DEMOCRATIZATION
2021, VOL. 28, NO. 2, 308–332
https://doi.org/10.1080/13510347.2020.1818721

Routledge
Taylor & Francis Group

Check for updates

Democracy and human development: issues of conceptualization and measurement

John Gerring [a], Carl Henrik Knutsen [b], Matthew Maguire [c], Svend-Erik Skaaning [d], Jan Teorell [e] and Michael Coppedge [f]

[a]Department of Government, University of Texas, Austin, TX, USA; [b]Department of Political Science, University of Oslo, Oslo, Norway; [c]School of Management, San José State University, San José, CA, USA; [d]Department of Political Science, Aarhus University, Aarhus, Denmark; [e]Department of Political Science, Lund University, Lund, Sweden; [f]Department of Political Science, Kellogg Institute, Hesburgh Center, University of Notre Dame, Notre Dame, IN, USA

ABSTRACT

This study attempts to reconcile competing positions in an important debate about the relationship between regime type and human development. We contend that this empirical relationship is contingent upon issues of conceptualization and measurement in democracy. First, the relationship is more likely to be perceived when democracy is measured in a nuanced fashion, taking account of gradations of democracy and autocracy. Second, some aspects of democracy – those associated with competitive elections – are more strongly associated with human development than others. Third, the components of electoral democracy interact in a reinforcing manner. Finally, the impact of democracy on human development is a distal relationship that depends upon a country's entire regime history. Our approach draws on several new datasets that interrogate change across a century, enhancing empirical leverage on this important question. To measure human development, we employ the Gapminder project, covering most sovereign countries from 1900 to 2012. To measure democracy, we draw on Varieties of Democracy data, which measure democracy in a highly differentiated fashion for most sovereign countries from 1900 to the present. An extensive set of analyses offer strong corroboration for the argument.

ARTICLE HISTORY Received 7 October 2019; Accepted 23 August 2020

KEYWORDS Democracy; human development; infant mortality; elections; institutions

Introduction

Does democracy improve human wellbeing? Debate over this question generally focuses on how regime type affects economic growth. Yet per capita GDP does not purport to provide a summary measure of human welfare, and is especially poor at capturing the welfare of less advantaged citizens.[1] To measure the welfare of the poor one must turn to indicators focused on poverty, life-enhancing policies (e.g. vaccination), or more direct measures of wellbeing such as health.[2] Human development measures reflect

CONTACT John Gerring jgerring@austin.utexas.edu
Supplemental data for this article can be accessed https://doi.org/10.1080/13510347.2020.1818721
© 2020 Informa UK Limited, trading as Taylor & Francis Group

the status of the underprivileged and form the backbone of the Millennium Development Goals.

A growing literature examines the role of institutions in fostering human development. Some studies find that democracy improves quality of life.[3] Others challenge this assessment.[4] It is by no means a settled question, empirically or theoretically.

Sceptics point out problems of causal identification arising from highly trended variables, sample bias, and non-robustness in the relationship. In addition, they cast doubt on the mechanisms that might plausibly connect democracy to human development. First, voters may focus on more salient outcomes such as employment and economic growth.[5] If so, the mechanism of electoral accountability is called into question. Second, economically disadvantaged citizens carry less weight, even in democracies[6]; resources are captured by citizens who are better able to organize and make demands on the state.[7] Third, democratization may induce conflict and instability,[8] which presumably impairs human development. Fourth, newly democratized polities are weakly institutionalized and thus inclined to adopt clientelistic or populist policies rather than undertake long-term investments in human capital.[9] Finally, even if a democratically elected government enacts legislation favouring human development, low state capacity may undermine efficient implementation.[10]

This study attempts to reconcile competing positions in this debate. Decisions about operationalization often have important consequences.[11] We contend that the empirical relationship between democracy and human development, proxied by mortality-based indices such as infant mortality, is contingent upon issues of conceptualization and measurement in the independent variable – democracy. Four features bear upon the relationship. First, the relationship is clearer when measures account for gradations of democracy. Second, some aspects of democracy – namely, those associated with competitive elections – are more robustly associated with human development than others (e.g. popular participation, civil society, individual liberty). Third, the various components of electoral democracy interact in a reinforcing manner. Finally, this is a distal relationship that depends upon a country's entire regime history, not simply its current regime status.

In summary, democracy's impact on human development is most plausible when (i) the concept is measured as an interval variable, (ii) measures centre on democracy's electoral components, (iii) these components are aggregated multiplicatively to reflect interaction effects, and (iv) a country's historical stock of democracy is taken into account.

