



Global Standards, Local Knowledge: The Varieties of Democracy

Version 1.0 (October 22, 2012)

Authors & Principal Investigators: **Michael Coppedge** (University of Notre Dame); **John Gerring** (Boston University); and **Staffan I. Lindberg** (University of Gothenburg & University of Florida)

V-Dem Institute

Institutional Homes:



Supported by:



TABLE OF CONTENTS

V-Dem Team	4
International Advisory Board	4
<u>I. RATIONALE</u>	10
Existing Democracy Indices	10
Definition	12
Precision	12
Sources and Coverage	14
Expert Coding	15
Aggregation	16
Validity and Reliability Tests	18
<u>II. A NEW APPROACH</u>	20
Conceptual Scheme	20
Clarifications	25
<u>III. DATA COLLECTION, AGGREGATION, VALIDATION</u>	27
History of Politics	27
Coding Types	28
Expert Coding	29
Coding Diagnostics	32
Combining Expert Ratings	33
Aggregation Techniques	35
Validation	37
<u>IV. PROJECT MANAGEMENT</u>	38
Project Team	38
Advisory Board	39
V-Dem Institute, Database, Web Site	39
Sustainability	40
International Collaboration and Outreach	41
<u>V. WORK PLAN AND FUNDING</u>	42
The First Phase and the Pilot Study	42
The Second (Current) Phase	43
The Third Phase	43
Funding	43
<u>VI. PAYOFFS</u>	44
General Points	44
Impact Evaluation	47
The Sociology of Measuring Democracy	49
Evaluating V-Dem	51

APPENDIX A: CODEBOOK***PRINTED SEPARATELY******APPENDIX B: COUNTRIES******PRINTED SEPARATELY******APPENDIX C: TAXONOMY******PRINTED SEPARATELY***

V-Dem Team

Further biographical details of the V-Dem team are found on the website (www.v-dem.net).

Principal Investigators

Michael Coppedge

Expertise: Democracy, Dimensions, & Diffusion; Latin America
Professor, Department of Political Science
Faculty Fellow, [Kellogg Institute for International Studies](#)
University of Notre Dame
www.nd.edu/~mcoppedg/crd/coppedge.1@nd.edu

John Gerring

Expertise: Democracy & Development, Research Design
Professor, Department of Political Science
Boston University
<http://people.bu.edu/jgerring/About.html>
jgerring@bu.edu

Staffan I. Lindberg

Expertise: Democracy, Institutions, & Elections; Africa
Associate Professor, Department of Political Science
University of Gothenburg & University of Florida
Research Fellow at the [Quality of Government Institute](#)
<http://www.pol.gu.se/personal/larare-och-forskare/lindberg--Staffan-i/www.clas.ufl.edu/users/sil/sil@ufl.edu> or xlista@gu.se

Project Managers

David Altman

Expertise: Direct Democracy; Latin America
Associate Professor, Department of Political Science
Pontificia Universidad Catolica de Chile
daltman@puc.cl

Michael Bernhard

Expertise: Civil Society; Sovereignty; Central & Eastern Europe
Raymond and Miriam Ehrlich Eminent Scholar
Dept. of Political Science, University of Florida
bernhard@ufl.edu

Steven Fish

Expertise: Legislatures; Post-Soviet States
Professor, Department of Political Science
University of California, Berkeley
sfish@berkeley.edu

Allen Hicken

Expertise: Parties and Party Systems; Asia
Associate Professor, Department of Political Science
Faculty Associate, Center for South East Asian Studies and Center for Political Studies
University of Michigan
ahicken@umich.nedu

Matthew Kroenig

Expertise: Legislatures; Western Europe
Assistant Professor, Department of Government
Georgetown University
mhk32@georgetown.edu

Drew Linzer

Expertise: Measurement Methods
Assistant Professor, Department of Political Science
Emory University
dlinzer@emory.edu

Kelly McMann

Expertise: Subnational Government; Russia & Central Asia

Associate Professor, Department of Political Science
Case Western Reserve University

kelly.mcmann@case.edu

Pamela Paxton

Expertise: Formal & Descriptive Representation

Professor, Department of Sociology
Adams Centennial Professor in the Liberal Arts
University of Texas, Austin

ppaxton@prc.utexas.edu

Megan Reif

Expertise: Election Fraud and Violence

Assistant Professor
Department of Political Science
University of Colorado Denver

megan.reif@ucdenver.edu

Holli Semetko

Expertise: Media; Western Europe

Professor, Department of Political Science
Vice Provost for International Affairs

Director, Office of International Affairs and the Halle

Institute

Emory University

holli.semetko@emory.edu

Svend-Erik Skanning

Expertise: Civil Liberties & HR; Western Europe

Associate Professor, Department of Political Science
Aarhus University

skanning@ps.au.dk

Jeffrey Staton

Expertise: The Judiciary; Latin America

Associate Professor, Department of Political Science
Director of Graduate Admissions & Placement
Emory University

jeffrey.staton@emory.edu

Jan Teorell

Expertise: Democratization, Democracy, & the Executive; Europe

Associate Professor, Department of Political Science
Lund University

<http://www.svet.lu.se/>

jan.teorell@svet.lu.se

Regional Managers

Richard Asante

Expertise: Anglophone West Africa

Research Fellow & Lecturer
Institute of African Studies
University of Ghana

Rasante@ug.edu.gh

Sékou Camara

Expertise: Central Africa

scamar01@hotmail.com

Jorge Vargas Cullell

Expertise: Central America

Director of Programa Estado de la Nación,
Costa Rica

jorgevargas@estadonacion.or.cr

Tiago Fernandes

Expertise: Southern Europe

Assistant Professor, Department of Political Science
New Lisbon University, Portugal

Tiago.Fernandes@EUI.eu

Shpetim Gashi

Expertise: Balkans

Vice President
Council for Inclusive Governance

shpetim.gashi@gmail.com

Carlos Gervasoni

Expertise: South America

Profesor Investigador Ordinario, Department of
Political Science & International Relations
Universidad Torcuato Di Tella, Buenos Aires,
Argentina

cgervasoni@utdt.edu

Mohsin Hashim

Expertise: Bangladesh

Associate Professor of Political Science
Director, Dana Honors Program
Muhlenberg College, USA

mhashim@muhlenberg.edu

Asma Shakir Khawaja

Expertise: South-Central Asia

Assistant Professor, Peace and Conflict Studies,
National Defence University, Islamabad

asmakhawaja@gmail.com

Carl Henrik Knutsen

Expertise: Western Europe

Postdoctoral Fellow, University of Oslo
Associated Researcher at ESOP, University of Oslo
Associated Researcher at CSCW, PRIO

c.h.knutsen@stv.uio.no

Sokol Lleshi

Expertise: Eastern Europe

Doctoral Candidate in Political Science,
Doctoral School of Political Science,
Central European University

koli_ll@yahoo.com

Kate Schuenke Lucien

Expertise: Haiti

Doctoral candidate in Political Science,
University of Notre Dame

kschuenk@nd.edu

Njekwa Mate

Expertise: Southern Africa

Head/Lecturer
School of Humanities and Social Sciences
Political and Administrative Studies Department
Lusaka, The University of Zambia, Zambia

nmateza@yahoo.co.uk

Daniel G Ogbaharya

Expertise: Horn of Africa

Instructor, Western Illinois University

dg-ogbaharya@wiu.edu

Ralph Premdas

Expertise: Anglophone Caribbean

Professor of Public Policy,
Sir Arthur Lewis Institute Of Social and Economic
Studies (SALISES),
University of the West Indies, Trinidad and Tobago

ralphpremdas@hotmail.com;

ralph.premdas@sta.uwi.edu

Solofo Randrianja

Expertise: Madagascar

Political Historian at Univ of Toamasina

solalao@hotmail.com

Abdoul Karim Saidou

*Expertise: Anglophone Central &
Eastern Africa*

Assistant Professor of Political Science
Obafemi Awolowo

University, Ile Ife, Nigeria

afolabiolugbemiga@yahoo.ca

Afolabi Olugbemiga Samuel

Expertise: Anglophone Central & Eastern Africa
Assistant Professor, Department of Political Science
Obafemi Awolowo University, Ile Ife, Nigeria

afolabiolugbemiga@yahoo.ca

Carlos Shenga

Expertise: Lusophone Africa

Director, Centre for Research on Governance and
Development, Mozambique
Researcher, Centre for Social Science Research,
University of Cape Town, South Africa

cshenga@yahoo.com

Martin Tanaka

Expertise: Andean Countries Principal Investigator
Instituto de Estudios Peruanos

Associate Professor

Department of Social Sciences

Pontificia Universidad Católica, Lima, Peru

mtanaka@iep.org.pe

Julio C. Teehankee

Southeast Asia

Associate Professor and Chair for the International
Studies Department

De La Salle University, Manila, Philippines

jteehankee@gmail.com

Kharis Templeman

Expertise: East Asia & Southeast Asia

PhD, Political Science

University of Michigan, USA

kharist@umich.edu

Medet Tiulegenov

Expertise: Central Asia

Assistant Professor

International and Comparative Politics Department
American University of Central Asia

tiulegenov@gmail.com

Project Coordinators

Krystin Krause
University of Notre Dame
Contact7@v-dem.net

Cecilia Pe Lero
University of Notre Dame
contact5@v-dem.net

Josefine Pernes
University of Gothenburg
Contact9@v-dem.net

Natalia Stepanova
University of Gothenburg
contact2@v-dem.net

Key Research Assistants

Vlad Ciobanu
University of Gothenburg
contact3@v-dem.net

Ionut Gusoi
University of Gothenburg
contact8@v-dem.net

Eva Karlsdóttir
University of Gothenburg
evathorawork@gmail.com

Keith Weghorst
University of Florida, Gainesville
keith.weghorst@ufl.edu

Josh Yesnowitz
Boston University
contact6@v-dem.net

Laleh Kazemi Veisari
University of Gothenburg
kazemi.laleh@gmail.com

Undergraduate Research Assistants

Victoria Anglin
University of Notre Dame
vanglin@nd.edu

Alexaundria Barnes
University of Notre Dame
abarnes5@nd.edu

Gable Brady
University of Notre Dame
gbrady@nd.edu

Christopher Brown
University of Notre Dame
cbrown35@nd.edu

Jordan Brown
University of Notre Dame
jbrown22@nd.edu

Micah Burbanks-Ivey
University of Notre Dame
Mburbank@nd.edu

Daniel Colston
University of Notre Dame
dcolston@nd.edu

Ana Dionne-Lanier
University of Notre Dame
adionnel@nd.edu

Steven Fisher
University of Notre Dame
sfisher4@nd.edu

Taylor Freeman
University of Notre Dame
tfreema3@nd.edu

Monica O'Hearn
University of Notre Dame
mohearn@nd.edu

Matthieu Huon
University of Notre Dame
mhuon1@nd.edu

Amy McDill
University of Notre Dame
amcdill@nd.edu

Anthony Mendoza
University of Notre Dame
amendoz3@nd.edu

Angel Mira
University of Notre Dame
amira@nd.edu

Chris Newton
University of Notre Dame
cnewton@nd.edu

AnnaLee Rice
University of Notre Dame
arice3@nd.edu

Ryan Schultheis
University of Notre Dame
rschult2@nd.edu

Marcilena Shaeffer
University of Notre Dame
mshaeffe@nd.edu

Yiren Shen
University of Notre Dame
yshen2@nd.edu

Qineng Shi
University of Notre Dame
qshi2@nd.edu

Tim Sullivan
University of Notre Dame
tsulliv8@nd.edu

Breanna Thomas
University of Notre Dame
bthoma11@nd.edu

Celeste Villa-Rangel
University of Notre Dame
cvillara@nd.edu

Zhizhou Wang
University of Notre Dame
zwang7@nd.edu

Njeri Williams
University of Notre Dame
nwilli11@nd.edu

Former Research Assistants:

Yannik Bendel, University of Gothenburg, Annika Fredén, University of Gothenburg, Nicki Khorram-Manesh, University of Gothenburg, Sandra Botero, University of Notre Dame, Chad Kiewiet de Jonge, University of Notre Dame, Haixiao (Erik) Wang, University of Notre Dame, Lauren Metayer, University of Notre Dame

International Advisory Board

Aitzaz Ahsan - *Senior Advocate, Supreme Court of Pakistan*

Abdullah M. Alsaidi - *Senior Fellow, International Peace Institute*

Larry Diamond - *Senior Fellow, Hoover Institution.*

Sonia Fleury – *Professor, Department of Political Science, Brazilian School of Public Administration and Business, Rio de Janeiro.*

Emmanuel Gyimah-Boadi – *Professor, Department of Political Science, University of Ghana; Executive Director, the Ghana Center for Democratic Development (CDD-Ghana); co-Director, Afrobarometer project*

Cyd Ho Sau-lan – member, Legislative Council, Hong Kong; chair, Hong Kong Human Rights Monitor; founding councilor, World Future Council; co-Convener, Project Civil Referendum

Niraja Gopal Jayal - Professor, Centre for the Study of Law and Governance, Jawaharlal Nehru University, New Delhi

Jeremy K. B. Kinsman – Director, Community of Democracies project; Regent's Lecturer, University of California, Berkeley; Diplomat-in-Residence, the Woodrow Wilson School, Princeton University

Augustin Loada – Professor, Departments of Law and Political Science, University of Ouagadougou, Burkina Faso; Executive Director, Centre pour la Gouvernance Démocratique (CGD), Ouagadougou

Soledad Loaeza – Professor of Political Science, Center for International Studies, El Colegio de México, Mexico City; member, Board of Advisers, International Institute for Democracy and Electoral Assistance (IDEA)

Pratap Bhanu Mehta – president, Centre for Policy Research, New Delhi; member, Government of India's National Security Advisory Board

Chung-in Moon - Professor, Department of Political Science, Yonsei University, Seoul; Editor-in-Chief, Global Asia

Saiful Mujani - Executive Director and cofounder, Lembaga Survei Indonesia (LSI, Indonesian Survey Institute); Professor of Social and Political Sciences, State Islamic University of Jakarta; Director of Research, Freedom Institute, Jakarta

Ziya Önis - Professor of International Political Economy; Director of the Center for Research on Globalization and Democratic Governance (GLODEM), Koç University, Istanbul

Oyun Sanjaasuren – member, Great State Assembly (parliament) of Mongolia; Chair of the Zorig Foundation; President, Mongolian Geological Association

Lilia Shevtsova – Senior Associate and Chair, Russian Domestic Politics and Political Institutions Program, Carnegie Endowment for International Peace, Moscow Center; Professor, Department of Political Science, Moscow State Institute of International Affairs of the Ministry of Foreign Affairs of the Russian Federation

G. J. W. Van Oven - Vice-President, Court of Appeal in the Hague; President, Judges for Judges Foundation

I. Rationale

In the wake of the Cold War, democracy has gained the status of a mantra.¹ Yet there is no consensus about how to conceptualize and measure it well enough to support meaningful and accurate comparisons through time and across countries. Skeptics may question whether such comparisons are possible. Nevertheless, there is a clear need to measure democracy so that we can mark its progress and setbacks, explain the process of democratization, reveal its consequences, and affect its future course. The need is so strong that crossnational indices such as Freedom House and Polity are used all over the globe even though they are known to suffer major shortcomings.

For policymakers, activists, academics, and citizens around the world the conceptualization and measurement of democracy matters. Billions of dollars in foreign aid intended to promote democracy and governance in the developing world are contingent upon judgments about how democratic a polity is, its recent history, future prospects, and the likely causal effect of particular forms of assistance. Likewise, a large body of work in social science deals with these same issues, i.e., the nature, causes, consequences, and trajectories of democracy around the world.

This document lays out a new approach to conceptualization and measurement of democracy, dubbed *Varieties of Democracy* (V-Dem). In Part I, we review the weaknesses of extant approaches. In Part II, we lay out our approach, which may be characterized as historical, multidimensional, and disaggregated. In Part III, we discuss the process of data collection. In Part IV, we describe project management. In Part V, we discuss what we have accomplished to date and what remains to be accomplished. In Part VI, we suggest the payoffs our approach may bring to the study of democracy and some outcome indicators that will be used to assess V-Dem's impact on several communities it is intended to serve.

Three appendices to this document are printed separately: (a) the *Codebook*, with a comprehensive list of indicators, (b) *Countries* covered by V-Dem, with notes pertaining to years covered and any special circumstances that may apply (e.g., with respect to a country's pre-independence history), and (c) a *Taxonomy* showing how each principle of democracy relates to lower-level components and indicators. Each is available (or will shortly be available) for download on the V-Dem web site (v-dem.net), along with data from the Pilot study, as described below.

Existing Democracy Indices

Critiques of democracy indices are legion.² Here, we will touch briefly on seven key issues of

¹ This paper integrates material previously published in Coppedge, Gerring, et al. (2011). The authors thank the many people who have generously provided comments and feedback at various stages of this project.

² See Acuna-Alfaro (2005), Beetham (1994), Berg-Schlosser (2004a, 2004b), Bollen (1993), Bollen and Paxton (2000), Bowman, Lehoucq, and Mahoney (2005), Coppedge, Alvarez, and Maldonado (2008), Foweraker and Krznaric (2000), Gleditsch and Ward (1997), Lindberg (2006), McHenry (2000), Munck (2009), Munck and

conceptualization and measurement: definition, precision, sources and coverage, coding, aggregation, and tests of validity and reliability.

Our discussion focuses on the most prominent efforts including the Political Rights, Civil Liberties, Nations in Transit, and Countries at the Crossroads indices, all sponsored by Freedom House; the Polity2 variable from the Polity IV database; a binary measure of democracy and dictatorship (“DD”) constructed by Adam Przeworski and colleagues; a binary measure constructed by Michael Bernhard, Timothy Nordstrom, and Christopher Reenock (“BNR”); a binary measure constructed by Carles Boix and colleagues (forthcoming); a multidimensional index produced by the Economist Intelligence Unit (“EIU”); the Bertelsmann Transformation Index (“BTI”) funded by the Bertelsmann Foundation; NELDA (Hyde, Marinov 2012); and the Democracy Barometer.³ Glancing reference will be made to other indices and many of the points made in the following discussion probably apply more broadly.⁴ The following exercise does not purport to provide a comprehensive review; it merely calls attention to the sort of problems common to this field.⁵

In our discussion, an *index* will be understood as a highly aggregated, composite measure of democracy such as Polity2, while an *indicator* will be understood as a more specific, disaggregated element of democracy such as turnout.

Verkuilen (2002), Treier and Jackman (2008), Vermillion (2006). For work focused more generally on governance indicators, see Arndt and Oman (2006), Besancon (2003), Kurtz and Schrank (2007), Sudders and Nahem (2004), Thomas (2010), USAID (1998).

