valid measure of labor rights. The piece nicely illustrates how critical evaluations of data quality can lead to improved data and improved data, in turn, to greater confidence in causal inference.

Finally, in his piece on the measurement of democratization, Matthijs Bogaards examines the empirical

implications of operational choices data users have been taking in the comparative study of political regime change. Since the study of political democratization rests upon the foundation of categorical concepts, comparative scholars have routinely been transforming ordinal scales of democracy and authoritarianism into qualitative categories of regimes and regime change. Bogaards surveys the bewildering multiplicity by which comparative scholars have been operationalizing types of political regimes and instances of regime change on the basis of the continuous measures provided by Polity and Freedom House, the two most widely used data sources on political democracy. As he shows, regardless of whether scholars conceive democratization as a change in regime categories or as a significant move along a continuum of regime properties, their efforts to translate ordinal data into measures of qualitative change have been fraught with conceptual as well as factual uncertainties. He illustrates the empirical implications different operational choices bear by reviewing the recent debate on democratic transitions and war. As the theory posits that political democratization carries the risk of triggering interstate wars, the author revises all instances of interstate war over the past two centuries and examines whether they were preceded by democratizing change according to *any* of the multiple measures of democratization scholars have derived from either Polity or Freedom House. As he finds, the identification of positive cases of democratization that are succeeded by war is highly sensitive to operational choices. The selection of data matters as much as the selection of rules of aggregation (in the definition of regime types) and rules of equivalence (in the definition of continuous democratization). In other words, measurement matters, at each step of the process.

Guest Editors’ Note

First versions of the articles included in this mini-symposium were presented at the workshop “The Numbers We Use, the World We See: Evaluating Cross-National Datasets in Comparative Politics,” European Consortium of Political Research (ECPR), Joint Sessions of Workshops, Rennes, France, April 11-16, 2008. We thank the participants for their invaluable inputs to the project. In addition, we are most grateful to PRQ editors Amy Mazur and Cornell Clayton for their constructive criticism and critical encouragement.

Notes

1. The recent APSA project on Democracy Audits and Governance Indicators convened by the association's new president, Henry Brady (see <http://sites.google.com/site/democracyaudit/>), gives the new awareness to data problems in political science its official seal of approval. For critical overviews on critical issues of cross-national measurement, see Adcock and Collier (2001), Munck and Verkuilen (2002), Herrera and Kapur (2007), and Lieberman (forthcoming).

2. More modest in its ambition, the Indicators Project of the APSA project on Democracy Audits and Governance Indicators (see note 1) lists ninety-five datasets on dimensions of democratic governance, together with some basic information about coverage and access. Less institutionalized data surveys have been produced in diverse areas of research. Landman and Häusermann (2003) offer a critical inventory of data on human rights and democratic governance, while Eck (2005) surveys data on armed conflicts and Jäckle (2008) on government termination. In the comparative study of public opinion, Rubal and Ferrín (2008) offer a systematic comparison of cross-national survey measures on political trust, Facello (2008) on political satisfaction, and Montero et al. (2008) on religiosity.

3. For debates on the measurement of democracy, see, among others, Beetham (1995); Bollen (1980); Bollen and Paxton (2000); Bowman, Lehoucq, and Mahoney (2005); Casper and Tufis (2003); Cheibub, Gandhi, and Vreeland (forthcoming); Collier and Adcock (1999); Collier and Levitsky (1997); Coppedge (2007); Coppedge, Alvarez, and Maldonado (2008); Gleditsch and Ward (1997); Hadenius and Teorell (2005); Inkeles (1991); Mainwaring, Brinks, and Pérez-Liñán (2001); Munck (2009); and Munck and Verkuilen (2002).

4. Based on standards of qualitative research, Lieberman (forthcoming) develops systematic criteria to assess the quality of the empirical evidence that sustains cross-national data and reviews exemplary datasets that follow best practices.