Some of these issues have been vetted previously. Wang, Mechkova and Andersson[12] use a fine-grained measure with extensive coverage and a measurement strategy that comports with (i) and (ii). They find that improvements in electoral aspects of democracy (when initial democracy levels are high) have positive effects on health outcomes even when accounting for level of corruption. These findings corroborate the argument that democracy promotes human development in a literature where studies point in all directions. Our theoretical argument highlights how the relationship between democracy and human development depends upon *all four* items listed above. We also offer a more disaggregated examination of the components of electoral democracy, different aggregation schemes, specification choices that deal with autocorrelation, and additional tests of potential mechanisms and confounders (including non-electoral aspects of democracy).

Thus, we bring together several elements of conceptualization and measurement hitherto treated in isolation and explicitly link them to an argument about how democracy affects human development. We show that a truly robust relationship between democracy and human development appears only when the four highlighted elements are combined into a single measure. This is why previous work, which accounts for only one or two elements, typically produces non-robust findings. In this fashion, we are able to reconcile seemingly contradictory findings from recent studies.

Our empirical approach draws on several datasets that allow us to interrogate evidence globally across a century. To measure human development, we employ data compiled by the Gapminder project that measures mortality – infant mortality, child mortality, and life expectancy – for most sovereign countries from 1900 to 2012. To measure democracy, we draw on data from Varieties of Democracy (V-Dem).[13]

We begin with a framework for understanding the concept of democracy and its relationship to human development. Next, we discuss the conceptualization and measurement of democracy and human development. Thereafter, we present numerous tests, first focusing on extant democracy indices compared to our preferred Multiplicative Polyarchy Index (MPI) and then considering different ways to aggregate and disaggregate MPI's components.

Arguments

Conceptualizing democracy's relationship to human development

Democracy may be linked to human development through several mechanisms. This follows from democracy's protean quality, which is open to various definitions, each suggesting somewhat different mechanisms. For example, if democracy is defined by strong rule of law and low corruption,[14] one principal mechanism may be the quality of governance. However, Wang, Mechkova and Andersson[15] demonstrate that empirical support for the good governance argument rests on data from very short time series with limited variation. When a longer time span is analysed based on V-Dem data, electoral democracy shows a more robust relationship with health than corruption when included in the same models. This relationship tends to persist over time.

In our discussion, we highlight two general theoretical frameworks that have been especially prominent in the literature. The first conceptualizes democracy around notions of citizen empowerment, and the second centres on elite-level contestation.

Citizen empowerment

According to one view, democracy affects human development through the empowerment of lay citizens and civic associations.[16] One avenue of empowerment is a free media. Granted press freedom, news outlets will report on policy disasters such as widespread famine, enhancing their salience in the public mind and invigorating public dissent.[17] Likewise, by disseminating mundane information pertaining to public health (e.g. benefits of lavatories compared to open-field defecation), quality of life may be improved.[18]

Another avenue of empowerment centres on civil society. Social connectedness (aka social capital) should have positive repercussions for public health, providing "people with a basis for cooperation that is mutually advantageous, a source of aid or assistance, a means of staying well informed about health issues, and a source of self-esteem".[19]

Popular participation in politics may also directly affect public health. Wigley and Akkoyunlu-Wigley[20] cite evidence from epidemiological studies showing that "the extent to which individuals perceive they have control over their lives plays a significant role in determining their health".

Although intuitive, the empowerment narrative may be challenged. First, there are questions about whether empowerment stands prior or posterior to human development.[21] Numerous studies suggest that health boosts economic performance,[22] and it might even affect social engagement and political participation.[23] One must therefore be cautious about inferring too much from correlations between participation and health. Second, some of the aforementioned channels do not operate independently of elite behaviour. Insofar as a free press helps to avert policy disasters, it is through incentivizing politicians to take particular actions – this matter is central to our alternative theory. Finally, improving nationwide conditions for human development requires vast resources. It is unclear how citizen empowerment could muster these resources or manage their distribution on a permanent basis, especially in poor countries with limited infrastructure. Only the state has sufficient resources and managerial capacity to make significant and sustained improvements in the quality of life for millions of citizens across a national territory.

Elite contestation

We contend that the relationship between democracy and human development involves masses and elites within a structure of electoral accountability, allowing state resources to be mobilized for a common purpose.