³ Freedom House employs two indices, “Political Rights” and “Civil Liberties” (sometimes they are employed in tandem, sometimes singly) each of which extends back to 1972 and covers most sovereign and semi-sovereign nations (see www.freedomhouse.org). Polity IV (Marshall and Jaggers 2007) also provides two aggregate indices, “Democracy” and “Autocracy,” usually used in tandem (by subtracting one from the other), which provides the Polity2 variable. Coverage extends back to 1800 for most sovereign countries with populations greater than 500,000 (www.cidcm.umd.edu/inscr/polity). DD (Alvarez, Cheibub, Limongi, and Przeworski 1996; Cheibub, Gandhi, and Vreeland 2010) codes countries dichotomously (democracy/dictatorship), including most sovereign countries from 1946 to the present. BNR (2001) construct a binary measure that extends (for all sovereign nations) from 1919 to the present. Boix et al. (forthcoming) is similar to DD, though with much better historical coverage and less strict adherence to the turnover rule. The EIU index is composed of five core dimensions and sixty sub-components, which are combined into a single index of democracy (EIU 2010). Coverage extends to 167 sovereign or semisovereign nations for 2006, 2008, and 2010. The Democracy Barometer (Buhlmann et al. 2011) examines the quality of democracy within thirty consolidated democracies (mostly in the OECD) from 1995 to the present with 100 quantitative indicators.

⁴ E.g., Competition and Participation variables (Vanhanen 2000), the Polyarchy index (Coppedge and Reinicke 1990), Contestation and Inclusiveness indices (Coppedge, Alvarez, and Maldonado 2008), the Political Regime Change [PRC] dataset (Gasirowski 1996; updated by Reich 2002), the Democratization Dataset (Schneider and Schmitter (2004a), Unified Democracy Scores (Pemstein, Meserve, and Melton 2010, the Democracy Barometer (Buhlmann et al. 2011), and indicators based on Altman and Pérez-Liñán (2002), Arat (1991), Bollen (1980, 2001), Bowman, Lehoucq, and Mahoney (2005), Hadenius (1992), and Moon et al. (2006).

⁵ Detailed surveys can be found in Hadenius and Teorell (2005a), Landman (2003), and Munck and Verkuilen (2002). See also Acuña-Alfaro (2005), Beetham (1994), Berg-Schlosser (2004a, 2004b), Bollen (1993), Bollen and Paxton (2000), Bowman, Lehoucq, and Mahoney (2005), Casper and Tufis (2003), Elkins (2000), Foweraker and Krznaric (2000), Gleditsch and Ward (1997), McHenry (2000), Munck and Verkuilen (2002), and Treier and Jackman (2008).

Definition

Democracy, understood in a very general way, means rule by the people. This common denominator claims a long heritage stretching back to the Classical age (Held 2006). All usages of the term also presume sovereignty. A polity must enjoy some degree of self-government in order for democracy to be realized.

Beyond these core definitional elements there is great debate. The debate covers both descriptive and normative aspects; what political regimes are (or reasonably could be) and what they ought to be.⁶ Since definitional consensus is necessary for obtaining consensus over measurement, the goal of arriving at a single universally accepted measure of democracy is impossible as long as this great debate continues.

For example, the Polity2 index rates the United States as fully democratic throughout the twentieth century and much of the nineteenth century. This is a fair conclusion if one disregards the composition of the electorate—from which women and most blacks were excluded—in one's definition of democracy (Paxton 2000). Similar challenges could be levied against other indices that omit consideration of attributes that some regard as definitional. These omissions are particularly glaring where democracy is defined in a minimal fashion.

More encompassing conceptions of democracy may be criticized for including elements that fall far from the core meaning of the term. For example, the Political Rights index includes questions pertaining to corruption, civilian control of the police, the absence of widespread violent crime, willingness to grant political asylum, the right to buy and sell land, and the distribution of state enterprise profits (Freedom House 2007). The authors of the index would argue that it measures freedom, not democracy; nevertheless, it is frequently used as an index of democracy. If such attributes were part of a larger project to measure many distinct dimensions of democracy, such choices could be defensible; but when they are combined into a single index alongside indicators that probably lie on other dimensions, they are problematic.

In other instances, it is the judgments embraced within an index that are problematic. For example, the EIU index regards mandatory voting as reflecting negatively on the quality of democracy in a country. While this provision infringes upon individual rights and in this respect may be considered undemocratic, it also enhances turnout and, arguably, the quality of representation. Its status in enhancing rule by the people is therefore controversial, depending on one's conception of democracy.

The methodological problems affecting contemporary indices therefore begin at the level of definition. Our solution, discussed in more detail below, is to create indices of *all* the main varieties of democracy found in the literature, as well as their component parts, so that countries can be compared over a range of definitions.

Precision

Many of the leading democracy indices are insensitive to important gradations in the degree or quality of democracy. If one purpose of any measurement instrument is discrimination

⁶ Studies of the concept of democracy are legion. See, e.g., Beetham (1994, 1999), Collier and Levitsky (1997), Held (2006), Lively (1975), Saward (2003), Weale (1999). All emphasize the far-flung nature of the core concept.

(Jackman 2008), extant democracy indices fall short of the ideal.

At the extreme, binary measures such as DD, BNR, and Boix et al. reduce democracy to a dummy variable useful for purposes such as analyzing the duration of democratic regimes (Collier and Adcock 1999). However, this dichotomous coding lumps together polities that exhibit disparate regime qualities (Bollen, Jackman 1989; Elkins 2000). For example, the DD index recognizes no distinctions within the large category of countries that have competitive elections and occasional leadership turnover. Papua New Guinea and Sweden thus receive the same score, despite evident differences in the quality of elections, civil liberties, and barriers to competition. Thus, although binary indices serve an important and indispensable function, they cannot be used to capture fine distinctions across regimes or through time.

Continuous measures appear to be more sensitive to gradations of democracy/autocracy because they have more ranks. Freedom House scores democracy on a seven-point index (13 points if the Political Rights and Civil Liberties indices are combined). Polity provides a total of 21 points if the Democracy and Autocracy scales are merged (an aggregation procedure suggested by the data providers), creating the Polity2 variable. Appearances, however, can be deceiving. Polity scores, for example, bunch up at a few thresholds (notably -7 and +10), suggesting that the scale is not as sensitive as it purports to be. The EIU index and Democracy Barometer are the most smoothly continuous of the extant indices. However, as with Polity, aggregate scores are reflections of underlying components that can be quite different (and therefore are obscured by averaging).⁷

In sum, the precision or reliability of all indices is often too low to justify confidence that a country that scores a few points higher than another is actually more democratic (Armstrong 2011; Pemstein et al. 2010). In fact, according to one of the most rigorous analyses, Polity scores are so imprecise that one cannot be confident that the United States was any more democratic than the top 70 of 153 countries in 2000 (Treier and Jackman 2008). Note that most extant indices are bounded and there is no way to distinguish the quality of democracy among countries that have perfect scores. This is acceptable so long as there really is no difference in the quality of democracy among these countries, or within a single country (receiving the same score) through time. However, this assumption may be doubted. According to Polity, Sweden's democracy has not changed since the extension of male suffrage in 1917 in Polity; yet, there have been several significant reforms since that time, which presumably have impacted the quality of Swedish democracy. In 2004, Freedom House assigned the highest score on its Political Rights index to countries as dissimilar as Andorra, Bulgaria, Denmark, Israel, Mauritius, Nauru, Panama, South Africa, Uruguay, and the United States. There are substantial differences in the quality of democracy among these diverse polities and others within each level of the Freedom House scale (Armstrong 2011); however, these differentiating elements are missing from extant democracy indices (Lijphart 1999).

To some extent, problems of imprecision can be overcome by the application of Bayesian models (Armstrong 2011; Bollen 2001; Pemstein et al. 2010) or factor analysis (Coppedge et al.

⁷ Questions can also be raised about whether these indices are properly regarded as interval scales; see Treier and Jackman (2008). We do not envision an easy solution to this problem although Pemstein, Meserve, and Melton (2010) offer some intriguing ideas.

2008), which allow one to combine information from many extant indices or from multiple sub-components of a single index and to assign confidence intervals to each point estimate. However, these techniques do not compensate for the shortcomings of the underlying data, which tends to be aggregated at a very high level and therefore inappropriate for ascertaining small differences through time or across countries.

Sources and Coverage

Many democracy indices are limited in temporal or country coverage. Nations in Transit covers only the post-communist states. Countries at the Crossroads covers seventy countries (beginning in 2004) deemed to be strategically important and at a critical juncture in their trajectory. The Democracy Barometer covers 30 countries, located exclusively in the OECD. The Political Rights and Civil Liberties indices stretch back to 1972 and include most sovereign and semisovereign states. DD begins in 1946, BNR begins in 1919, BTI begins in 2003, EIU begins in 2006, and the Democracy Barometer begins in 1990.

Only a few democracy indices stretch back further in historical time—notably, Polity (1800–), Boix et al. (1800–), and Tatu Vanhanen’s (2000) index of democratization (1810–). We suspect that the value of these indices stem partly from their fairly comprehensive historical coverage—though Polity excludes states with fewer than 500,000 inhabitants and none of these indices attempt to include colonies prior to independence, even if they enjoyed substantial self-government.

The sources employed to provide coding for extant indices are often problematic. For example, the Political Rights and Civil Liberty indices rely heavily on secondary accounts such as the *New York Times* and Keesing’s Contemporary Archives for coding in the 1970s and 1980s. These historical sources, while informative, do not provide equally comprehensive coverage of every country in the world, introducing potential bias into the resulting indices. In later eras, these indices have relied much more on country expert coding. However, the change from one source of evidence to another—coupled with some possible changes in coding procedures—may have compromised the continuity of the time-series. No effort has been made to revise previous scores so that they are consistent with current coding criteria and expanded knowledge of past regimes.⁸

Some indices such as the EIU rely heavily on polling data, which is available on a non-comparable and highly irregular basis for 100 or so nation-states. For other countries (about half of the population covered by the EIU) these data must be estimated by country experts or imputed. Procedures employed for this estimation are not made publicly available.⁹ Although surveys of citizens are important for ascertaining attitudes, they are not available for every country in the world, and in no country are they available on an annual basis. Moreover, use of such surveys severely limits the historical reach of any democracy index, since the origin of

⁸ Gerardo Munck, personal communication (2010).

⁹ Reliance on survey data also raises even more difficult questions about validity, i.e., whether the indicators measure what they are supposed to measure. There is surprisingly little empirical support for the notion that respondents are able to assess their own regimes in a cross-nationally comparable way or that they tend to live under regimes that are congruent with their own values.

systematic surveying stretches back only a half-century (in the US and parts of Europe) and is much more recent in most countries. Other survey-based questions are of questionable relevance for understanding the quality of democracy in a polity. It is of course interesting to know whether citizens regard their country as democratic, whether they support democratic institutions and practices, and whether they subscribe to democratic norms such as tolerance. But it is not clear whether such attitudes make a country more or less democratic. It is striking that popular support for democracy, as gauged by standardized surveys such as the Afrobarometer and LAPOP, sometimes varies inversely with standard indices of democracy (such as Polity2 and the Political Rights index) that rest on the institutional features of polities (cite?).

Expert Coding

Several indices (including Freedom House, EIU, and BTI) rely on expert coding. Such judgments can be made fairly reliably if there are clear and concrete coding criteria (Munck and Verkuilen 2002). Unfortunately, this is not always the case. For example, the Nations in Transit expert survey poses five sub-questions to answering the question, “Is the country’s governmental system democratic?”:

1. Does the Constitution or other national legislation enshrine the principles of democratic government?
2. Is the government open to meaningful citizen participation in political processes and decision-making in practice?
3. Is there an effective system of checks and balances among legislative, executive, and judicial authority?
4. Does a freedom of information act or similar legislation ensure access to government information by citizens and the media?
5. Is the economy free of government domination?¹⁰

Quite aside from the debatable validity of equating democracy with separation of powers and a free-market economy, these are not easy questions to answer, and their difficulty stems from the general, vague, or ambiguous terms in which they are posed. One cannot judge whether the “principles of democratic government” are “enshrined” without specifying what those principles are. Are these principles enshrined if they are written on parchment but not practiced? What does it mean to be “open to meaningful citizen participation”? What is the basis for determining whether checks and balances are “effective”? What degrees and kinds of government regulation and ownership can be permitted before economic “freedom” is infringed upon? Wherever the questions do not define these criteria, respondents must rely on their own hidden beliefs and assumptions. This creates a danger that coding decisions about *particular* topics—e.g., press freedom—will reflect an overall sense of how democratic a country is rather than an independent evaluation of the level of press freedom. In this respect, “disaggregated” indices may actually be considerably less disaggregated than they appear. It is the ambiguity of the questionnaires underlying these surveys that foster this sort of premature aggregation.

¹⁰ Report on Methodology, downloaded from www.freedomhouse.hu/images/nit2009/methodology.pdf.

Aggregation

Since democracy is a multi-faceted concept all composite indices must wrestle with the aggregation problem— which indicators to combine into a single index, whether to add or multiply them, and how much to weight them. It goes without saying that different solutions to the aggregation problem lead to quite different results (Munck and Verkuilen 2002; Munck 2009; Goertz 2006). This is a very consequential decision, and because it is especially central to the V-Dem project we shall spend a bit of time discussing it.

The aggregation rules used by almost all measures are additive, with an (implicit or explicit) weighting scheme; they are sums, averages, or weighted averages of various components. It is far from obvious that this is the most appropriate aggregation rule. Others recommend that one should consider the various sub-components of democracy as necessary (non-substitutable), mutually constitutive (interactive), or both (Goertz 2006: 95–127; Munck 2009; Schneider 2010). In order for any aggregation scheme to be successful, rules must be clear, they must be operational, and they must reflect an accepted definition of democracy. Otherwise, the resulting measure is not valid. Although most indices have fairly explicit aggregation rules, they are sometimes difficult to comprehend and consequently to apply (e.g., Polity). They may also include “wild card” elements, allowing the coder free rein to assign a final score, in accordance with his or her overall impression of a country (e.g., Freedom House). And they may be altogether inappropriate when the relevance or weight of one component depends on the values of other components.

A more inductive approach may also be taken to the aggregation problem. Michael Coppedge, Angel Alvarez, and Claudia Maldonado (2008) apply an exploratory factor analysis of a large set of democracy indicators, identifying two dimensions that they label Contestation and Inclusiveness. Other writers analyze extant indices as reflections of a (unidimensional) latent variable.¹¹ These inductive approaches allow for the incorporation of diverse data sources and may provide uncertainty estimates for each point score. However, problems of definition are implicit in any factor-analytic or latent-variable index, for an author must decide which indicators to include in the sample— requiring a judgment about which extant indices are measuring “democracy” and which are not—and how to interpret commonalities among the chosen indicators. This is not solvable simply by referring to the labels assigned to the indicators in question, as many of the most well known and widely regarded democracy indices are packaged as “rights,” “liberties,” or “freedom” rather than democracy, and do not necessarily measure exactly the concepts they purport to measure. Moreover, while factor-analytic and latent variable approaches allow for the incorporation of multiple sources of data, thereby reducing some sources of error, they remain biased by any systematic error common to the chosen data sources.

Another approach to the aggregation problem is to collect primary data at a disaggregated level, letting end-users decide whether and how to aggregate it. Democracy

¹¹ Pemstein, Meserve, and Melton (2010); Bollen and Jackman (1989); Bollen and Paxton (2000); Treier and Jackman (2008).

assessments (aka audits) provide detailed indicators for a single country.¹² This sort of detailed inquisition into the quality of democracy is very much in the spirit of our enterprise, except that the collected data are limited to a single country (and often to recent years) and are not comparable across other countries where detailed assessments have been conducted.

Several specific topics integral to democracy have been successfully measured on a global scale. And several broader ventures to measure democracy in a comprehensive and disaggregated fashion have been proposed, but not fully implemented (e.g., Beetham et al. 2001). These efforts at disaggregation push in the right direction. However, they are problematic on one or several accounts: (a) the number of indicators are very small; (b) the indicators are still highly abstract and hence difficult to operationalize; (c) the underlying components, while conceptually distinct, are gathered in such a way as to compromise their empirical independence; (d) the information necessary to code the indicator is not available across nations or prior to the contemporary era; or (e) the indicators are not released to the general public.

Consider the Polity index, which is ostensibly disaggregated into five components: competitiveness of participation, regulation of participation, competitiveness of executive recruitment, openness of executive recruitment, and constraints on executive. Although each of these components is described at length in the Polity codebook (Marshall and Jaggers 2007), it is difficult to say precisely how they would be coded in particular instances, or how the stated aggregation principles lead to an overall score for a given country in a given year (Munck and Verkuilen 2002). Even in disaggregated form (e.g. Gates et al. 2006), the Polity index is highly abstract, and therefore open to diverse interpretations.

The two principal Freedom House measures—Civil Liberties and Political Rights—are similarly difficult to interpret. Since 2006, Freedom House has released coding scores for the components of Civil Liberties and Political Rights. The Political Rights index is shown to be the product of (a) Electoral Process, (b) Pluralism and Participation, and (c) Functioning of Government. The Civil Liberties index comprises (a) Freedom of Expression, (b) Association and Organizational Rights, (c) Rule of Law, and (d) Personal Autonomy and Individual Rights. This represents a step towards disaggregation, yet intercorrelations among the seven components are extremely high—Pearson's $r = 0.86$ or higher. This by itself is not necessarily problematic; it is possible that all democratic (or nondemocratic) things go together. However, the high intercorrelations of the Freedom House indicators coupled with their ambiguous coding procedures suggest that these components may not be entirely independent of one another. It is impossible to discard the possibility that country coders have a general idea of how democratic each country is, and that this idea is reflected in consistent scores across the multiple indicators. Components that are scored separately on a questionnaire may not be independently coded.

The EIU (2010) index does a better job of disaggregating components, which are reported for five dimensions: electoral process and pluralism, civil liberties, the functioning of government, political participation, and political culture. Correlations are still quite high,

¹² E.g., Beetham (1994, 2004), Diamond and Morlino (2005), Landman and Carvalho (2008), Proyecto Estado de la Nación (2001).

ranging from .74 to .93 (except for the cultural variable, which is distinctive). Moreover, the specificity of the questions makes the claim of independence among these five variables plausible. Unfortunately, EIU is unwilling to divulge data for the sixty specific questions that compose the five dimensions, so it is difficult to judge the accuracy and independence of the index. It may be useful, or it may be not, but we cannot know for sure.

Validity and Reliability Tests

Worries about validity in extant democracy indices are nourished by periodic appraisals focused on specific countries. A recent study by scholars of Central America alleges major flaws in coding for Costa Rica, El Salvador, Honduras, Guatemala, and Nicaragua in three crossnational indices—generated by Polity, Vanhanen, and Gasiorowski (1996) — errors that, the authors suspect, also characterize other indices and other countries (Bowman, Lehoucq, and Mahoney 2005).