5. For a concise general discussion of source bias, see Rohlfing (2008, 12-17).

References

- Adcock, Robert, and David Collier. 2001. Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review* 95 (3): 529-47.
- Beetham, David, ed. 1995. *Defining and measuring democracy*. Newbury Park, CA: Sage.
- Benoit, Kenneth, and Michael Laver. 2007. Estimating party policy positions: Comparing expert surveys and hand-coded content analysis. *Electoral Studies* 26:90-107.
- Bogaards, Matthijs. 2007. Elections, election outcomes, and democracy in southern Africa. *Democratization* 14 (1): 73-91.
- Bollen, Kenneth A. 1980. Issues in the comparative measurement of political democracy. *American Sociological Review* 45 (3): 370-90.
- Bollen, Kenneth A., and Pamela Paxton. 2000. Subjective measures of liberal democracy. *Comparative Political Studies* 33 (1): 55-86.
- Bowman, Kirk, Fabrice Lehoucq, and James Mahoney. 2005. Measuring political democracy: Case expertise, data adequacy, and Central America. *Comparative Political Studies* 38 (8): 939-70.
- Burkhart, Ross E. 2010. Inequality data and explanations of democracy. *Political Concepts* 41, IPSA Committee on Concepts and Methods, Working Paper Series. www.concepts-methods.org.
- Canache, Damarys, Jeffrey J. Mondak, and Mitchell A. Seligson. 2001. Meaning and measurement in cross-national research on satisfaction with democracy. *Public Opinion Quarterly* 65 (4): 506-28.
- Casper, Gretchen, and Claudiu Tufis. 2003. Correlation versus interchangeability: The limited robustness of empirical findings on democracy using highly correlated datasets. *Political Analysis* 11 (2): 196-203.
- Chandra, Kanchan, and Steven Wilkinson. 2008. Measuring the effect of "ethnicity." *Comparative Political Studies* 41 (4-5): 515-63.
- Cheibub, José Antonio, Jennifer Gandhi, and James Raymond Vreeland. Forthcoming. Democracy and dictatorship revisited. *Public Choice* 142.
- Collier, David, and Robert Adcock. 1999. Democracy and dichotomies: A pragmatic approach to choices about concepts. *Annual Review of Political Science* 2:537-65.
- Collier, David, and Steven Levitsky. 1997. Democracy with adjectives: Conceptual innovation in comparative research. *World Politics* 49 (3): 430-51.
- Coppedge, Michael. 2007. Thickening thin concepts: Issues in large-N data generation. In *Regimes and democracy in Latin America: Theories and methods*, ed. Gerardo L. Munck, 105-22. Oxford: Oxford University Press.
- Coppedge, Michael, Angel Alvarez, and Claudia Maldonado. 2008. Two persistent dimensions of democracy: Contestation and inclusiveness. *Journal of Politics* 70 (3): 632-47.
- Eck, Kristine. 2005. A beginner's guide to conflict data: Finding and using the right dataset. Uppsala Conflict Data Program, UCDP Papers 1, Uppsala, Sweden.
- Elkins, Zachary. 2000. Gradations of democracy? Empirical tests of alternative conceptualizations. *American Journal of Political Science* 44 (2): 287-94.
- Facello, Chiara. 2008. How to measure political (dis)satisfaction? European Consortium of Political Research (ECPR), Joint Sessions of Workshops, Rennes, France, April 11-16.
- Gandhi, Jennifer. 2008. What are we missing? Electoral data in democracies and non-democracies. *Political Concepts* 26, IPSA Committee on Concepts and Methods, Working Paper Series. www.concepts-methods.org.
- Gleditsch, Kristian S., and Michael D. Ward. 1997. Double take: A reexamination of democracy and autocracy in modern polities. *Journal of Conflict Resolution* 41 (3): 361-83.
- Hadenius, Axel, and Jan Teorell. 2005. Assessing alternative indices of democracy. *Political Concepts* 6, IPSA Committee on Concepts and Methods, Working Paper Series. www.concepts-methods.org.
- Hawken, Angela, and Gerardo L. Munck. 2009. Do you know your data? Measurement validity in corruption research. Unpublished typescript, Pepperdine University and University of Southern California, Malibu, CA, and Los Angeles.
- Herrera, Yoshiko M., and Devesh Kapur. 2007. Improving data quality: Actors, incentives, and capabilities. *Political Analysis* 15 (4): 365-86.
- Inkeles, Alex, ed. 1991. *On measuring democracy: Its consequences and concomitants*. New Brunswick, NJ: Transaction Publishers.