Consider the incentives facing leaders. Competitive elections establish a relationship of accountability between electors (principals) and leaders (agents) such that principals punish agents who do not perform as expected.[24] And long-established democracies are likely to have more institutionalized parties and party systems, which, in turn, may improve accountability and facilitate development-enhancing policies. Hence, when leaders compete for voters' approval in free elections, they will orient their policies to please their constituents. Insofar as electorates favour human development, democratic governments should seek to satisfy that desire.[25]

A key mechanism lies in public policies adopted by governments. A simple median-voter model suggests that competitive elections pressure politicians to institute redistributive policies in order to address social inequality.[26] Meanwhile, a large theoretical literature suggests that voters reward incumbents at the polls for resisting predation and providing public goods.[27]

Empirical studies suggest a strong relationship between democracy and policies with a redistributive or public goods orientation. Such policies include education[28] – especially primary education[29] – and infrastructure projects focused on the masses (e.g. sanitation and clean water) rather than privileged urban clienteles (e.g. hospitals).[30] Most studies also find that democracy enhances aggregate social spending and public sector size,[31] which might correlate with overall level of redistribution or public goods. There is, therefore, ample reason to expect that democracy affects public policies designed to improve human development.

Whether these policies achieve their stated goal is another matter. Classrooms may be filled while teachers are absent.[32] Health care expenditures may not reach rural areas where they are most needed.[33] Many factors connive to inhibit delivery of public

services to the poor, attenuating the connection between social spending and human development.[34]

Despite these inefficiencies, we expect such policies to make a big difference in the lives of poor people. Conditional cash transfer programmes, for example, seem to increase enrolment rates, improve preventive health care, and raise household consumption.[35] Generally, leaving aside "poverty trap" situations, we expect that the ease of improving someone's condition is inversely proportional to the severity of their condition. That is, the welfare of the poor is more responsive to government policy than the welfare of the rich. So, even where service delivery is flawed, we expect a relationship between policy effort and human development outcomes, and we anticipate such efforts to be bigger when the degree of electoral democracy is higher.

Conceptualizing and measuring democracy

We turn now to measurement of democracy, a diffuse concept with many ingredients and many possible aggregation techniques. Democracy is sometimes viewed as a binary feature,[36] and for some outcomes, a single-threshold measure is appropriate. However, when considering human development outcomes, we see no reason to suppose that the impact of contestation conforms to a threshold model. A minimal amount of contestation is good for human development, but greater contestation is even better.[37] The relationship between electoral contestation and human development should therefore be continuous and monotonic (though not necessarily linear).

Relatedly, we argue that various elements associated with electoral democracy enhance human development. This includes each institutional aspect of what Dahl[38] called *polyarchy*. Accordingly, we measure all factors that might affect responsiveness and accountability between leaders and citizens through competitive elections. This is presumably maximized when (1) elections are clean and not marred by fraud or systematic irregularities, (2) the chief executive is selected (directly or indirectly) through elections, (3) suffrage is extensive, (4) political and civil society organizations operate freely, and (5) there is freedom of expression, including access to alternative sources of information. Table B1 details each component, using measures drawn from V-Dem.[39]

Elections lie at the centre of this conception of democracy. Accordingly, we hypothesize that the purely electoral components of polyarchy will have greater impact on human development than the more peripheral components (elements 4–5 referenced previously above). However, elections do not occur in a vacuum; other aspects of polyarchy should play important supporting roles.

For example, in a polity where only a portion of adult citizens can vote, politicians are incentivized to ignore the needs of disenfranchized citizens. Since these citizens are invariably less well-off, this might adversely affect policies that enhance human development. Likewise, wherever the power of voters is attenuated through practices of vote-buying, coercion, fraud, and other tools of manipulation, we expect the disenfranchized to be predominantly poor. Citizens with higher incomes have more power resources for calling out and contesting electoral manipulation. It is thus less costly for politicians to buy votes and otherwise manipulate electoral outcomes in poor (and often rural) districts.[40] This sort of informal bias should have the same distributional effect as formal disenfranchisement, incentivizing politicians to ignore the needs of less advantaged citizens. In short, weaknesses in electoral democracy are likely to bring (de-

distributive) policy consequences. Certainly, middle- and upper-income voters occupy privileged positions in virtually *any* polity. Nonetheless, the degree to which the political system is tilted towards the wealthy is affected by the quality of democracy. When the provisions of polyarchy are in place, class bias should be mitigated.