Surprisingly, inter-coder reliability tests are not common practice among democracy indices. Freedom House does not conduct such tests, or at least does not make them public. Polity used to do so, but it required a good deal of hands-on training before coders reached an acceptable level of coding accuracy. This suggests that other coders might not reach the same decisions simply by reading Polity's coding manual. And this, in turn, can contribute to the problem of conceptual validity, in which key concepts are not well matched to the empirical data.

These critiques notwithstanding, defenders of contemporary indices often point out that extant indices are highly intercorrelated. Indeed, the correlation between Polity2 and Political Rights – the dominant indices, by most accounts – is a respectable 0.88 (Pearson's r). Yet, closer examination reveals that the consensus is largely the product of countries lying at the democratic extreme—Canada, Sweden, the United States, *et al.* When countries with the top two scores on the Freedom House Political Rights scale are eliminated, Pearson's r drops to 0.63. This is not an impressive level, especially when one considers that scholars and policymakers are usually interested in precisely those countries lying in the middle and bottom of the distribution—countries that are undemocratic or imperfectly democratic.¹³

These measurement differences are part of the reason for some of the divergent findings in empirical work where democracy is a key variable. Przeworski and Limongi (1997), for example, found that per capita income was not associated with transitions to democracy, but Zachary Elkins (2000) showed that their result depended on using a binary measure of democracy; when graded measures such as Freedom House or Polity were substituted, the correlation between income and transitions returned to significance. More generally, Gretchen Casper and Claudiu Tufis (2003) showed that few explanatory variables (beyond per capita income) have a consistently significant correlation with levels of democracy when different democracy indices are employed. In fact, the *only* predictor that both remained significant regardless of how democracy was measured and survived all their robustness checks was Rae's index of party-system fractionalization (an aspect of party competition, which is arguably a

¹³ For extensive cross-country tests see Hadenius and Teorell (2005a).

component of democracy itself).¹⁴ Thus, we have good reasons to suspect that extant indices suffer problems of validity and reliability and that these problems are consequential. They impact what we think is going on in the world.

¹⁴ See also Hadenius and Teorell (2005a) and Högström (2013).

II. A New Approach

The task of constructing a global index of democracy that is valid, precise, and universally accepted is practically impossible, for all the reasons we have discussed. To be sure, if all we needed to do were to identify major regime changes, gross differences in levels of democracy, or track trends in the average global level of democracy, we could probably get by with existing measures. But development agencies, international organizations, NGOs, journalists, educators, and researchers need measures that can do more than this. Existing measures of democracy are especially inadequate for measuring smaller changes and differences in the degree of democracy or autocracy; empirically analyzing relationships among various elements of democracy; and evaluating the effectiveness of targeted democracy promotion efforts. Polity, Freedom House, and their counterparts are inadequate for these tasks. Just as critically, none of them provides a variety of measures that would enable scholars, institutions, policy-makers, and civil society actors to measure exactly what *they* mean by “democracy” so that they can support chosen comparisons and contrasts with reliable empirical data.

Four features, considered together, distinguish our approach relative to extant indices. First, we approach democracy as a multi-dimensional concept with many components. Second, we collect information on indicators relevant to democracy at a highly disaggregated level and make all data (aggregated and disaggregated) freely available. Third, we extend indicators of democracy back through modern history to 1900. Finally, we offer multiple aggregation techniques for all higher-order concepts, each with clear and transparent rules that can easily be replicated by end-users. In this section, we introduce the V-Dem conceptual scheme.

Conceptual Scheme

There is no consensus on what democracy at large means, beyond the vague notion of rule by the people. Political theorists have been emphasizing this point for some time, and empiricists would do well to take the lesson to heart (Gallie 1956; Held 2006; Shapiro 2003: 10–34).

Even so, there appears to be relative consensus on the various normative principles that constitute this protean term. Seven key principles seem paramount: electoral, liberal, majoritarian, consensual, participatory, deliberative, and egalitarian, as summarized in Table 1. Each represents a different way of understanding what is required for “rule by the people.” Thus, while no single principle can reasonably purport to embody all the meanings of democracy, these seven principles, taken together, offer a fairly comprehensive accounting of the concept of democracy as it is employed today.¹⁵

The *electoral* principle of democracy—identified with contestation, competition, and elite, minimal, realist, or Schumpeterian democracy—is the idea that democracy is achieved

¹⁵ The consensus is only relative: it could be debated if the list should consist of seven, six, or five principles, whether they should be called “principles” or “dimensions,” and whether “majoritarian” and “consensual” should both be listed when they can be thought of as opposite poles on a single dimension. This list of seven principles is intended to provide assurance that we are trying to take a comprehensive inventory of core values of democracy that is likely to include almost all the components that any user would want to have measured.

through competition among leadership groups, which vie for the electorate's approval during periodic elections before a broad electorate. Parties and elections are the critical instruments in this largely procedural account of the democratic process. Of course, many additional factors might be regarded as important for ensuring and enhancing electoral contestation, e.g., civil liberties, an active media, a written constitution, an independent judiciary (to enforce the rules of the game), and so forth. However, these factors are viewed as secondary to electoral institutions (Dahl 1956; Przeworski et al. 2000; Schumpeter 1950).

The *liberal* principle of democracy stresses the intrinsic importance of transparency, civil liberty, rule of law, horizontal accountability (effective checks on rulers), and minority rights. These are seen as defining features of democracy, not simply as aids to political competition. The liberal model takes a “negative” view of political power insofar as it judges the quality of democracy by the limits placed on government. Principles and procedures must be established so as to ensure that rule by the majority does not result in the oppression of minorities or the loss of individual liberties.¹⁶

The *majoritarian* principle of democracy (aka responsible party government) reflects the principle that the will of the majority should be sovereign. The many should prevail over the few. To facilitate this, political institutions must centralize and concentrate, rather than disperse, power (within the context of competitive elections), e.g., strong and centralized parties, a unitary constitution, plurality electoral laws (or PR with high statutory thresholds), and so forth.¹⁷

The *consensual* (aka pluralist) principle of democracy favors multiple veto-points and the inclusion of as many political perspectives as possible, rather than decisionmaking by narrow majorities or pluralities. Consensual democracy therefore emphasizes proportional electoral laws, large party systems, supermajority cabinets, supermajority decision rules, the separation of executive and legislative powers, a federal constitution, and other institutions that require the national head of government to share power with other political actors and bodies.¹⁸

The *participatory* principle of democracy is usually viewed as a lineal descendant of the “direct” (i.e., non-representative) model of democracy, as derived from the experience of Athens—though elements of this model may also be discerned in “republican” thought and in the experience of many small communities throughout the world and throughout human history (Pocock 1975). The motivation for participatory democracy is uneasiness about delegating authority to representatives. Direct rule by citizens is preferred, wherever practicable. And within the context of representative government, the participatory component is regarded as the most democratic element of the polity. This model of democracy thus highlights the importance of voting, but also of citizen assemblies, party primaries, referenda, juries, social movements, public hearings, town hall meetings, and other forums of citizen

¹⁶ See Dahl (1956 on “Madisonian Democracy”; see also Gordon (1999), Hamilton, Madison and Jay (1992), Hayek (1960), Held (2006, ch. 3), Hirst (1989), Mill (1958), Vile (1998).

¹⁷ See American Political Science Association (1950), Bagehot (1963), Ford (1967), Goodnow (1900), Lijphart (1999), Lowell 1889), Ranney (1962), Schattschneider (1942), Wilson (1956).

¹⁸ See Lijphart (1999), Mansbridge (1983).

engagement.¹⁹

The *deliberative* principle of democracy focuses on the process by which decisions are reached in a polity. A deliberative process is one in which public reasoning focused on the common good motivates political decisions—as contrasted with emotional appeals, solidary attachments, parochial interests, or coercion. According to this principle, democracy requires more than a mindless aggregation of existing preferences; there should be respectful dialogue at all levels—from preference formation to final decision—among informed and competent participants who are open to persuasion (Dryzek 2010: 1). “The key objective,” writes David Held (2006: 237), “is the transformation of private preferences via a process of deliberation into positions that can withstand public scrutiny and test.” Some political institutions serve a specifically deliberative function, such as consultative bodies (hearings, panels, assemblies, courts); polities with these sorts of institutions might be judged more deliberative than those without them. However, the more important issue is the degree of deliberativeness that can be discerned across all powerful institutions in a polity (not just those explicitly designed to serve a deliberative function) and among the citizenry.²⁰

The *egalitarian* principle of democracy stresses that formal political rights and civil liberties are insufficient for political equality. The polity should also address material and immaterial inequalities that inhibit the actual exercise of these rights and liberties. Ideally, groups – as defined by income, wealth, education, ethnicity, religion, caste, race, language, region, gender, sexual identity, or other ascriptive characteristics – should have approximately equal participation, representation, agenda-setting power, protection under the law, and influence over policymaking and policy implementation. If such equality does not already exist, the egalitarian principle requires state efforts to make the distribution of socio-economic resources, education, and health more equal so as to enhance political equality. (This principle does not entail equality of power between leaders and citizens, as leaders in all polities are by definition more powerful.)²¹

Naturally, the conceptual scheme presented in Table 1 does not capture all the theoretical distinctions at play in the complex concept of democracy. It does not fully capture the distinction between direct and representative democracy, or among different principles of representation (Pitkin 1967). It does not capture the intellectual history or etymology of the concept, though traces of that history will be glimpsed in each principle (Dunn 1995; Held 2006;

¹⁹ See Barber (1988), Benelo and Roussopoulos (1971), Dewey (1916), Fung and Wright (2003), Macpherson (1977), Mansbridge (1983), Pateman (1976), Rousseau (1984), Young (2000).

²⁰ See Bohman (1998), Cohen (1989), Elster (1998), Fishkin (1991), Fung (2005), Gutmann and Thompson (1996), Habermas (1984, 1996), Held (2006, ch. 9), Offe (1997). A number of recent studies have attempted to grapple with this concept empirically; see Bächtiger (2005), Dryzek (2009), Mutz (2008), Ryfe (2005), Steiner et al. (2004), Thompson (2008).

²¹ See Ake (1999), Berman (2006), Bernstein (1961, 1996), Dahl (1982, 1989), Dewey (1916, 1930), Dworkin (1987, 2000), Gould (1988), Lindblom (1977), Meyer (2007), Offe (1982), Sen (1999), Walzer and Miller (1995). Many of the writings cited previously under participatory democracy might also be cited here. Taking a somewhat different stand on this issue, Beetham (1999) and Saward (1998: 94-101) do not request an equal distribution of resources. Rather, they consider access to basic necessities in the form of health care, education, and social security to be democratic rights as they make participation in the political process possible and meaningful.

Shapiro and Hacker-Cordon 1999). All typologies are limited in some respects. Nevertheless, the typology captures a good deal of the action surrounding current debates on democracy. Each principle is logically distinct and—at least for some theorists—independently valuable. For example, some writers believe that enhanced avenues for participation are good for democracy even in the absence of liberal aspects of democracy.

Moreover, we suspect that there is a good deal of divergence across these seven principles among the world's polities. Some countries will be particularly strong on the electoral principle; others will be strong on the egalitarian principle, and so forth. Thus, the typology provided in Table 1 is likely to prove a useful empirical device, allowing one to chart variation in political institutions through time and space.

Table 1:
Principles of Democracy

I. Electoral (*Aka* elite, minimal, realist, Schumpeterian)

Ideals: contestation, competition.

Question: are important government offices filled by free and fair multiparty elections?

Institutions: elections, political parties, competitiveness, turnover.

II. Liberal

Ideals: limited government, horizontal accountability, individual rights, civil liberties, transparency.

Question: is power constrained and individual rights guaranteed?

Institutions: independent media, interest groups, and judiciary; written constitution with explicit guarantees.

III. Majoritarian (*Aka* responsible party government)

Ideals: majority rule, centralization, vertical accountability.

Question: does the majority (or plurality) rule?

Institutions: consolidated and centralized, with special focus on the role of political parties.

IV. Consensual (*Aka* pluralist)

Ideal: power sharing, multiple veto-points

Question: how numerous, independent, and diverse are the groups and institutions that participate in policymaking?

Institutions: multiparty system, proportional electoral laws, supermajorities, oversized cabinets, federalism

V. Participatory

Ideal: government by the people.

Question: do ordinary citizens participate in politics?

Institutions: election law, civil society, local government, direct democracy.

VI. Deliberative

Ideal: government by reason.

Question: are political decisions the product of public deliberation?

Institutions: media, hearings, panels, other deliberative bodies.

VII. Egalitarian (*Aka* social)

Ideal: political equality.

Question: are all citizens equally empowered?

Institutions: socio-economic and political factors that generate conditions for political equality.

The typology constitutes one step in the direction of disaggregation. A second step is the

identification of meso-level *components*. These specify different aspects of the seven principles, and are too numerous to list here (see *Taxonomy*). They also exist at multiple levels; i.e., a single component may have several sub-components.

The final step in disaggregation is the identification of *indicators*, the empirical referents of components. In identifying indicators we look for features (a) that are related to at least one principle of democracy; (b) that bring the political process into closer alignment with the core meaning of democracy (rule by the people); and (c) that are measurable—directly or indirectly (via coder judgments) across polities and through time.

Indicators take the form of nominal data (classifications, text, dates) and ordinal or interval scales. Some refer to *de jure* aspects of a polity – that which is stipulated by statute or constitutional law (including the unwritten constitution of places like the United Kingdom and well-established principles of constitutional law, as developed through a common law system). Others refer to *de facto* aspects of a polity – that which is true in practice. (Frequently, we suspect that the latter is at variance with the former. We are inclined to prefer *de facto* indicators, but we collect *de jure* indicators as well so that they can be compared.)

There are approximately 400 indicators in the V-Dem questionnaire. Each question, along with possible response-types, is listed in the *Codebook*. More will be said about coding procedures in the next section.

To summarize, the V-Dem conceptual scheme has several levels of specificity:

Core concept (1)
 Principles (7)
 Components (##)
 Sub-Components (##)
 Indicators (~400)

A *Taxonomy* (not yet constructed) contains strictly hierarchical depictions of the conceptual scheme for each principle. This kind of organization is useful for enumerating all the indicators by classifying them under the most relevant sub-component, component, and principle. Of course, a strict taxonomy cannot capture the conceptual schemes fully because some indicators are relevant for measuring multiple sub-components, some sub-components are relevant for multiple components, and some components are relevant for multiple principles. For example, the sub-component “executive selection and removal by election” is relevant for both electoral democracy and liberal democracy; civil liberties are relevant for both participatory and egalitarian democracy; and several components—judicial independence, subnational elections, number of parties in the cabinet, and so on—are relevant for both majoritarian and consensual democracy, albeit in opposite ways. These more complex relationships are best depicted with interactive software. A preliminary version of one such visualization is available at the V-Dem site (<http://www.v-dem.net>). The reader is encouraged to browse this visualization to see the relationships that we envision across principles, components, and indicators.

Clarifications

Several important clarifications apply to the V-Dem conceptual scheme.

First, our attempt to operationalize democracy through principles, components, and indicators does not attempt to incorporate the *causes* of democracy (except insofar as some elements of our far-flung concept might affect other elements). For example, regime-types may be affected by economic development (Epstein et al. 2006), by colonial experiences (Bernhard et al. 2004), or by attitudes and political cultures (Almond, Verba 1963/1989; Hadenius and Teorell 2005b; Welzel 2007). However, these elements are not regarded as *constitutive* of democracy. That is, we do not regard a polity as more democratic than another because it is richer, has a British colonial heritage, or has a culture that seems to support democratic practices. (In the latter case, it is the fact of exhibiting those practices, not the fact of having a culture that supports them, that makes a polity more or less democratic.)

Second, our quest to conceptualize and measure democracy should not be confused with the quest to conceptualize and measure *governance*.²² Of course, many elements of democracy might also be considered elements of good governance. But others probably have little relevance for the quality of governance and some may have a negative impact. Thus, to say that an indicator or component provides a measure of democracy is not to say that it advances the cause of good governance – or, for that matter, of justice or the good (Arneson 2004). The relationship of democracy to other desired outcomes is an empirical matter, one which we anticipate V-Dem will be able to shed light on (see section VI.) However, measures of good governance – unless they happen to dovetail with measures of democracy – must be gathered from other sources.

Third, we recognize that some indicators and components (listed in the *Codebook* and the *Taxonomy*) are more important in guaranteeing a polity's overall level of democracy than others. This of course depends upon one's model of democracy. The point is that inclusion in V-Dem does not presuppose a judgment of relative importance. All it means is that a particular component or indicator is relevant to the operationalization of at least one principle of democracy.

Fourth, principles, components, and indicators of democracy sometimes conflict with one another. At the level of principles, there is an obvious conflict between the majoritarian and consensus principles, which adopt contrary perspectives on most institutional components: fragmented power satisfies the consensus ideal but inhibits the possibility of majority rule. One can easily perceive conflicts across other principles as well. Similar conflicts are in evidence at lower levels of aggregation. For example, protection of individual liberties can impose limits on the will of the majority; and the existence of strong civil society organizations can have the effect of pressuring government to restrict the civil liberties enjoyed by marginal groups (Isaac *n.d.*). Furthermore, the same institution may be differently viewed according to different

²² Rose-Ackerman (1999), Thomas (2010). Inglehart and Welzel (2005) argue that *effective* democracy – as opposed to purely formal or institutional democracy – is linked to state capacity. A formally democratic country that has no capacity, or is ridden with corruption, is not democratic in the full sense of the term. In order to represent this very thick concept of democracy they multiply the Freedom House indices by indices of corruption (drawn from Transparency International or the World Bank), producing an index of effective democracy. See Knutsen (2010) for critical discussion.

conceptions of democracy. For example, the common practice of mandatory voting is clearly offensive to the liberal model (where individual rights are sacrosanct and include the right not to vote), but is vindicated by the participatory model (since it has a demonstrated effect in boosting turnout where sanctions are more than nominal).

Such contradictions are implicit in democracy's multidimensional character. No wide-ranging empirical investigation can avoid conflicts among democracy's diverse attributes. However, with separate indicators representing these different facets of democracy it will be possible to examine potential tradeoffs empirically—an important dilemma for policymakers and academics to grapple with.

Fifth, our proposed set of principles, components, and indicators, while fairly comprehensive, is by no means entirely comprehensive. The protean nature of *democracy* resists closure; there are always potentially new principles/components/indicators that, from one perspective or another, may be associated with this essentially contested term. Moreover, some conceptions of democracy are difficult to capture empirically, and virtually impossible to track over time and across countries on a global scale. This limits the scope of any empirical endeavor.