- Jäckle, Sebastian. 2008. Datasets for the analysis of government termination. European Consortium of Political Research (ECPR), Joint Sessions of Workshops, Rennes, France, April 11-16.
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing social inquiry: Scientific inference in qualitative research*. Princeton, NJ: Princeton University Press.
- Kittel, Bernhard. 2006. A crazy methodology? On the limits of macro-quantitative social science research. *International Sociology* 21 (5): 647-77.
- Krueger, Alan B., and David D. Laitin. 2004. "Misunderestimating" terrorism. *Foreign Affairs* 83 (5): 8-13.
- Landman, Todd, and Julia Häusermann. 2003. Map-making and analysis of the main international initiatives on developing indicators on democracy and good governance. Unpublished report, University of Essex, Human Rights Centre, and Statistical Office of the Commission of the European Union (Eurostat). www.todd-landman.com.
- Laitin, David, and Daniel Posner. 2001. The implications of constructivism for constructing ethnic fractionalization indices. *APSA-CP Newsletter* 12 (1): 13-17.
- Lieberman, Evan S. 2002. Taxation data as indicators of state-society relations: Possibilities and pitfalls in cross-national research. *Studies in Comparative International Development* 36 (4): 89-115.
- Lieberman, Evan S. 2010. Bridging in qualitative-quantitative divide: Best practices in the development of historically-oriented replication datasets. *Annual Review of Political Science* 13.
- Mainwaring, Scott, Daniel Brinks, and Aníbal Pérez-Liñán. 2001. Classifying political regimes in Latin America, 1945-1999. *Studies in Comparative International Development* 36 (1): 37-65.
- Montero, José Ramón, Guillermo Cordero, María Celeste Ratto, and Gema García Albarcete. 2008. Assessing independent variables in cross national datasets: The case of religiosity. European Consortium of Political Research (ECPR), Joint Sessions of Workshops, Rennes, France, April 11-16.
- Munck, Gerardo L. 2009. *Measuring democracy: A bridge between scholarship and politics*. Baltimore: Johns Hopkins University Press.
- Munck, Gerardo L., and Richard Snyder. 2004. How the concepts we use and the way we measure them shape the world we see. Presented at American Political Science Association (APSA), 100th Annual Meeting, Chicago, September 2-5.
- Munck, Gerardo L., and Jay Verkuilen. 2002. Conceptualizing and measuring democracy: Evaluating alternative indices. *Comparative Political Studies* 35 (1): 5-34.
- Ríos-Figueroa, Julio, and Jeffrey K. Staton. 2008. Unpacking the rule of law: A review of judicial independence measures. Political Concepts 21, IPSA Committee on Concepts and Methods, Working Paper Series. www.concepts-methods.org.
- Rivera, Mauricio. 2010. Estudios sobre represión estatal en regímenes democráticos: Teorías, métodos, hallazgos y desafíos. *Política y Gobierno* 17 (1): 59-95.
- Rohlfing, Ingo. 2008. Do you know your data? Criteria for dataset quality. European Consortium of Political Research (ECPR), Joint Sessions of Workshops, Rennes, France, April 11-16.
- Rubal, Martiño, and Mónica Ferrín. 2008. Measuring trust and its determinants. European Consortium of Political Research (ECPR), Joint Sessions of Workshops, Rennes, France, April 11-16.
- Schedler, Andreas. 2009. Academic market failure: Data quality and availability in comparative politics. Working Paper 215, CIDE, Department of Political Studies, Mexico City.
- Schedler, Andreas. 2012. The measurer's dilemma: Coordination failures in cross-national political data collection. *Comparative Political Studies* 45 (2): forthcoming.
- Schedler, Andreas, and Rodolfo Sarsfield. 2007. Democrats with adjectives: Linking direct and indirect measures of democratic support. *European Journal of Political Research* 46 (5): 637-59.
- Schneider, Aaron. 2003. Decentralization: Conceptualization and measurement. *Studies in Comparative International Development* 38 (3): 32-56.
- Wilson, Sven E., and Daniel M. Butler. 2007. A lot more to do: The sensitivity of time-series cross-section analyses to simple alternative specifications. *Political Analysis* 15:101-23.