Finally, it is important to acknowledge that principles and higher-order components, while much easier to define than *democracy*, are still resistant to authoritative conceptualization. Our objective in Table 1 and in the Taxonomy is to identify the most essential and distinctive attributes associated with these concepts. Even so, we are keenly aware that others might make different choices, and that different choices will be required for different tasks. The purpose of the typology and the taxonomies is guidance, not legislation. It demonstrates how the various elements of V-Dem hang together, according to a particular set of inter-relationships. We expect other writers will assemble and dis-assemble these parts in whatever fashion suits their needs and objectives. In this respect, V-Dem has the modular quality of a Lego set.

In particular, some end-users of V-Dem will require more aggregated concepts than we have provided here. They may wish, for example, to assemble a maximal (ideal-type) definition of electoral democracy, including many more components than we have included in our spare (minimal) definition.²³ Defining a principle in a maximal fashion generates a concept that may be regarded as a holistic conception of democracy; that is, it may plausibly be viewed as a definition of democracy at-large. End-users may also wish to combine different principles into an overall conception of democracy – e.g., *electoral + liberal*. Again, we think that the modular quality of V-Dem offers a flexible tool that should be suitable for many purposes.

²³ See Gerring (2012b: ch 5) on minimal and maximal definitions and Coppedge (1999) on thick and thin definitions.

III. Data Collection, Aggregation, Validation

The viability of any dataset hinges critically on its method of data collection. V-Dem aims to achieve transparency, precision, and realistic estimates of uncertainty with respect to each data point. In these respects, as in others, we believe that V-Dem compares favorably with extant democracy indicators. In order to assess these issues, however, it is important to lay out our approach in some detail.

History of Politics

Our principal concern is with the operation of political institutions within large and fairly well-defined political units which enjoy a modicum of sovereignty or which serve as operational units of governance (e.g., colonies of overseas empires). We shall refer to these units as *polities* or *countries*.²⁴

We are not concerned merely with the present and recent past of these *polities*. In our view, understanding the present – not to mention the future – requires a rigorous analysis of history. The regimes that exist today, and those that will emerge tomorrow, are the product of complex processes that unfold over decades. Although regime changes are sometimes sudden, like earthquakes, these dramatic events are perhaps best understood as a combination of pent-up forces that build up over long spans of time, not simply the precipitating factors that release them. Democratization trends, causes, effects, and sequences cannot be understood without historical data.²⁵

The advantage of our topic (in contrast with other historical measurement tasks such as national income accounts) is that much of the evidence needed to code features of democracy is well preserved in books, articles, newspapers archives, and in living memory. Democracy is, after all, a high-profile phenomenon. Although a secretive regime may hide the true value of goods and services in the country, it cannot disguise the existence of an election. And those features of an election that might prejudice the outcome towards the incumbent are difficult to obscure completely. Note that virtually everyone living in that country, studying that country, or covering that country for some foreign news organization or aid organization has an interest in tracking this result.

Thus, we regard the goal of historical data-gathering as essential and also realistic, even if it cannot be implemented for every indicator in the *Codebook*. Some historical indicators are better than none at all. Furthermore, if V-Dem can demonstrate empirically that the kind of

²⁴ We are not measuring democracy within very small communities (e.g., neighborhoods, school boards, municipalities, corporations), in contexts where the political community is vaguely defined (e.g., transnational movements), or on a global level (e.g., the United Nations). This is not to say that the concept of democracy should be restricted to formal and well-defined *polities*. It is simply to clarify our approach, and to acknowledge that different strategies of conceptualization and measurement may be required for different subject areas.

²⁵ This echoes a persistent theme presented in Capoccia and Ziblatt (2010), Teorell (2011), and in other historically grounded work (Nunn 2009; Mahoney and Rueschemeyer 2003; Pierson 2004; Steinmo, Thelen, and Longstreth 1992).

indicators that can be coded for the past are highly correlated with indicators that can be coded for the present, proxy historical indicators can be constructed. Finally, V-Dem's pilot phase proved that it is possible to extend our detailed, disaggregated indicators far back in modern history, even for little-studied countries (e.g., Suriname, Myanmar, Albania) and even during years prior to formal independence (i.e., under colonial rule).

V-Dem therefore aims to gather data whenever possible back to 1900 for all territories that can claim a sovereign or semi-sovereign existence (they enjoyed a degree of autonomy at least with respect to domestic affairs) and served as the operational unit of governance (separable from other territories). The latter criterion means that they are governed differently from other territories and we might reasonably expect many of our indicators to vary across these units. Thus, in identifying political units we look for those that have the highest levels of autonomy and/or are operational units of governance. These sorts of units are referred to as "countries," even if they are not fully sovereign. This means that Eritrea is coded as an Italian colony (1900-41), a province of Italian East Africa (1936-41), a British holding administered under the terms of a UN mandate (1941-51), a federation with Ethiopia (1952-62), a territory within Ethiopia (1962-93), and an independent state (1993-). Sometimes, historical polities are coded separately because their existence is not tied clearly to a contemporary nation-state. Thus, we code the Ottoman Empire and the Federation of Rhodesia and Nyasaland (aka Central African Federation) as separate countries.

Whenever a contemporary country, *A*, was once a part of another country (empire or nation-state), *B*, that was ruled in a fairly uniform fashion, the coding for *B* is applied to *A* in order to create continuous historical coding for *A*. Thus, former members of Yugoslavia (Serbia et al.) receive the coding for Yugoslavia from 1912 to 1991. In this fashion, a panel dataset is created that is mostly "balanced" from 1900 to the present. Of course, it is not entirely balanced: some countries cannot be coded back to 1900 because their status in a previous era is divided or unclear. For further details, see *Countries*.

V-Dem provides time-series codings that reflect historical changes as precisely as possible. The coding procedure allows coders to specify the exact dates (day/month/year) corresponding to changes in an institution, rather than lumping all events together within a given year. Likewise, elections are coded as events happening on a specific date rather than as an indefinite period of time. Naturally, date-specific data can be aggregated at 12-month intervals, which may be essential for time-series where country/years form the relevant units of analysis.

Coding Types

The ~400 indicators listed in the *Codebook* fall into three main types. Type (A) data is gathered from extant sources (other datasets or secondary sources) and requires no original coding. The collection of this data is supervised by the Principal Investigators and Project Managers and carried out by RAs connected to the project. Sources are listed in the *Codebook* (and will eventually be integrated into the on-line database). This sort of coding comprises 113 indicators, roughly one-third of the total. The *Codebook* also lists 31 additional ("A+") indicators, whose collection is contingent upon additional funding.

Type (B) data is gathered from country-specific sources and does not require coding decisions, being factual in nature. This data is gathered by Country Research Assistants under the supervision of Regional Managers. This sort of coding comprises 65 indicators, roughly 1/6 of the total.

Type (C) data requires some degree of judgment about the state of affairs obtaining in a particular country at a particular point in time. This genre of indicator is coded by Country Experts – generally academics or policymakers who are nationals of and/or residents in a country, with deep knowledge of that country and of a specific substantive area. This sort of coding comprises 175 indicators, about half of the total.

Type (D) data is created from (A), (B), or (C) coding. This includes “cumulative” indicators such as “number of elections since 1900.” It also includes more aggregated variables such as components and principles, as described above. There is no limit, in principle, to the number of higher-order variables that might be generated as part of the V-Dem project but Appendix A lists 12 such indicators at present.

Expert Coding

Type (C) coding – by Country Experts – is the most difficult, since it involves judgment on the part of the coder. Accordingly, a number of steps are taken to minimize error and to gauge the degree of imprecision that remains.²⁶

We seek a minimum of five Country Experts to code each country-year for every indicator. The 175 C-indicators are organized into twelve sections and four clusters, as follows:

1. Elections
 - Political parties/electoral systems
 - Direct democracy
2. Executive
 - Legislature
 - Deliberation
3. Judiciary
 - Civil liberty
 - Sovereignty
4. Civil society organizations
 - Media
 - Political equality

We suggest (but do not require) that each Country Expert code at least one cluster. It is left to the Regional Managers in consultation with the coders, to decide who is most appropriate for which sections of the questionnaire and how many sections a single coder will be assigned. Note that if five Country Experts coded all four indicators for each country, V-Dem would use about 1,000 coders, in total, to collect the data on C-indicators across the world. However,

²⁶ For a perceptive discussion of the role of judgment in coding see Schedler (2012a).

most experts possess the requisite expertise to code only one or two of the clusters. This means that, in practice, each country is typically coded by 15 to 25 Country Experts.²⁷ This in turn, means that V-Dem will likely engage the services of 2,000 to 5,000 Country Experts.

Informed by our experience with the pilot study, the following procedure is used to identify and select Country Experts. First, we identify a list of potential coders for a country (typically 60-100 names per country). This list is compiled by Regional Managers in consultation with Country Research Assistants, based on their intimate knowledge of a country. Central RAs (located at Gothenburg or Notre Dame) also contribute to this list, using information about potential country experts gathered from the web. Other members of the project team provide additional names if and when they have country-specific expertise.

Regional Managers and Country RAs thus play a critical role in the data collection process. V-Dem's approach is to recruit Regional Managers who are nationals of, and residents in, one of the countries in each region whenever possible. Country RAs are almost always nationals or residents in the country to be coded. This is critical to our effort to identify and recruit Country Experts who have expertise in various sections of the survey.

For each potential coder/Country Expert we compile basic information – country of origin, current location, highest educational degree, current position, and area of expertise. We also take note of any possible biases that might affect their ability to code questions in a dispassionate manner.

The selection and enrollment of Country Experts from the initial list is handled by Regional Managers in consultation with the Principal Investigators and the Country Research Assistant. Five criteria are applied to this choice.

The most important selection criterion, naturally, is expertise in the section of the survey they are assigned to code. This is usually signified by an advanced degree in the social sciences, law, or history; a record of publications; and positions in civil society that establish their expertise in the chosen area (e.g. a well-known and respected journalist). Naturally, potential coders are drawn to areas of the survey that they are most familiar with, and are unlikely to agree to code topics they know little about. So, self-selection works to achieve our primary goal of matching questions in the survey with country-specific expertise.

The second criterion is origin in the country to be coded. V-Dem's goal is that a minimum of three out of five (60%) Country Experts should be nationals and/or permanent residents of the country they code (preferably both). Exceptions may be required for a few countries where it is difficult to find in-country coders who are both qualified and dispassionate. This criterion should help avoid Western/Northern biases in the coding, which can also come from self-selection biases in who makes the migration to Western/Northern universities.

The third criterion is the prospective coder's seriousness of purpose. By this, we mean a person's willingness to devote time to the project, to deliberate carefully over the questions asked in the survey, and to report their honest judgment. Sometimes, personal

²⁷ Note that in some rare cases, small and under-studied countries (e.g., Yemen) it is necessary to ask individual experts to code the whole set of surveys, simply because experts on the various parts of the survey are not available.

acquaintanceship is enough to convince a Regional Manager that a person is fit, or unfit, for the job. Sometimes, this feature becomes apparent in communications that precede the offer to work on V-Dem.²⁸

The fourth criterion is impartiality. V-Dem aims to recruit coders who will answer questions in the survey in an impartial manner. This means avoiding those who might be beholden to powerful actors – by reason of coercive threats or material incentives – or who serve as spokespersons for a political party or ideological tendency (in some instances, such as North Korea, this may entail avoiding all in-country coders). Where this is difficult, or where the reality is difficult to determine, we aim to include a variety of coders who, collectively, represent an array of views and political perspectives on the country in question.

The final criterion is obtaining diversity in professional background among the coders chosen for a particular country. This entails a mixture of professionals who are active in political affairs (e.g., in the media or civil society organizations) along with academics who study these topics. It also means finding experts who are located at a variety of institutions, universities and research institutes.

After weighing these five criteria, an initial cast of (typically) 15-25 Country Experts are contacted. If the quota of five Country Experts per section of the survey for each country is not filled, we work down the list of potential Country Experts until the quota is obtained. Those who fail to begin or complete the survey in a reasonable time are replaced by others, following the same procedure. Coders receive a modest honorarium for their work, in accordance with the number of questions or surveys they have answered.

A number of steps are taken to assure informed consent and confidentiality among participants. The on-line survey provides full information about the project (including this document) and the use of the data, so that coders are fully informed. It also requires prospective coders to choose one of three levels of confidentiality and certify that they accept the terms of the agreement. They can access the surveys only with a username and password that we assign. Their data is stored on a firewall-protected server managed by the University of Notre Dame. Any data released to the public that includes information about experts excludes information that might be used to identify coders.

Of course, finding the right coders is only the first step toward removing bias from a survey. The more important feature of a survey is the construction of the survey itself. In crafting indicators, we have sought to construct questions whose meaning is clear and specific and not open to a wide variety of interpretations. They should mean the same thing (more or less) in each context and not suffer from temporal or spatial non-equivalence. The V-Dem pilot test (described below) served as an initial test of our questionnaire, prompting quite a few revisions in the next round of surveys. The revised questions for C-coding then went through

²⁸ This last problem is not a significant threat to coding validity in the pilot study or the first phase of the main data collection. Few individuals seems to have been motivated to conduct this arduous coding assignment for purely monetary reasons. The reason for this is that V-Dem pays very little relative to what highly qualified experts could earn for the same amount of work from other pursuits. Further strengthening this point, there seems to be no relationship between the wealth of the country and our ability to recruit coders: we have faced challenges getting experts to agree to conduct coding for the poorest as well as the richest countries in the world.

several rounds of review with the Project Managers and outside experts before emerging in the current and final form.

Importantly, experts are required to report a level of confidence for each coding, an indication of their subjective level of uncertainty. This is scored on a scale from 0 to 100. A second opportunity for registering uncertainty is in the “Remarks” fields that lie at the end of each section of the survey. Here, experts can comment (in prose) on any aspect of the indicators or codings that s/he found problematic or difficult to interpret.

Once a survey (i.e., a section of the questionnaire filled out by a coder) has been completed, the data is analyzed using an algorithm designed to raise flags for dubious patterns. Our system also generates graphs of each coder’s ratings and the results are distributed to, and reviewed by the Regional Manager and by RAs at project headquarters. In rare circumstances, where a particular coder’s responses are grossly out of line with other coders’ responses to a question, or where the responses show evidence of inattention (e.g., where the same coding is repeated for a century, even in the face of obvious changes in the country’s history), that coder is contacted and asked to reconsider. However, for the most part, decisions by coders are respected. We fully expect that for difficult questions – those for which a factual basis for coding is not available – a range of responses is possible. It is desirable, under the circumstances, for the coding results to reflect this uncertainty.

Coding Diagnostics

After coding is complete, a number of diagnostic tests will be conducted in order to gauge whether there might be some form of systematic bias in the coding. Simple tests are conducted by regressing responses for each question against various characteristics of the coder (collected in a post-survey questionnaire [see *Codebook*]) along with country and year dummies, then observing whether any of these characteristics are significant predictors.

A similar test is conducted at higher levels of aggregation by including the coder characteristics as predictors in more fully specified multilevel measurement models for various components. This procedure is repeated across all questions of type (C). Relevant background features include: (1) sex, (2) age, (3) highest degree received, (4) location of graduate work (Anglo-Europe or not), (5) place of birth (in-country or abroad), (6) years spent in country, (7) years spent away from country, (8) location of employer (in country or not), (9) employer (a 9-part classification, tested as a series of dummies), (10) overall estimate of level of democracy in country that has been coded, (11) overall estimate of level of democracy in selected countries around the world (understood as separate predictors), (12) stance towards free markets, and (13) preferred models of democracy. (For further details on these questions see the *Codebook*.)

Several of these questions deserve further explication. #10 derives from a question that asks coders to assign a point score on a 0-100 scale summarizing the level of democracy in the country they are coding about which country in the world the coder perceives as most democratic, overall (using whatever understanding of democracy they themselves prefer). This will be transformed into standard deviations from the mean (Z scores) for a given country and then combined into a common variable – differentiating coders who feel more optimistic about democracy in their country from those who feel more pessimistic and allowing us to test

whether these preconceptions have any influence over their responses.

#11 is similar, except that here we select prominent countries from around the world that embody varying characteristics of democracy/autocracy and ask coders to provide an estimate of the overall level of democracy in these countries (on the same 100-point scale). Chosen exemplars include Costa Rica, Cuba, India, Nigeria, North Korea, Russia, Saudi Arabia, South Africa, Sweden, Switzerland, the United Kingdom, the United States, and Venezuela. Answers to each question constitute a separate interval variable, generating twelve variables.

#12 gauges the coder's views of the relationship between markets and democracy. "It is sometimes argued that free markets and strong protection for property rights are essential for the establishment of democracy." The coder is then asked whether s/he agrees or disagrees with this statement, along a Likert scale.

#13 asks explicitly about the coder's views of democracy. We list and briefly explain the seven principles of democracy outlined in a previous section of this document – electoral, liberal, majoritarian, consensus, participatory, deliberative, egalitarian – asking coders to rate their degree of support for that principle on a Likert scale.

Another sort of diagnostic test concerns possible biases associated with a specific country, rather than with specific coders. It could be, for example, that coders for one country have higher standards, or different standards, than coders for another country. Perhaps in some countries intellectuals are generally critical of the regime while in other countries intellectuals are generally supportive – attitudes that have little to do with the actual quality of democracy in these countries. If so, this might infect V-Dem coding with systematic bias. (Although most questions in the V-Dem questionnaire are highly specific and thus less liable to judgments about the overall "gestalt" of a regime, a handful of questions are more abstract and open-ended.)

In order to determine whether this might be the case we plan to conduct *lateral* coding in the first V-Dem update. (An update, as explained below, will cover the most recent year as well as several years previous. Thus, our first update might include the years 2008-13.) This means that, in addition to our usual complement of country-specific coders, we will enlist additional experts to code a portion of the survey for *all* countries across a region. Thus, we might ask South American legal scholars to code the judicial questions on the V-Dem survey for all countries in South America (for the years 2008-13). These scholars are likely to be familiar with practices across neighboring countries and are therefore well-positioned to make comparative judgments, e.g., is the judiciary in Brazil more independent from political pressures than the judiciary in Uruguay? In identifying regions we will be sure to include "overlap" countries so that we have a basis for comparison across regions. For example, an overlap country in the Americas might be Mexico. So long as coders for each region of the Americas include Mexico in their coding we will have a metric by which to establish equivalence across regions.

Combining Expert Ratings

Having five coders for each type (C) question is immensely useful, as it allows us to identify wayward coders as well as to conduct inter-coder reliability tests. (As noted, these sorts of tests

– standard in most social science studies – are rarely if ever employed in extant democracy indices.) However, we also need to combine these ratings in order to provide a single “best estimate” for each question.

To combine expert ratings for a particular country/indicator we employ a well understood and robust technique, the linear opinion pool (O’Hagan et al. 2006, 181-184). This procedure gives greater weight to scores about which the experts feel more confident. For interval-level data, this would be a weighted average based on the experts’ prior beliefs that their ratings are correct. Most of our indicators, however, are ordinal or nominal, so combinations are a bit more complicated. We calculate the priors for each scale level surrounding an expert’s rating, based on the confidence level the expert assigned to his or her preferred rating. We then calculate a weighted sum of these priors for all the experts who answered the question for that country-date. The combined rating is the one with the highest total probability. Naturally, we also report what that probability is – which may be useful for subsequent analyses.

The priors are calculated from the experts’ confidence levels and the number of levels in each variable. For a two-level scale, zero confidence corresponds to 50-50 probabilities; and to a similarly flat distribution for scales with multiple levels. For nominal items, the priors for non-selected scores are distributed evenly among all the alternatives; for ordinal items, they favor the nearest alternatives over those that are farther from the preferred score. Table 2 illustrates the procedure. Although much of the math is hidden in this table, one can appreciate that preferred scores have the highest prior for each expert, adjacent scores have priors that diminish with distance, high-confidence ratings have greater weight, and the summary score (2) is the one with the greatest sum of weighted priors, which sum to 1.0. The two shaded cells contain the information that would be reported for this observation in the dataset. We will also analyze inter-coder reliability using more familiar statistics such as Cronbach’s alpha and other versions of the intra-class correlation coefficient (ICC).

Table 2:
Combining Expert Scores for an Ordinal Indicator

Expert	Rating (1-3)	Confidence	<i>Weighted Priors</i>		
			P(x=1)	P(x=2)	P(x=3)
1	1	50	0.14	0.05	0.02
2	2	90	0.01	0.36	0.01
3	2	95	0.01	0.39	0.01
Total			0.16	0.80	0.04

Our data will be the first democracy data (and among the first in national-level comparative social science) to include score-specific reliability estimates. Reliability estimates can be used in several ways. They will enable users to select only the most reliable data. They can be used to weight analyses to correct for heteroscedasticity when our indicators are the

dependent variables and reliability is correlated with a regressor. They will also guide the project team as we improve the data over time, indicating which survey items need modification or coding by additional experts. And they can be used to improve index construction, as discussed below.

Aggregation Techniques

The conceptual scheme presented in the *Taxonomy* lays out inter-relationships among principles, components, and indicators at the level of membership – what belongs with what. We must also consider the problem of weighting, i.e., how to assign weights to the various elements of an index (a principle or component). This will be referred to as a problem of aggregation.

As part of this project, we plan to experiment with multiple aggregation schemes for each principle and component. Ultimately, we may decide that one or several of these aggregation techniques are superior, in which case this will be presented in the final version of V-Dem. This is a matter that can only be determined at a later stage of the project when data collection is substantially complete. At present, it is sufficient to indicate the various approaches that will be implemented and assessed. Importantly, each approach will be accompanied by an estimate of aggregation *uncertainty*, that is, the extent to which the aggregate score is sensitive to the choice of aggregation technique.

A first step is to decide whether an indicator can be treated as a *cause* indicator or an *effect* indicator (i.e., an intrinsic or defining attribute). As an example, consider the concept of judicial independence. Institutions such as life tenure, fixed budgets, and a consensual judicial appointment process might be regarded as “causes” of judicial independence. By contrast, purges of the judiciary, court packing, and other attacks by the government on the judiciary are direct manifestations of (the absence of) judicial independence, and as such may be considered “effects.” This distinction matters because effect indicators, but not cause indicators, are expected to be intercorrelated. This provides an empirical tool for validating aggregation schemes.

A second step is deciding upon the relationships among the various indicators that comprise a higher-order concept (a component or principle). If indicators are regarded as (binary) sufficient conditions then the aggregation rule is to take the maximum: a score of 1 for any single indicator will suffice for membership in the concept by reason of substitutability (Goertz 2006). If indicators are regarded as (binary) necessary conditions, by contrast, the aggregation rule is to take the minimum – equivalent to multiplying across a set of binary indicators. A third option establishes other kinds of interactive relationships among indicators. A fourth option is to rank-order all (binary) indicators according to their centrality to a concept, from most central to most peripheral, generating a lexical scale.²⁹ A fifth option regards each

²⁹ A score on this cumulative (ordinal) scale is achieved by counting up a country’s score on each indicator, starting from the most central and proceeding to the most peripheral, but arresting the process whenever a 0 is encountered. Thus, more peripheral attributes of a concept do not count in its aggregation unless more central prerequisites are satisfied. The mathematical formula for a cumulative scale with three binary (0,1) indicators (A, B, and C, with A most central and C least central) is $A + AB + ABC$. See Gerring, Skaaning (2012).

indicator as independent and equally important for the definition of a (non-binary) higher-order concept, in which case a simple averaging across standardized indicators may be adopted as an aggregation rule. A sixth option is to use factor or principal components analysis to assign factor weights based on correlations between each indicator and a presumed latent variable.

A seventh technique is based on item response theory (IRT) models, which are a statistical approach for aggregating discrete outcomes into continuous underlying scales in a logically coherent way. IRT produces latent scales by estimating how a series of observed variables interrelate across a group of cases – here, countries and years. The technique estimates both a series of thresholds or “difficulty” parameters and a set of “discrimination” parameters for each item. Cases that cross only “easy” thresholds are scored low on the scale and those that can cross “difficult” thresholds are scored high. The “discrimination” of each item indicates how informative that item is for placing cases along the latent scale. If an item has low discrimination (that is, close to zero), then knowledge of that item is only minimally informative for determining a case’s location on the scale. Since many of our ratings are recorded as ordinal variables with more than two possible outcomes, we will use a modification of the basic dichotomous-outcome IRT model, known as a graded-response model (e.g., Treier and Jackman 2008, Pemstein et al 2010). This aggregation technique has the virtue of generating scores with meaningful interpretations. For example, a low score on a civil liberties scale might mean that a country provided freedom from torture and servitude but not any of the harder-to-achieve civil liberties; a relatively high score mean that it provided all of them except the most difficult, access to justice.

Latent variable measurement models will allow us to specify each indicator as either a cause or an effect of latent components, and the components as reflections of more general latent principles (Armstrong 2011; Bollen 1993, Treier and Jackman 2008, Pemstein et al 2010). Specifically, we will use a new technique, Bayesian latent-variable modeling with Monte Carlo Markov Chain (MCMC) estimation, which was developed by Drew Linzer for a judicial independence application with Jeffrey Staton (Linzer & Lewis 2011; Linzer and Staton 2011). This technique is especially useful for constructing scales from time-series, cross-sectional indicators that are bounded at the extremes, have different numbers of ordinal thresholds, and are missing observations. A variant of this kind of measurement model could use the completely disaggregated, expert-level data in some of these models, which would enable us to model and correct for sources of expert bias. Our online questionnaire includes a short battery of questions about aspects of the experts’ background that we suspect may affect their judgment: gender; age; education; type of employer; and country of citizenship, residence, education, and employment.

Latent-variable models can also be used to answer additional questions such as the following: How reliable are the country experts, and why? How much weight should each indicator be given when measuring each component, and how much weight should each component have when measuring each principle? Are ratings more reliable for recent years than for earlier years? How much more reliable are ratings of well-known countries compared to obscure countries? Are simpler indices good approximations of the most rigorous ones? Which components of democracy are measured by extant indices, and how well? And perhaps most importantly, how far can the number of dimensions of democracy be reduced without

sacrificing information about theoretically important differences, what is the nature of the dimensions that make both conceptual and empirical sense, and how have these dimensions evolved over the past century?

Naturally, the options reviewed here may sometimes be employed in combination with each other. For example, one might argue that the freeness/fairness of elections matters little if suffrage is severely limited. If so, then the latter must be treated as a necessary condition of the former – though other indicators would come into play in judging the relative freeness and fairness of an election with mass suffrage.

Aggregating the data in different ways will reveal whether the more mathematically sophisticated methods are necessary for producing valid, reliable, and meaningful indices, as opposed to simpler additive indices that would be easier for most users to understand. We consider this an exciting new frontier for research in conceptualization and measurement that has not been feasible to implement until now, given the limitation so of extant democracy indices.

Validation

A project of this magnitude requires validation. To be sure, an informal process of validation will undoubtedly ensue once the project is complete. End-users will write reviews in academic journals and – we hope – will also contribute to the wiki blogs that we plan to develop for the V-Dem web site. Here, we want to provide the outlines of a more formal test, one that could easily be applied prior to the project's completion (given sufficient funding).

V-Dem indicators (that are not purely factual in nature) are subject to an internal process of verification by virtue of the fact that they are coded independently by multiple country experts. Inter-coder reliability tests, along with self-reported uncertainty, may therefore be viewed as partial validation of the results – allowing us to arrive at a measure of certainty for each point estimate. Naturally, one could also replicate the coding of indicators with a new set of coders. It is unlikely that this procedure would yield substantially different outcomes, but we plan to conduct such replications during future updates (by re-coding several prior years with each update).

Another use of inter-coder reliability is to compare reliability across different levels of aggregation. Although most questions asked of country experts are highly specific, a few are more general. For example, we ask coders to assess the degree to which a given election was free and fair, taking all matters into consideration. In addition, one question (in the post-survey questionnaire) asks coders to assign a point score (on a scale from 0-100) summarizing a country's overall level of democracy. This is asked of the coder's country of expertise as well as a handful of well-known countries around the world (the United States, Russia, India,...). This is the most aggregated question in the survey and is included solely for diagnostic purposes. Accordingly, we can distinguish at least three levels of aggregation among the questions asked of country experts. This allows us to test a fundamental measurement hypothesis in this study, namely that questions at a lower level of aggregation will achieve a higher level of inter-coder agreement than questions at a higher level of aggregation. If we find this to be the case, a key design element of the study is validated.

Remaining validation exercises focus at the meso-level (components). Note that components serve as the conceptual building blocks for all V-Dem principles; so in testing components we are also offering a meso-level test of the principles. One approach to validity testing compares V-Dem components with other indices that purport to measure similar concepts. For example, the correlation between a preliminary V-Dem index of civil liberties and the Freedom House index of civil liberties is -.87 (a negative sign because in Freedom House data, higher scores mean “less free”). Additional tests will be conducted with other indices that correspond (more or less) to components identified in the Taxonomy, e.g., the World Justice Project’s Rule of Law Index (Agrasti et al. 2011).

However, techniques of convergent validity are limited in their utility for two reasons. First, we have serious doubts about the validity of many standard indices, for reasons discussed in Part I. Second, standard indices tend to hover at a higher level of aggregation, thus impairing comparability between V-Dem components and non-V-Dem indices. Indeed, only a few extant indices are close enough in conception and construction to provide an opportunity for direct corroboration with V-Dem components. Consequently, we must look elsewhere for validation of the V-Dem project.

We are contemplating an approach modeled loosely on Bowman, Lehoucq, and Mahoney (2005). In that study, several political scientists with deep knowledge of Central American politics examined codings for three indices of democracy – Gasiorowski (1996), Polity IV (Marshall & Jaggers, 2002), and Vanhanen (2000) – as they apply to countries in Central America over the past century. They find multiple examples of idiosyncratic coding for these cases, i.e., codings that have little or no face validity. It is certainly feasible to extend this system of expert validation to V-Dem components by asking a neutral body unconnected with V-Dem to conduct such a validation test based on a handful of randomly drawn countries. A pool of external experts could then be asked to validate V-Dem component scores for a single country over time, and in comparison with other countries with which the expert is familiar.

IV. Project Management

What follows is a brief overview of the main components of V-Dem – the project team, the advisory board, the web site, sustainability into the future, and outreach to the international community.

Project Team

A complex undertaking such as this requires coordinating the work of many individuals, each of which brings a unique skill or knowledge base to the project. A full list with details on the project team can be found at the website www.v-dem.net.

Principal Investigators (PIs) – Coppedge, Gerring, and Lindberg – manage the project team and oversee all activities. This is the highest decision-making body and functions by

consensus. The PIs began work on this project in 2008.

Project Managers (PMs) assume responsibility for at least one category of indicators (sovereignty, voting and representation, elections, the executive, the legislature, the judiciary, political parties and the party system, the media, civil society associations, direct democracy, subnational politics, civil liberty, political equality, deliberation, and measurement issues). This division of labor enabled us to define concepts, identify existing data sources, and write survey questions that reflect the wide-ranging work and accumulated wisdom on these topics. It also allowed for a decentralized system of informal consultation among experts (academics and practitioners) in these areas. Project managers are also responsible for assuring cross-country equivalence for their indicators, so that a given question is not interpreted differently in varying country contexts. Most of the PMs have been with the project since 2009.

Regional Managers (RMs) are responsible for recruiting and managing expert coders, overseeing coding, and checking for problems of equivalence at the regional level. Occasionally, someone is engaged to manage only a single country, in which case we refer to him/her as a *Country Manager*. The first RMs were recruited in Fall 2011.

Country Experts (CEs) code type (C) indicators, as discussed above. They are generally holders of a PhD and usually residents or citizens of the country they are coding. They also have specialized knowledge in at least one of the sections of the V-Dem survey. The CEs have been recruited on a rolling basis since the start of the pilot study in Spring 2011.

Research Assistants (RAs) assist at all levels except expert coding. They are generally pursuing graduate studies, and will work under the supervision of members of the project team. *Central RAs* are located at the University of Gothenburg and the Kellogg Institute at University of Notre Dame (the two headquarters). Central RAs not only assist with the collection of A-type data but also works with project management, quality control, web maintenance, communication, and outreach. A *Country RA (CRA)* is selected by the Regional Manager for each country to gather the type (B) data for that country. *Regional RAs* work under the supervision of the Regional Managers.

Advisory Board

V-Dem's advisory board includes a mix of academics, jurists, and practitioners from around the world. Board members provide guidance on all aspects of the project – though the board is not responsible for the content or conduct of V-Dem. While communication is normally conducted through email and Skype, the board will meet periodically in the coming years to facilitate in-depth discussion.

V-Dem Institute, Database, Web Site

The V-Dem program is hosted in a shared fashion between the Department of Political Science at University of Gothenburg (Sweden) and the Kellogg Institute at the University of Notre Dame (USA). The V-Dem Institute is being established at the Department of Political Science, University of Gothenburg in order to coordinate research, data collection, the working papers series, annual conferences, country- and thematic reports, guest scholarship and internship

programs, and outreach & dissemination of the project. The Kellogg Institute hosts the website, the database, and also helps coordinate data collection and research.

The V-Dem website can be viewed at www.V-Dem.net. The survey interface and underlying relational database has been designed following our guidance, by the Center for Research Computing (CRC) at the University of Notre Dame. The site includes web-based interfaces that allow users to download pilot data and to visualize these data in line graphs or histograms. It is currently being expanded to allow visitors to select variables, countries, and dates for download and to visualize data in scatterplots and maps.

All data is entered on-line (exceptions are made for those Country Experts who might not have internet access). All data (except confidential information about the coders) will be made public when coding for all countries is complete, though data for selected countries may be released prior to the completion of the project at the request of funders. Before the release of the comprehensive dataset, we anticipate an embargo period of up to 12 months to allow the core researchers of the project team prioritized access.

After coding is complete, we plan to establish an open and transparent system of display and dissemination. This might include a blog or Wiki-style format in which interested individuals could comment on the scores assigned to a country. This commentary might take the form of additional information—perhaps unknown to the coders—that speaks to the viability of the coding. Or it might take the form of extended discussions about how a particular question applies to the circumstances of that country, providing further information upon which to base estimates of uncertainty. It will also help to refine the questionnaire in future iterations of the project.

The V-Dem website will also provide additional services besides download of the data and online analysis tools. V-Dem indicators, coding schemes, aggregation formulas, as well as podcasts of selected V-Dem talks, a V-Dem Working Paper series with emerging results, and possibly a series of thematic and country reports/briefs on democracy are elements that will be featured in years to come.

Sustainability

V-Dem must be continually revised and updated in order to remain useful for scholars, practitioners, and citizens. Users will be encouraged to contribute suggestions for revision, as discussed above, and these suggestions will be systematically reviewed. The set of indicators comprising V-Dem may be expanded, and some indicators may require revision. In the latter case, revisions will need to be implemented systematically (for all countries and years in the database) so that equivalence across countries and through time is maintained.

Most important, V-Dem will need to be iterated on an annual or semiannual basis so that the database is up to date. We may use these updates as opportunities to re-code prior years (e.g., the 5 years prior to the update), offering a further check on inter-coder (or intra-coder) reliability. In any case, we would employ the same coders whenever possible, further enhancing the efficiency of each update (since coders would be already familiar with the database and the on-line coding system).

International Collaboration and Outreach

The current project is international in character and composition to a much greater extent than any other democracy index. V-Dem has dual headquarters on either side of the Atlantic – at the Kellogg Institute at the University of Notre Dame (USA), and at the V-Dem Institute in the Department of Political Science at the University of Gothenburg (Sweden). The project team – consisting of PIs, PMs, RMs, and Central & Country RAs – includes nationals of virtually every country in the world. The advisory board includes nationals of Brazil, Burkina Faso, Canada, Ghana, Hong Kong, India, Indonesia, Mexico, Mongolia, the Netherlands, Pakistan, Russia, South Korea, Turkey, the United States, and Yemen. Altogether, members of the project team and the advisory board connect the V-Dem project to several dozen universities located in every region of the world. Most important, V-Dem enlists a total of 2,000 to 5,000 experts from every country in the world to serve as coders for our questionnaire.

We anticipate that the direct involvement of academics and professionals in V-Dem will result in numerous projects, many of which are collaborative in nature. Indeed, many of these collaborative projects have already begun. In future decades, we anticipate that the V-Dem database and web site will become a focal point for policymakers, practitioners, students, and scholars across the social sciences who are interested in democracy.

To facilitate this, we plan to create a visiting scholars-program for selected individuals who have served as country experts, regional managers, or advisors in the project so that they may visit the V-Dem Institute at the University of Gothenburg for a period of a month or so. During this time, they would be expected to complete research on a theme related to the project.

The V-Dem Institute plans to hold annual conferences at University of Gothenburg, as well as smaller workshops, as means to further stimulate innovative collaborations and as additional venues for generating new ideas and publicizing the results of research conducted with V-Dem data.

We also plan to introduce the project at social science conferences throughout the world. This includes the International Political Science Association meetings and meetings of regional political science associations (Africa, East Asia, Southeast Asia, Europe, Latin America, the United States). Project team members and members of the advisory board who live or work in each region will assist in that region's conferences.

Publications in high-profile academic journals and press outlets of a more popular nature will ensure that all relevant communities of interest are informed of the availability of V-Dem and how it can be used to help them do their work better.

We are also establishing a working paper series, to be posted on the Varieties of Democracy web site. All data will be a public good but we will also ask users of the data to contribute initial drafts of their work to the V-Dem Working Paper series. We anticipate that, once publicly available, V-Dem data will generate an increasing number of working papers, journal articles, chapters, and books by scholars around the world, contributing to a growing stream of interrelated research.

Traffic to the web site, as channeled by Google and other search engines and links on related sites, offers a further channel for outreach – making V-Dem accessible to anyone who

searches on key terms (e.g., “democracy”). Funding permitting, we plan to commission the creation of a V-Dem App for smart devices that would enable users to check scores for selected countries, variables, and years, and even to generate color-coded graphs and maps on demand. The V-Dem App will also provide access to working papers, discussion for a, and podcasts of selected V-Dem talks.

V. Work Plan and Funding

The progress of V-Dem may be summarized in three phases.

The First Phase and the Pilot Study

First, there is the conception and management of the project itself, including the pilot study. This phase is substantially complete.

With the support of Sweden’s Ministry of Foreign Affairs, the Riksbankens Jubileumsfond (Sweden), the Quality of Government Institute at the University of Gothenburg, and the Kellogg Institute and other bodies at the University of Notre Dame, we conducted a pilot study in twelve countries from May 2011 to January 2012. These included two countries from each of six regions: Mexico and Suriname in the Americas, Japan and Burma in Asia, Russia and Albania in the post-Communist region, South Africa and Ghana in Africa, Egypt and Yemen (North and South) in the Middle East, and Sweden and Switzerland in Europe. These pairs were chosen to represent countries in each region that, in our judgment, would be fairly easy, and fairly hard, to code. (For example, in the Americas we judged that Mexico would be fairly easy and Suriname fairly difficult.)

The pilot study progressed rapidly. Ninety experts submitted 461,000 country-date-question-expert observation ratings to the database in just seven months, including all 12 countries and all survey questions. Preliminary analyses indicate that the data already collected are of good quality. For example, we ask all experts not only to choose a response to each question but also to estimate their confidence in all their responses. The average summary certainty level in the pilot data is 84 percent, which means that based on the agreement among experts and their confidence in their ratings, we can be, on average, 84 percent confident that the summary scores are correct. Cronbach’s alpha averages .709 (conventionally considered in the “acceptable” range) for the 131 pilot-study type (C) questions for which we have enough data to calculate it. We analyzed the pilot study data to pinpoint the weakest questions and modify or drop them, and used this information to revise the questionnaire extensively in order to improve reliability before the current phase of data collection. In addition, very preliminary analyses have shown that aggregated scores discriminate well not only across countries and over time, but also across components. For example, a country can do well with respect to civil society but less well on administration of elections or legislative strength. Such distinctions are among the most important contributions of our approach.

The Second (Current) Phase

The second phase of V-Dem, comprising data collection for the entire world from 1900 to 2012, began in March 2012. We aim to conclude data collection in 2014. This of course is contingent upon obtaining necessary funding. At present, we have secured funding for data collection in roughly 120 of the 206 countries/territories we plan to code.

Concurrently, the PIs are working on issues of aggregation – how to build components from indicators and broader indices from components. This stage also involves cleaning the data, preparing it for public access (through the V-Dem web site).

The Third Phase

Beginning in 2014 (if all goes well), V-Dem will enter a third phase where three somewhat separate activities will run simultaneously.

- A) *Updates*. Annual or biannual updates of the data, building on our established network of RMs, RM-RAs, CRAs, and CEs;
- B) *Outreach & Dissemination*. This involves the V-Dem institute reaching out to potential end-users (including the community of democracy practitioners) with information, training, workshops, and conferences possibly involving producing country- and thematic reports for the policy community; and
- C) *The V-Dem Research Program*. The goal is to establish a coherent research program that addresses issues of conceptualization and aggregation, sequences of democratization, and the causes and effects of democracy. This research program would include scholars from the project team as well as other collaborators, and would culminate in reports written for publication in scholarly journals and in more popular venues.

Funding

The principal funding for the pilot phase of the program came from the Ministry of Foreign Affairs-Sweden. The V-Dem pilot phase was co-funded by also by Riksbankens Jubileumsfond, and the Quality of Government Institute. The hosting universities, the Kellogg Institute of Notre Dame and the University of Gothenburg, have provided substantial seed funds for the pilot phase and both institutions have committed to additional co-funding for the present and the coming years.

The Second Phase is currently half-funded. The Research Council of Norway (through a collaborative arrangement with Håvard Hegre, Oslo University) is currently financing data collection for 17 countries; the Canadian Development Agency is financing another 10 countries at present; an additional 40 countries is expected to be funded through a partnership with International IDEA and with funding from Ministry of Foreign Affairs-Sweden; the Ministry of Foreign Affairs/Denmark is contributing to data collection in another set of 14 countries, and finally the European Commission/DEVCO is preparing a grant to fund the data collection in 30-40 countries. Thus we are in the process of having secured funding for the data collection in over half of the countries in the world already.

In addition, we are applying for funding from United States National Science Foundation, the Danish Council for Independent Research, and discussions have begun with other potential funders in Europe and North America including the emerging European Endowment for Democracy (EED), Norway's Ministry for Foreign Affairs, and the United States Agency for International Development (USAID). Through the partnership with IDEA on data collection and outreach, we hope to secure funding for an additional set of countries.

For the Third Phase, applications with the Riksbankens Jubileumsfond for data collection (USD 1.3mn) and the V-Dem Research Program (USD4.5mn), are pending and will be decided in the Fall 2012.

VI. Payoffs

Having described V-Dem in some detail, the payoffs anticipated from our project may now be briefly reviewed. We begin with some general discussion, summarizing many of the points made in previous sections of this project description. We proceed to a more specialized topic, the possible role of V-Dem indicators in the impact evaluation of democracy assistance programs. We then approach the sociology of democracy measurement – why a project of V-Dem's scope has never before been attempted. Finally, we lay out some benchmarks by which the impact of the V-Dem project on the development community (encompassing both policymakers and academics) can be monitored in the coming years – V-Dem's monitoring and evaluation (M&E) plan.

General Points

We begin with the aggregation problem. Many problems of conceptualization and measurement stem from the decision to represent democracy as a single point score (based on a binary, ordinal, or interval index). Of course, summary measures have their uses. Sometimes one wants to know whether a country is democratic or non-democratic or how democratic it is overall (Bayer, Bernhard 2010). V-Dem will help with this task insofar as it offers the base-level indicators that any index requires. No longer will it be necessary for new indices to start from scratch, collecting data for all countries and all time-periods.

Even so, the goal of summarizing a country's regime type is elusive. As we have seen, extant democracy indices suffer from serious problems of conceptualization and measurement. While many new indices have been proposed over the past several decades—all purporting to provide a single point score that accurately reflects countries' regime status—none has been successful in arriving at an authoritative and precise measurement of this challenging concept.

In our view, the traditional approach falls short because its self-assigned task is impossible. The highly abstract and contested nature of democracy impedes effective operationalization. This is not a problem that can be solved—at least not in a conclusive fashion. Naturally, one can always impose a particular definition upon the concept, insist that this is democracy, and then go forward with the task of measurement. But this is unlikely to

convince anyone not already predisposed to the author's point of view. Moreover, even if one could gain agreement over the definition and measurement of democracy, an important question remains about how much useful information about the world this highly aggregated concept would provide.

A more productive approach to this topic is to recognize the multiple principles of democracy and, within each principle, to disaggregate. At lower levels of abstraction the concept becomes more tractable. While the world may never agree on whether the overall level of democracy in India is summarizable as a "4" or a "5" (on some imagined scale), we may agree on scores for this and other countries at the level of principles, components, and indicators.

The importance of creating consensus on these matters can hardly be over-emphasized. The purpose of a set of democracy indicators is not simply to guide rich-world policymaking bodies such as USAID, the World Bank, and the UNDP. As soon as a set of indicators becomes established and begins to influence international policymakers, it also becomes fodder for dispute in other countries around the world. A useful set of indicators is one that claims the widest legitimacy. A poor set of indicators is one that is perceived as a tool of western influence or a mask for the forces of globalization (as Freedom House is sometimes regarded). The hope is that by reducing the elements of democracy down to levels that are more coherent and operational it may be possible to generate a broader consensus around this vexed subject. Moreover, insofar as indicators might be used to construct aggregate indices, at least everyone would know what lies behind a country's score in a particular year since the underlying indicators used to compose an index would be available for inspection. (Presumably, the aggregation principles would be available as well.)

Another advantage is the degree of precision and differentiation that a disaggregated set of indicators offers relative to extant composite indices. While holistic measures of democracy float hazily over the surface of politics, the principles, components, and indicators of a disaggregated dataset are comparatively specific and precise. Contrasts and comparisons become correspondingly acute.

Recall that the data we are collecting is extremely disaggregated, down to the level of country-date-variable-expert observations. When completed, it will become one of the largest country-level databases in existence. Note that there are approximately 400 indicators, of which 175 (C-type indicators) will be coded by 5 experts per item. In addition, a series of perhaps 20-40 composite indices (D-type variables) will be generated from this underlying data. Each of these 500 or so variables cover 206 countries/territories, most of which are coded over the course of 112 years. The result will be a dataset of about 22 million data points. For anyone interested in issues relating to democracy, the data in this disaggregated form will be a vast resource for exploration and testing. It will allow policymakers and researchers to clarify how, specifically, one country's democratic features differ from others in the same region, or across regions, and how a country's features have changed over time.

This detailed knowledge of cases is especially helpful in the context of country assessments. How can policymakers determine which aspects of a polity are most in need of assistance? While Freedom House and Polity offer only a few dimensions of analysis (and these are highly correlated and difficult to distinguish conceptually), V-Dem envisions numerous

parameters. It seems clear that for assessing the potential impact of programs focused on different elements of a polity it is helpful to have indicators at hand that offer a differentiated view of the subject. Intuitively, the greatest effectiveness is achieved when program interventions are targeted on the weakest element of democracy in a country. A large set of fully differentiated indicators would make it possible to both identify those elements and test the assumption behind such choices.

Relatedly, V-Dem will allow policymakers and researchers to track a single country's progress or regress through time. One will be able to specify which facets of a polity have improved, and which have remained stagnant or declined. This means that the longstanding question of regime transitions would be amenable to empirical tests. When a country transitions from autocracy to democracy (or vice versa), which elements come first? Are there common patterns, a finite set of sequences, prerequisites? Or, is every transition unique? Do transition patterns effect the consolidation of democracy? With a large set of indicators measured over many years, it would become possible for the first time to explore transition sequences.³⁰ Does a newly vibrant civil society lead to more competitive elections, or to an authoritarian backlash? Do accountable elected officials create an independent judiciary, or does an independent judiciary make officials accountable? Similar questions could be asked about the relationships among citizenship, voting, parties, civil society, and other components of democracy, perhaps with the assistance of sequence-based econometrics (Abbott 1995; Abbott and Tsay 2000; Wu 2000).

Note that insofar as one wishes to judge trends, trend lines are necessary. A single snapshot of the contemporary world reveals nothing about the direction or speed at which countries are moving toward, or away from, democracy. Even trends in a short span of recent years can be very misleading, as many democratization paths contain many years of stasis punctuated by sudden movements toward or away from democracy. Assessments of global trends require even more data, as some countries move in opposite directions in any given year; "waves" of democratization exist only on average, with many exceptions (Huntington 1991).

One would also be able to test democracy's causal effect as an independent variable (Carbone 2009). Does democracy hinder economic growth, contain inflation, promote public order, or ensure international peace? Answering such questions requires a lengthy time-series because the effects of these factors play out over many years. They also require a great deal of disaggregation because we need to know, as specifically as possible, which elements of democracy are related to which results. This is helpful from the policy perspective as well as from the analytic perspective, so that we can gain insight into causal mechanisms. Whether democracy is looked upon as an independent (causal) variable, or a dependent (outcome) variable, we need to know which aspect of this complex construct is at play.

Recent work has raised the possibility that democracy's effects are long-term, rather than (or in addition to) short-term (Gerring et al. 2005). It seems quite probable that the short-term and long-term effects of democracy are quite different. Plausibly, long-term effects are

³⁰ Sequencing is explored by Schneider and Schmitter (2004b) with a smaller set of indicators and a shorter stretch of time. See also McFaul (2005).

more consistent, and more positive along various developmental outcomes, than short term effects. Consideration of these questions demands a historical coding of the key variables.

A final set of advantages of our proposed approach to democracy derives from its data-collection strategy. This is open to public scrutiny and commentary. Procedures of collection and aggregation are replicable. Experts with deep knowledge of each country's history are to be enlisted. Several experts are asked to code each indicator wherever the latter requires a degree of judgment. Inter-coder reliability tests, along with other gauges of reliability, are systematic. Periodic revisions are planned rather than ad hoc, and where re-coding is required it is applied to the whole dataset so that consistency is preserved through time and across polities.

With democracy indicators, as with most things, the devil is in the details. Getting these details right, and providing full transparency at each stage, should enhance the precision and validity of the measurement instrument, as well as its legitimacy in the eyes of potential end-users.

Impact Evaluation

Policymakers are increasingly concerned with evaluating the impact of the programs that they fund – both individually and at an aggregate level. Does democracy promotion work (in the sense of enhancing prospects for democracy)? Does it lead to other undesirable outcomes such as greater instability?³¹ These are difficult, but nonetheless fundamental, issues. Thus, we want to consider carefully to what extent V-Dem indicators might contribute to knowledge about program evaluation.

Let us begin at the program level. Imagine a program focused on improving capacity among civil society groups in a country. Ideally, an element of randomization is involved so as to allow for an experimental or quasi-experimental design. Regardless of research design, the *proximal* impact of this program is best measured with very specific indicators – presumably focused on the groups that are being offered assistance (or might be offered assistance) – collected for this purpose. V-Dem has little to offer at the micro-level.

Suppose that the funders of this program also wish to know whether this program – or some set of programs – have impacts at a societal level, i.e., across civil society at-large. Here, V-Dem indicators focused on civil society may prove extremely useful. It will certainly be more useful than highly aggregated extant indices, which do not afford the opportunity to examine how particular aspects of democracy may have changed in response to an intervention.

Of course, the funder could also decide to collect their own (custom-designed) societal-level indicators for this purpose. However, there are several reasons to suppose that V-Dem might offer an advantage over custom-made macro-indicators (at the country level). First, this would significantly increase the cost of program evaluation. Second, one faces the problem of bias since those who construct and measure the outcomes of concern typically have a stake in

³¹ For various sides of this issue see Bollen, Paxton, Morishima (2005), Burnell (2008), Carothers (1999, 2006), CEUDAP (2008), Chua (2003), Cox, Ikenberry, Inoguchi (2000), Crawford (2001), Epstein, Serafino, Miko (2007), Finkel, Perez-Linan, Seligson (2007), Goldsmith (2001), Green, Kohl (2007), Guilhot (2005), Hearn (2000), McMahon (2001), Youngs (2010).

the outcome. (Usually, this imposes a “positive” bias on the analysis.) Third, in order to discern subtle changes in a trend-line that might be a product of a specific intervention it is usually necessary to construct a long time-series, extending back well before the initiation of the program. This may be difficult to do, or at least costly. Fourth, a custom set of indicators is likely to be restricted to a single country, prohibiting comparisons with neighboring countries. These comparisons often help in detecting spurious causal relationships. So, all things considered, we can identify several advantages to using V-Dem to measure country-level outcomes. The only serious disadvantage would be a situation in which a relevant outcome-measure is not included in the panoply of V-Dem indicators. Here, a monitoring and evaluation plan might supplement V-Dem with a few custom-designed indicators.

Now let us consider macro-level interventions. By this, we mean the entire set of democracy promotion efforts undertaken by a large agency such as USAID, the World Bank, the UNDP, or a group of agencies the democracy assistance programs funded by the DAC countries. Our interest might be country-specific (how have DAC democracy promotion efforts affected the quality of democracy in Haiti?) or crossnational (how have DAC democracy promotion efforts affected the quality of democracy across the developing world?). At this level, we believe that V-Dem offers a significant advance over extant democracy indices (Polity, Freedom House, et al.). First, V-Dem allows one to investigate the impact of policies on specific indicators and components of democracy (rather than looking only at democracy at-large). It could be, for example, that democracy promotion is more successful in bolstering the quality of elections than in enhancing the quality of civil society. If so, this would provide important information for funders to consider. Likewise, if funding categories are disaggregated – so that programs focused on elections can be distinguished from programs focused on civil society – then these specific connections can be probed. This is the approach taken by a recent and highly influential study of USAID democracy promotion (Finkel, Perez-Linan, and Seligson 2007). A shortcoming of this study is that the outcomes are often not well-calibrated to the inputs, a product of the paucity of disaggregated indicators of democracy when Finkel et al. conducted their analysis.

The business of inferring causal relationships between (non-randomized) inputs and outputs measured at the country level is fraught. However, policymakers must consult the available evidence about what works, and what doesn’t, even if that evidence does not approach the experimental ideal. They cannot resort to hunches. While randomization can assist in measuring proximal outcomes at the program level this sort of research design is not viable at the macro level and with respect to distal outcomes. Here, we must enlist observational data. And here, V-Dem offers a significant improvement over extant indicators and therefore may assist policymakers in evaluating and improving democracy assistance programs.

One further point may be added. One of the problems of evaluating democracy promotion efforts is determining whether a program was actually responsible for observed improvements when improvements may have happened anyway, due to other conditions. In democracy promotion, the bias is toward positive findings because of the general trend toward improvement in democracy throughout the world at the present time. Negative findings are less likely to crop up for the wrong reasons and are therefore more trustworthy. Positive findings are even more likely when using highly aggregated indices of democracy such as Polity

or Freedom House, which tend to capture overall tendencies that are likely to be associated with cross-national trends. Such analyses tend to overlook specific dimensions or attributes of a regime that do not change or that move in a negative direction. V-Dem indicators are designed to capture these discrepant tendencies (as well as the main tendencies). This makes V-Dem indicators useful for causal inference in general, and especially for helping policymakers identify situations in which policies are *not* having their intended effect.

The Sociology of Measuring Democracy

V-Dem has many potential uses, as outlined above, and its utility extends equally to policymakers and academics. Both of these constituencies are engaged in a common enterprise, an enterprise that has been impeded by the lack of a sufficiently discriminating instrument.

In this context, one might ponder the sort of problems that would arise for macroeconomists, finance ministers, and policymakers at the World Bank and IMF if they possessed only one highly aggregated indicator of economic performance: an index of “prosperity,” for example. As good as GDP is (and there are of course difficulties), it neither measures the whole concept of prosperity nor says anything specific. It would not go very far without the existence of additional variables that measure the components of this macro-level concept. This is in fact where the economics discipline stood 80 years ago: aside from census data, which generated some information about income and employment once a decade, and some trade figures (because trade was taxed), there were only “boutique” economic indicators compiled by lone academic economists for specific research projects. This situation changed with the Great Depression, when Simon Kuznets was appointed to set up an office in the United States Department of Commerce that would collect national accounts data in a systematic fashion (Carson 1975).

At the current time, it has been estimated that the United States government collects data on 45,000 economic indicators, with private data providers adding perhaps another four million indicators tracking various aspects of the American economy (Silver 2012: 185). Of course, other countries’ economies are not so minutely observed so it is not possible to make systematic comparisons across all of these indicators. Even so, the options for comparative analysis of macroeconomic indicators across the OECD and beyond are remarkable.

In the field of political analysis, we find ourselves in approximately the position of economists before the Depression. We have some crude sense of how democratic a country is; but we have no systematic knowledge of how a country scores on the various components of democracy, and our historical knowledge is even weaker.

Now let us step back from the details of our project to ask a larger question in the sociology of knowledge. If a new set of democracy indicators promises so many returns one might wonder why it has not already been developed. After all, academics and policy makers have been struggling with issues of conceptualization and measurement for quite some time,

and the problems noted at the outset of this paper are widely acknowledged.³²

It is important to remember that producing a dataset of this immense scope is time-consuming and expensive, requiring the participation and coordination of many researchers. While the downstream benefits are great, no single scholar or group of scholars has the resources or the incentives to invest. Indeed, the academic disciplines do not generally reward members who labor for years to develop new data resources.³³

Moreover, few national or international organizations have the funding and the motivation to collect global and historical data on a highly disaggregated level. While a host of national institutions (e.g., finance, commerce, agriculture, and labor departments, along with national banks) and international financial institutions (e.g., the World Bank, International Monetary Fund, World Trade Organization, Food and Agriculture Organization, and Organization for Economic Cooperation and Development) supervise the collection of economic data, there are no analogous institutions charged with collecting political data.

Thus, the topic of democracy has been left to several meagerly financed organizations who must go begging each year (or every several years) to renew their funding so that coding can be continued.³⁴ Freedom in the World, including the construction of the Political Rights and Civil Liberties indices (along with associated staff, consultants, printing, and promotion costs), is budgeted at \$500–600,000 annually.³⁵ Polity consumes about \$120,000 annually.³⁶ These are paltry sums, especially when compared with the vast sums of money spent by OECD countries on democracy promotion.

A brief review of these expenditures may be helpful in putting our subject into the proper context. According to figures compiled by the Development Assistance Committee, OECD countries spent over \$13 billion (USD) in 2009 for programs to support the quality of democracy and governance in the developing world. This includes contributions from Canada (\$409 million), Denmark (\$213 million), Sweden (\$533 million), the United Kingdom (\$802 million), the United States (\$3.9 billion), the European Union (\$1.6 billion), the United Nations Development Programme (\$296 million), and UNICEF (\$40 million).³⁷

³² For discussion of the difficulties of descriptive inference see Gerring (2012a).

³³ Note that while scholars who are discontented with the leading indicators of democracy periodically re-code countries of special concern to them (e.g., Acuna-Alfaro 2005; Berg-Schlosser 2004a, 2004b; Bowman, Lehoucq, and Mahoney 2005; McHenry 2000), this re-coding is generally limited to a small set of countries or a small period of time. The incentive structure for data collection is discussed compellingly in Herrera and Kapur (2007), and coordination failures in Schedler (2012b).

³⁴ The following statistics are drawn from a variety of sources, as cited. They represent the most recent figures available. Where data for several recent years is available these are averaged together to form an estimate of normal budget activity.

³⁵ Arch Puddington (Director of Research, Freedom House), personal communication (January 2011).

³⁶ Monty Marshall, personal communication (December 2010).

³⁷ Data generated from statistics provided by DAC, downloaded on January 10, 2012 (<http://stats.oecd.org/Index.aspx?DatasetCode=CRSNEW>). The following categories are included in the calculation of D&G spending: 15110 (Public Sector policy and administrative management), 15111 (Public finance management), 15112 (Decentralization and support to subnational government), 15113 (Anti-corruption organizations and institutions), 15130 (Legal and judicial development), 15150 (Democratic participation and civil

Arguably, the formulation, targeting, and overall effectiveness of these programs might be improved if a small fraction of this money were spent on the creation of better indicators of democracy, as outlined here.

In any case, the failure to adequately measure democracy is a product both of paltry resources and poor institutional incentives. Consequently, academics and policymakers have continued to employ—and complain about—Polity, Freedom House, DD, and other highly aggregated indices. It is our hope that users of these indices will recognize the public good aspect of enhanced measures of democracy, and that investments by individual scholars and funding institutions will help make this prospect a reality.

Evaluating V-Dem

In concluding this lengthy project description we need to address a basic question of monitoring and evaluation with respect to the V-Dem project. How will we know whether V-Dem has the payoffs that we claim for it?

This project is not a treatment that can be randomized across units and assessed in an experimental fashion. Nonetheless, it is important to offer some assessment of V-Dem's impact on the communities it is intended to serve – including policymakers, democracy activists, and academics. (Of course, our ultimate community of interest consists of citizens throughout the world, but this distal impact will be impossible to measure in any meaningful fashion.)

In order to do this, we intend to use quantitative indicators like the following:

1. Number of RMs, CRAs, and CEs engaged in and completing their work for V-Dem.
2. Number of data points, country-years, indicators, and indices available online.
3. Number of working papers posted on the V-Dem site.
4. Number of hits on the V-Dem web site, download of data, analyses carried out online.
5. Number of Google hits for V-Dem.
6. Number of Tweets for V-Dem.
7. Number of Facebook entries for V-Dem.
8. Number of citations of V-Dem in international news and policy journals such as *Journal of Democracy*, *Democratization*, *The Economist*, the *New York Times*, the *Financial Times*, *Le Monde*, as determined by searches of on-line indices for each publication.
9. Number of DAC countries that use V-Dem as a criterion for allocation of development assistance, for country assessments, or for M&E, as determined by informal survey.

Most of these indicators can be disaggregated by region, e.g., hits on the V-Dem site (as % of total) originating from East Asia, Southeast Asia, South Asia, MENA, Africa, Europe, North

society), 15151 (Elections), 15152 (Legislatures and political parties), 15153 (Media and free flow of information), 15160 (Human rights), 15170 (Women's equality organizations and institutions). For further information see Cornell (2012; forthcoming).

America, South America, and Pacifica. Many of these indicators can also be applied to Freedom House and Polity, allowing for a direct comparison with the two most widely used indices of democracy.

Each of the seven indicators will be collected annually (at the end of the calendar year), allowing us to track V-Dem usage over time. At the end of a decade (with 2011 understood as the base year), we expect to find increased usage of V-Dem – understood through absolute counts as well as relative to Freedom House and Polity – across all impact indicators.

In addition, a number of qualitative approaches will be enlisted in order to assess V-Dem's contribution to the community of scholars and professionals working on democracy. A page on the web site will be reserved for discussion of the project, allowing for comments from all end-users in a blog format. (The only restriction would be that participants identify themselves and that their comments be germane.) Equally important are formal reviews of the V-Dem project appearing in journals, development blogs, and working papers over the coming years. These will be posted on the V-Dem web site for easy access.

VII. References

- Abbott, Andrew. 1995. "Sequence Analysis: New Methods for Old Ideas." *Annual Review of Sociology* 21: 93–113.
- Abbott, Andrew, and Angela Tsay. 2000. "Sequence Analysis and Optimal Matching Methods in Sociology." *Sociological Methods and Research* 29: 3–33.
- Acuna-Alfaro, Jairo. 2005. "Measuring Democracy in Latin America (1972–2002)." Working Paper No. 5, Committee on Concepts and Methods (C&M) of the International Political Science Association (August).
- Agrast, Mark David; Juan Carlos Botero; Alejandro Ponce. 2011. *Rule of Law Index 2011*. The World Justice Project.
- Ake, Claude. 1999. *The Feasibility of Democracy in Africa*. Dakar: CODRESIA.
- Almond, Gabriel A.; Sidney Verba. 1963/1989. *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Newbury Park, CA: Sage.
- Altman, David, and Aníbal Pérez-Liñán. 2002. "Assessing the Quality of Democracy: Freedom, Competitiveness and Participation in Eighteen Latin American Countries." *Democratization* 9(2): 85–100.
- Alvarez, Michael, Jose A. Cheibub, Fernando Limongi, and Adam Przeworski. 1996. "Classifying Political Regimes." *Studies in Comparative International Development* 31(2): 3–36.
- American Political Science Association. 1950. *Toward a More Responsible Two-Party System*. New York: Holt, Rinehart.
- Arat, Zehra F. 1991. *Democracy and Human Rights in Developing Countries*. Boulder: Lynne Rienner.
- Armstrong II, David A. 2011. "Stability and Change in the Freedom House Political Rights and Civil Liberties Measures." *Journal of Peace Research* 48(5) 653–62.
- Arndt, Christiane, and Charles Oman. 2006. "Uses and Abuses of Governance Indicators." Paris: OECD Development Centre.
- Arneson, Richard J. 2004. "Democracy Is Not Intrinsically Just." In *Justice and Democracy*, ed. Keith Dowding, Robert E. Goodin, and Carole Pateman. Cambridge: Cambridge University Press.
- Bächtiger, André, ed. 2005. Special issue: "Empirical Approaches to Deliberative Democracy." *Acta Politica* 40(1–2).
- Bagehot, Walter. 1963 [1867]. *The English Constitution*. Ithaca: Cornell University Press.
- Barber, Benjamin. 1988. *The Conquest of Politics: Liberal Philosophy in Democratic Times*. Princeton: Princeton University Press.
- Bayer, Resat; Michael Bernhard. 2010. "A Test of the Relative Utility of Scalar and Dichotomous Measures: The Operationalization of Democracy and the Strength of the Democratic Peace." *Conflict Management and Peace Science* 27:1, 85–101.
- Beetham, David. 2004. "Towards a Universal Framework for Democracy Assessment." *Democratization* 11(2):1–17.
- Beetham, David, ed. 1994. *Defining and Measuring Democracy*. London: Sage.
- Beetham, David, Sarah Bracking, Iain Kearton, and Stuart Weir, eds. 2001. *International IDEA Handbook on Democracy Assessment*. The Hague: Kluge Academic Publishers.
- Beetham, David. 1999. *Democracy and Human Rights*. Cambridge: Polity.

- Benelo, George C. and Dimitrios Roussopoulos, eds. 1971. *The Case for Participatory Democracy: Some Prospects for a Radical Society*. New York: Viking Press.
- Berg-Schlosser, Dirk. 2004a. "Indicators of Democracy and Good Governance as Measures of the Quality of Democracy in Africa: A Critical Appraisal." *Acta Politica* 39(3): 248–278.
- Berg-Schlosser, Dirk. 2004b. "The Quality of Democracies in Europe as Measured by Current Indicators of Democratization and Good Governance." *Journal of Communist Studies and Transition Politics* 20(1): 28–55.
- Berman, Sheri. 2006. *Social Democracy and the Making of Europe's Twentieth Century*. Cambridge University Press,
- Bernhard, Michael, Christopher Reenock, and Timothy Nordstrom. 2004. "The Legacy of Western Overseas Colonialism on Democratic Survival." *International Studies Quarterly* 48, 225-50.
- Bernhard, Michael, Timothy Nordstrom, and Christopher Reenock. 2001. "Economic Performance, Institutional Intermediation, and Democratic Survival." *Journal of Politics* 63: 775–803.
- Bernstein, Eduard. 1961. *Evolutionary Socialism: A Criticism and Affirmation*, Random House.
- Bernstein, Eduard. 1996. *Selected Writings of Eduard Bernstein, 1900-1921*, Prometheus Books.
- Bertelsmann Foundation. [various years]. *Bertelsmann Transformation Index [2003–]: Towards Democracy and a Market Economy*. Washington, DC: Brookings Institution Press.
- Besancon, Marie. 2003. "Good Governance Rankings: The Art of Measurement." World Peace Foundation Reports 36. Cambridge, MA: World Peace Foundation.
- Bohman, James. 1998. "Survey Article: The Coming of Age of Deliberative Democracy." *Journal of Political Philosophy* 6(4): 400–425.
- Boix, Carles; Michael K. Miller, and Sebastian Rosato. (forthcoming). "A Complete Dataset of Political Regimes, 1800-2007." *Comparative Political Studies*.
- Boix, Carles, and Susan C. Stokes. 2003. "Endogenous Democratization." *World Politics* 55(4): 517–549.
- Bollen, Kenneth A. 1980. "Issues in the Comparative Measurement of Political Democracy." *American Sociological Review* 45: 370–90.
- Bollen, Kenneth A. 1993. "Liberal Democracy: Validity and Method Factors in Cross-National Measures." *American Journal of Political Science* 37(4): 1207–1230.
- Bollen, Kenneth A. 2001. "Cross-National Indicators of Liberal Democracy, 1950–1990." 2nd ICPSR version. Chapel Hill, NC: University of North Carolina and Ann Arbor, MI: Inter-university Consortium for Political and Social Research.
- Bollen, Kenneth A., and Pamela Paxton. 2000. "Subjective Measures of Liberal Democracy." *Comparative Political Studies* 33(1): 58–86.
- Bollen, Kenneth A., and Robert W. Jackman. 1989. "Democracy, Stability, and Dichotomies." *American Sociological Review* 54(4): 612–621.
- Bollen, Kenneth; Pamel Paxton; Rumi Morishima. 2005. "Assessing International Evaluations: An Example from USAID's Democracy and Governance Program." *American Journal of Evaluation* 26:2 (June) 189-203.
- Bowman, Kirk, Fabrice Lehoucq, and James Mahoney. 2005. "Measuring Political Democracy: Case Expertise, Data Adequacy, and Central America." *Comparative Political Studies* 38(8): 939–970.
- Bühlmann, Marc, Wolfgang Merkel, Lisa Müller, and Bernhard Weßels. 2011. "The Democracy Barometer: A New Instrument to Measure the Quality of Democracy and Its Potential for Comparative Research." *Comparative European Politics*.

- Burnell, Peter. 2008. "From Evaluating Democracy Assistance to Appraising Democracy Promotion." *Political Studies* 56
- Capoccia, Giovanni; Daniel Ziblatt. 2010. "The Historical Turn in Democratization Studies: A New Research Agenda for Europe and Beyond." *Comparative Political Studies* 43 (August/September) 931-68.
- Carbone, Giovanni. 2009. "The Consequences of Democratization." *Journal of Democracy* 20:2 (April) 123-37.
- Carothers, Thomas. 1999. *Aiding Democracy Abroad: The Learning Curve*. Washington, DC: Carnegie Endowment.
- Carothers, Thomas. 2006. "The Backlash Against Democracy Promotion." *Foreign Affairs* 85 (Mar/Apr).
- Carson, Carol S. 1975. "The History of the United States National Income and Product Accounts: The Development of an Analytical Tool." *Journal of Income and Wealth* 21: 153-181.
- Casper, Gretchen, and Claudiu Tufis. 2003. "Correlation Versus Interchangeability: The Limited Robustness of Empirical Findings on Democracy Using Highly Correlated Data Sets." *Political Analysis* 11: 196-203.
- Cheibub, Jose Antonio, Jennifer Gandhi, and James Raymond Vreeland. 2010. "Democracy and Dictatorship Revisited." *Public Choice* 143(1-2): 67-101.
- Chua, A.L. 2003. *The World on Fire: How Exporting Free Market Democracy Breeds Ethnic Hatred and Global Instability*. New York: Anchor Books.
- Cohen, Joshua. 1989. "Deliberation and Democratic Legitimacy." In *The Good Polity*, ed. Alan Hamlin and Phillip Petit. Oxford: Blackwell.
- Collier, David, and Robert Adcock 1999. "Democracy and Dichotomies: A Pragmatic Approach to Choices about Concepts." *Annual Review of Political Science* 2:537-565.
- Collier, David, and Steven Levitsky. 1997. "Democracy with Adjectives: Conceptual Innovation in Comparative Research." *World Politics* 49: 430-451.
- Committee on Evaluation of USAID Democracy Assistance Programs (CEUDAP). 2008. *Improving Democracy Assistance: Building Knowledge Through Evaluations and Research*. National Research Council.
- Coppedge, Michael, and John Gerring with David Altman, Michael Bernhard, Steven Fish, Allen Hicken, Matthew Kroenig, Staffan I. Lindberg, Kelly McMann, Pamela Paxton, Holli A. Semetko, Svend-Erik Skaaning, Jeffrey Staton, and Jan Teorell. 2011. "Conceptualizing and Measuring Democracy: A New Approach." *Perspectives on Politics* 9(1): 247-267.
- Coppedge, Michael, and Wolfgang H. Reinicke. 1990. "Measuring Polyarchy." *Studies in Comparative International Development* 25(1): 51-72.
- Coppedge, Michael, Angel Alvarez, and Claudia Maldonado. 2008. "Two Persistent Dimensions of Democracy: Contestation and Inclusiveness." *Journal of Politics* 70(3): 335-350.
- Coppedge, Michael. 1999. "Thickening Thin Concepts and Theories: Combining Large N and Small in Comparative Politics." *Comparative Politics* 31(4): 465-476.
- Cornell, Agnes. 2012. "Does regime type matter for the impact of democracy aid on democracy?" *Democratization*.
- Cornell, Agnes. forthcoming. *The Role of Foreign Aid in Democratization Processes*. Ph.D. diss. University of Gothenburg.
- Cox, M.; John Ikenberry; T. Inoguchi (eds). 2000. *American Democracy Promotion: Impulses, Strategies and Impacts*. Oxford: Oxford University Press.

- Crawford, Gordon. 2001. *Foreign Aid and Political Reform: A Comparative Analysis of Democracy Assistance and Political Conditionality*. Palgrave Macmillan.
- Dahl, Robert A. 1956. *A Preface to Democratic Theory*. Chicago: University of Chicago Press.
- Dahl, Robert A. 1982. *Dilemmas of Pluralist Democracy*. New Haven: Yale University Press.
- Dahl, Robert A. 1989. *Democracy and its Critics*. New Haven: Yale University Press.
- Dewey, John. 1916. *Democracy and Education*. New York: The Free Press.
- Dewey, John. 1930. *Individualism*. Minton, Balch & Company.
- Diamond, Larry and Leonardo Morlino, eds. 2005. *Assessing the Quality of Democracy*. Baltimore: Johns Hopkins University Press.
- Dryzek, John S. 2010. *Foundations and Frontiers of Deliberative Governance*. Oxford: Oxford University Press.
- Dryzek, John S. 2009. "Democratization as Deliberative Capacity Building." *Comparative Political Studies* 42(11): 1379–1402.
- Dunn, John. 1995. *Democracy: A History*. Boston: Atlantic Monthly Press.
- Dworkin, Ronald. 2000. *Sovereign Virtue: The Theory and Practice of Equality*, Cambridge, MA: Harvard University Press.
- Dworkin, Ronald., 1987. *Law's Empire*. Cambridge: Harvard University Press.
- Economist Intelligence Unit (EIU). 2010. "Democracy Index 2010: Democracy in Retreat." Economist Intelligence Unit. (http://www.eiu.com/public/democracy_index.aspx), accessed January 25, 2011.
- Elkins, Zachary. 2000. "Gradations of Democracy? Empirical Tests of Alternative Conceptualizations." *American Journal of Political Science* 44(2): 287–294.
- Elster, Jon, ed. 1998. *Deliberative Democracy*. Cambridge: Cambridge University Press.
- Epstein, David L.; Robert Bates; Jack Goldstone; Ida Kristensen; Sharyn O'Halloran. 2006. "Democratic Transitions." *American Journal of Political Science* 50:3 (July) 551–69.
- Epstein, S.; N. Serafino; F. Miko. 2007. *Democracy Promotion: Cornerstone of U.S. Foreign Policy?* Washington, DC: Congressional Research Service.
- Finkel, Steve, Anibal Pérez-Liñán, and Mitchell A. Seligson. 2007. "The Effects of U.S. Foreign Assistance on Democracy Building, 1990–2003." *World Politics* 59: 404–439.
- Fishkin, James S. 1991. *Democracy and Deliberation: New Directions for Democratic Reform*. New Haven: Yale University Press.
- Ford, Henry Jones. 1898/1967. *The Rise and Growth of American Politics*. New York: Da Capo.
- Foweraker, Joe, and Roman Krznaric. 2000. "Measuring Liberal Democratic Performance: An Empirical a Conceptual Critique." *Political Studies* 48(4):759–787.
- Freedom House. 2007. "Methodology," Freedom in the World 2007. New York. (http://www.freedomhouse.org/template.cfm?page=351&ana_page=333&year=2007), accessed September 5, 2007.
- Fung, Archon, and Erik Olin Wright, eds. 2003. *Deepening Democracy: Institutional Innovations in Empowered Participatory Governance*. London: Verso.
- Fung, Archon. 2005. "Deliberation before the Revolution: Toward an Ethics of Deliberative Democracy in an Unjust World." *Political Theory* 33(2): 397–419.
- Gallie, W. B. 1956. "Essentially Contested Concepts." *Proceedings of the Aristotelian Society* 56: 167–220.

- Gasiorowski, Mark J. 1996. "An Overview of the Political Regime Change Dataset." *Comparative Political Studies* 29(4): 469–83.
- Gates, Scott, Håvard Hegre, Mark P. Jones, and Håvard Strand. 2006. "Institutional Inconsistency and Political Instability: Polity Duration, 1800–2000." *American Journal of Political Science* 50(4):893–908.
- Gerring, John, Philip Bond, William Barndt, and Carola Moreno. 2005. "Democracy and Growth: A Historical Perspective." *World Politics* 57(3): 323–364.
- Gerring, John; Sven-Erik Skaaning. 2012. "A Cumulative Scale." Unpublished paper, Department of Political Science, Boston University.
- Gerring, John. 2012a. "Mere Description." *British Journal of Political Science* (forthcoming).
- Gerring, John. 2012b. *Social Science Methodology: A Unified Framework*, 2d ed. Cambridge: Cambridge University Press.
- Gleditsch, Kristian S., and Michael D. Ward. 1997. "Double Take: A Re-examination of Democracy and Autocracy in Modern Polities." *Journal of Conflict Resolution* 41: 361–83.
- Goertz, Gary. 2006. *Social Science Concepts: A User's Guide*. Princeton: Princeton University Press.
- Goldsmith, Arthur A. 2001. "Donors, Dictators and Democrats in Africa." *Journal of Modern African Studies* 39:3, 411-36.
- Goodnow, Frank J. 1900. *Politics and Administration*. New York: Macmillan.
- Gordon, Scott. 1999. *Controlling the State: Constitutionalism from Ancient Athens to Today*. Cambridge: Harvard University Press.
- Gould, Carol C. 1988. *Rethinking Democracy: Freedom and Social Cooperation in Politics, Economy, and Society*. Cambridge, U.K.: Cambridge University Press.
- Green, Andrew T.; Kohl, R.D. 2007. "Challenges of Evaluating Democracy Assistance: Perspectives from the Donor Side." *Democratization* 14(1):151-165.
- Guilhot, Nicolas. 2005. *The Democracy Makers: Human Rights and International Order*. New York: Columbia University Press.
- Gutmann, Amy, and Dennis Thompson. 1996. *Democracy and Disagreement*. Cambridge: Harvard University Press.
- Habermas, Jurgen. 1984. *The Theory of Communicative Action. Volume One: Reason and the Rationalization of Society*. Boston: Beacon Press.
- Habermas, Jurgen. 1996. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Cambridge: MIT Press.
- Hadenius, Axel and Jan Teorell. 2005b. "Cultural and Economic Prerequisites of Democracy: Reassessing Recent Evidence." *Studies in Comparative International Development* 39(4): 87–106.
- Hadenius, Axel, and Jan Teorell. 2005a. "Assessing Alternative Indices of Democracy." Committee on Concepts and Methods Working Paper Series, August.
- Hadenius, Axel. 1992. *Democracy and Development*. Cambridge: Cambridge University Press.
- Hamilton, Alexander, James Madison, and John Jay. 1992 [1787–88] . *The Federalist*. Ed. William R. Brock. London: Everyman/J.M. Dent.
- Hayek, Friedrich A. 1960. *The Constitution of Liberty*. London: Routledge & Kegan Paul.
- Hearn, Julie 2000. "Aiding Democracy? Donors and Civil Society in South Africa." *Third World Quarterly* 21, 815-30.
- Held, David. 2006. *Models of Democracy*, 3d ed. Cambridge: Polity Press.

- Herrera, Yoshiko M., and Devesh Kapur. 2007. "Improving Data Quality: Actors, Incentives, and Capabilities." *Political Analysis* 15(4): 365–386.
- Hirst, Paul Q., ed. 1989. *The Pluralist Theory of the State: Selected Writings of G.D.H. Cole, J.N. Figgis, and H.J. Laski*. New York: Routledge.
- Högström, John. 2013. "Does the Choice of Democracy Measure Matter? Comparisons between the Two Leading Democracy Indices, Freedom House and Polity IV." *Government and Opposition* (forthcoming).
- Huntington, Samuel P. 1991. *The Third Wave: Democratization in the Late Twentieth Century*. Norman, OK: University of Oklahoma Press.
- Hyde, Susan; Nikolay Marinov. 2012. "Which Elections Can Be Lost?" *Political Analysis* 20:191-210.
- Inglehart, Ronald and Welzel, Christian (2005). *Modernization, Cultural Change and Democracy: The Human Development Sequence*. Cambridge: Cambridge University Press.
- Isaac, Jeffrey C. n.d. "Thinking About the Quality of Democracy and its Promotion." Unpublished ms.
- Jackman, Simon. 2008. "Measurement." In *The Oxford Handbook of Political Methodology*, ed. Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier. Oxford, UK: Oxford University Press.
- Kelley, Judith. 2009. "The More the Merrier? The Effects of Having Multiple International Election Monitoring Organizations." *Perspectives on Politics* 7(1): 59–64.
- Knutsen, Carl Henrik. 2010. "Measuring Effective Democracy." *International Political Science Review* 31:2, 109-28.
- Krook, Mona Lee. 2009. *Quotas for Women in Politics: Gender and Candidate Selection Reform Worldwide*. Oxford: Oxford University Press.
- Kurtz, Marcus J. and Andrew Schrank. 2007. "Growth and Governance: Models, Measures, and Mechanisms." *Journal of Politics* 69(2): 538–554.
- Landman, Todd, and Edzia Carvalho. 2008. *Issues and Methods in Comparative Politics: An Introduction*, 3rd Edition. London: Routledge.
- Landman, Todd. 2003. "Map-Making and Analysis of the Main International Initiatives on Developing Indicators on Democracy and Good Governance." Unpublished manuscript, University of Essex.
- Lijphart, Arend. 1999. *Patterns of Democracy*. New Haven: Yale University Press.
- Lindberg, Staffan I. (ed.) 2009. *Democratization by Elections: A New Mode of Transition?* Johns Hopkins University Press.
- Lindberg, Staffan I. 2006. *Democracy and Elections in Africa*. Johns Hopkins University Press.
- Lindblom, Charles E. 1977. *Politics and Markets: The World's Political-Economic Systems*. New York: Basic Books.
- Lively, Jack. 1975. *Democracy*. Oxford: Basil Blackwell.
- Linzer, Drew A. and Jeffrey K. Staton. "A Measurement Model for Synthesizing Multiple Comparative Indicators: The Case of Judicial Independence." Paper Prepared for presentation at the 2011 Annual Meeting of the American Political Science Association, September 1-4, Seattle, Washington.
- Linzer, Drew A. and Jeffrey Lewis. 2010. "poLCA: Polytomous Variable Latent Class Analysis." R package version 1.2. <http://userwww.service.emory.edu/~dlinzer/poLCA>.
- Lively, Jack. 1975. *Democracy*. Oxford: Basil Blackwell.
- Lowell, A. Lawrence. 1889. *Essays on Government*. Boston: Houghton Mifflin.
- Macpherson, C.B. 1977. *The Life and Times of Liberal Democracy*. Oxford: Oxford University Press.
- Mahoney, James, and Dietrich Rueschemeyer, eds. 2003. *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press.

- Mansbridge, Jane. 1983. *Beyond Adversarial Democracy*. Chicago: University of Chicago Press.
- Marshall, Monty G. and Keith Jaggers. 2007. "Polity IV Project: Political Regime Characteristics and Transitions, 1800–2006." (<http://www.systemicpeace.org/inscr/p4manualv2006.pdf>), accessed February 1, 2011.
- McFaul, Michael. 2005. "Transitions from Postcommunism." *Journal of Democracy* 16(3): 5–19.
- McHenry, Dean E., Jr. 2000. "Quantitative Measures of Democracy in Africa: An Assessment." *Democratization* 7(2): 168–185.
- McMahon, Edward R. 2001. "Assessing USAID's Assistance for Democratic Development: Is it Quantity versus Quality?" *Evaluation* 7:4, 453–67.
- Meyer, Thomas. 2007. *The Theory of Social Democracy*. Cambridge: Polity Press
- Mill, John Stuart. 1958 [1865]. *Considerations on Representative Government*, 3d ed., ed. Curren V. Shields. Indianapolis: Bobbs-Merrill.
- Moon, Bruce E., Jennifer Harvey Birdsall, Sylvia Ciesluk, Lauren M. Garlett, Joshua H. Hermias, Elizabeth Mendenhall, Patrick D. Schmid, and Wai Hong Wong. 2006. "Voting Counts: Participation in the Measurement of Democracy." *Studies in Comparative International Development* 41(2): 3–32.
- Munck, Gerardo L. 2009. *Measuring Democracy: A Bridge between Scholarship and Politics*. Baltimore: John Hopkins University Press.
- Munck, Gerardo L., and Jay Verkuilen. 2002. "Conceptualizing and Measuring Democracy: Alternative Indices." *Comparative Political Studies* 35(1): 5–34.
- Mutz, Diana C. 2008. "Is Deliberative Democracy a Falsifiable Theory?" *Annual Review of Political Science* 11: 521–538.
- Nielson, Daniel L., Ryan M. Powers, and Michael J. Tierney. 2010. "Broad Trends in Foreign Aid: Insights from PLAID 1.9.1." Presented at the conference: Aid Transparency and Development Finance: Lessons and Insights from AidData. Oxford, March 22–25.
- Nunn, Nathan. 2009. "The Importance of History for Economic Development." *Annual Review of Economics* 1(1): 1–28.
- Offe, Clas. 1997. "Micro-aspects of Democratic Theory: what makes for the deliberative competence of citizens?" in: Hadenius, Axel (ed.). *Democracy's Victory and Crisis*. Cambridge: Cambridge UP: pp.81-104.
- Offe, Claus 1983. "Political Legitimation through Majority Rule?", *Social Research* 50 (4): 709-756.
- O'Hagan, Anthony, Caitlin E. Buck, Alireza Daneshkhah, J. Richard Eiser, Paul H. Garthwaite, David J. Jenkinson, Jeremy E. Oakley and Tim Rakow. 2006. *Uncertain Judgements: Eliciting Experts' Probabilities*. West Sussex: Wiley.
- Pateman, Carole. 1976. *Participation and Democratic Theory*. Cambridge: Cambridge University Press.
- Paxton, Pamela. 2000. "Women's Suffrage in the Measurement of Democracy: Problems of Operationalization." *Studies in Comparative International Development* 35(3): 92–110.
- Pemstein, Daniel, Stephen Meserve, and James Melton. 2010. "Democratic Compromise: A Latent Variable Analysis of Ten Measures of Regime Type." *Political Analysis* 18(4): 426–449.
- Pierson, Paul. 2004. *Politics in Time: History, Institutions, and Social Analysis*. Princeton: Princeton University Press.
- Pitkin, Hanna Fenichel. 1967. *The Concept of Representation*. Berkeley: University of California Press.
- Pocock, J.G.A. 1975. *The Machiavellian Moment: Florentine Political Thought and the Atlantic Republican*

- Tradition*. Princeton: Princeton University Press.
- Proyecto Estado de la Nación 2001. *Informe de la auditoría ciudadana sobre la calidad de la democracia en Costa Rica*. San José, Costa Rica: Proyecto Estado de la Nación.
- Przeworski, Adam, Fernando Limongi. 1997b. "Modernization: Theories and Facts." *World Politics* 49 (January) 155-83.
- Przeworski, Adam, Michael E. Alvarez, José Antonio Cheibub, and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Well-Being in the World, 1950–1990*. Cambridge: Cambridge University Press.
- Ranney, Austin. 1962. *The Doctrine of Responsible Party Government: Its Origins and Present State*. Urbana: University of Illinois.
- Reich, Gary. 2002. "Categorizing Political Regimes: New Data for Old Problems." *Democratization* 9:4 (Winter) 1-24.
- Rose-Ackerman, Susan. 1999. *Corruption and Government: Causes, Consequences, and Reform*. Cambridge: Cambridge University Press.
- Rousseau, Jean-Jacques. 1984. *A Discourse On Inequality*. Trans. by Maurice Cranston. Middlesex, England: Penguin..
- Ryfe, David M. 2005. "Does Deliberative Democracy Work?" *Annual Review of Political Science* 8: 49–71.
- Saward, Michael. 1998. *The Terms of Democracy*. Oxford: Polity.
- Saward, Michael. 2003. *Democracy*. Cambridge: Polity.
- Schattschneider, E.E. 1942. *Party Government*. New York: Rinehart.
- Schedler, Andreas. 2012a. "Judgment and Measurement in Political Science." *Perspectives on Politics* 10:1, 21-36.
- Schedler, Andreas. 2012b. "The Measurer's Dilemma: Coordination Failures in Cross-National Political Data Collection." *Comparative Political Studies* 45:2, 237–66.
- Schneider, Carsten Q. 2010. "Issues in Measuring Political Regimes." Working Paper 2010/12, Center for the Study of Imperfections in Democracy, Central European University. (<http://disc.ceu.hu/working-papers>).
- Schneider, Carsten Q., and Philippe C. Schmitter. 2004a. "The Democratization Dataset, 1974–2000." Unpublished manuscript, Department of Political Science, Central European University.
- Schneider, Carsten Q., and Philippe C. Schmitter. 2004b. "Liberalization, Transition and Consolidation: Measuring the Components of Democratization." *Democratization* 11(5): 59–90.
- Schumpeter, Joseph A. 1950 [1942]. *Capitalism, Socialism and Democracy*. New York: Harper & Bros.
- Sen, Amartya. 1999. "Democracy as a Universal Value." *Journal of Democracy* 10 (July) 3-17.
- Shapiro, Ian. 2003. *The State of Democratic Theory*. Princeton: Princeton University Press.
- Shapiro, Ian, and Casiano Hacker-Cordon, eds. 1999. *Democracy's Value*. Cambridge: Cambridge University Press.
- Silver, Nate. 2012. *The Signal and the Noise: Why So Many Predictions Fail – But Some Don't*. New York: Penguin.
- Steiner, Jurg, Andre Bachtiger, Markus Spornli, and Marco R. Steenbergen. 2004. *Deliberative Politics in Action: Analysing Parliamentary Discourse*. Cambridge: Cambridge University Press.
- Steinmo, Sven, Kathleen Thelen, and Frank Longstreth, eds. 1992. *Structuring Politics: Historical Institutionalism in Comparative Analysis*. Cambridge: Cambridge University Press.

- Sudders, Matthew, and Joachim Nahem. 2004. *Governance Indicators: A Users' Guide*. Oslo: UNDP.
- Teorell, Jan. 2011. "Over Time, Across Space: Reflections on the Production and Usage of Democracy and Governance Data." *Comparative Democratization* 9:1 (February) 1, 7.
- Thomas, Melissa A. 2010. "What Do the Worldwide Governance Indicators Measure?" *European Journal of Development Research* 22: 31–54.
- Thompson, Dennis F. 2008. "Deliberative Democratic Theory and Empirical Political Science." *Annual Review of Political Science* 11: 497–520.
- Treier, Shawn, and Simon Jackman. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* 52(1): 201–17.
- USAID. 1998. Handbook of Democracy and Governance Program Indicators. Technical Publication Series PNACC-390. Washington: USAID Center for Democracy and Governance.
- Vanhanen, Tatu. 2000. "A New Dataset for Measuring Democracy, 1810–1998." *Journal of Peace Research* 37: 251–265.
- Vermillion, Jim. 2006. "Problems in the Measurement of Democracy." *Democracy at Large* 3(1): 26–30.
- Vile, M.J.C. 1967/1998. *Constitutionalism and the Separation of Powers*. Liberty Fund.
- Walzer, Michael and David Miller. 1995. *Pluralism, Justice, and Equality*. Oxford: Oxford University Press.
- Weale, Albert. 1999. *Democracy*. New York: St. Martin's Press.
- Welzel, Christian. 2007. "Are Levels of Democracy Affected by Mass Attitudes? Testing Attainment and Sustainment Effects on Democracy." *International Political Science Review* 28(4): 397–424.
- Wilson, Woodrow. 1956 [1885]. *Congressional Government*. Baltimore: Johns Hopkins University Press.
- Wu, Lawrence L. 2000. "Some Comments on 'Sequence Analysis and Optimal Matching Methods in Sociology: Review and Prospect.'" *Sociological Methods and Research* 29(1): 41–64.
- Young, Iris M. 2000. *Inclusion and Democracy*. Oxford University Press, 2000
- Youngs, Richard (ed). 2010. *The European Union and Democracy Promotion: A Critical Global Assessment*. Baltimore: The Johns Hopkins University Press.