High-quality democracy serves not only to empower poor people but also to channel their power in a programmatic direction – towards public goods and away from clientelistic payoffs. Some such policies pertain to classic public goods that are by definition non-exclusionary, including vaccination programmes or measures to prevent the spread of deadly pandemics. Even if the poorest citizens are excluded from a governing coalition they likely benefit from such policies. Yet, many other relevant policies are "quasi-public goods" or "club goods", which can be targeted to particular voters, e.g. by locating hospitals in areas where supporters reside. [41] Improvements in the quality of democracy are likely to incentivize politicians to cater to broader groups of voters, while deterioration incentivize leaders to target spending on narrower – and typically wealthier – groups.[42] Testing this logic on targeting of welfare legislation, Knutsen and Rasmussen[43] find that programmes cover a broader array of social groups in democracies than in autocracies, and we surmise that this dynamic holds also for other policies, such as electricity infrastructure.[44] Hence, when the quality of electoral democracy increases, out-groups, including the poor, should benefit.

We argue that the foregoing components of polyarchy have an interactive and complementary relationship. While the quality of elections is most crucial, each feature enhances the value of the others concerning the impact on human development. Likewise, a single weak link may critically impair electoral contestation. Absent elections it matters little if the regime tolerates free association or free expression. If participation in elections is restricted to a single party, it matters little if suffrage is universal. If free elections exist, but executive power is held by an unelected body, government responsiveness is severely mitigated. Finally, if political and civil society organizations cannot operate freely, or absent freedom of expression, it is difficult to hold government officials accountable.

Elites can thus deploy a "menu of manipulation",[45] choosing different mechanisms to suppress competition according to what they think they can get away with or what is most effective in forestalling democratic accountability. Any one tool of manipulation may be sufficient for securing incumbency. The ingredients of electoral contestation must be aggregated in a multiplicative fashion to capture these complementarities.[46]

The five elements described previously above are multiplied together to form a Multiplicative Polyarchy Index (MPI). Accordingly, the impact of one component depends on the scores of all other components. (V-Dem's Polyarchy Index can be considered as the mean of MPI and an additively aggregated polyarchy index formed from identical indicators.)[47] Since three components – (1), (2) and (3) – have a true zero, a polity receives zero if any of these three sub-components is zero.

We list all countries in our sample and their scores in 2000 in Table D1 and display a histogram of these scores in Figure D1. We then compare the MPI with the most commonly used historical measure of democracy, Polity2 from the Polity IV dataset.[48] Scatterplots, shown in Figures D2 and D3, compare point estimates from the two indices in 1900 and 2000. Although the measures correlate fairly well in the contemporary era (Pearson's $r = 0.85$ in 2000), they are not highly correlated in 1900 ($r = 0.57$), presumably due to the inclusion of suffrage in MPI. Polity2 judges nineteenth-century United

States to be completely democratic, while it receives a low score on the MPI. Likewise, New Zealand, the only country with full suffrage for men and women in 1900, receives the top score (in that year) from the MPI. Other points of disagreement concern the quality of democracy. For example, in 2000, Russia receives a fairly high score from Polity2 and a fairly low score from MPI. Generally, MPI establishes a more demanding standard of democracy than Polity2.

We argue, finally, that democracy has both short- and long-term effects on human development. Insofar as democracy affects public policies, we can differentiate policies with fairly immediate effects (e.g. vaccinating infants) and policies involving investments to be realized in the future such as electrical grid improvements, sanitation facilities and the education of nurses and doctors. We must also consider the long-range effects of democratic institutions on political parties, as alluded to. To capture both proximal and distal effects, given that the variable of interest is sluggish (and hence inappropriate for a distributed lag model), we calculate a "stock" measure of democracy to measure a country's regime history. Specifically, we measure each year (prior to the year of observation) back to 1900 or the year of independence, discounting each year by various rates.[49]

In sum, democracy's impact on human development is most plausible when the concept is measured as an interval variable, when measures are focused on democracy's electoral components, when these components are aggregated in a multiplicative fashion, and when a country's historical stock is incorporated into the resulting index. Figure D1 displays a histogram of the MPI stock index for 2000.

Measuring a key aspect of human development: mortality

Human development can also be measured in various ways.[50] We focus on mortality-based health outcomes – infant mortality, child mortality, and life expectancy – for three reasons. First, mortality is of paramount importance to all people, and the enjoyment of all other goods depends upon it. Second, mortality is relatively easy to measure since deaths are generally recorded, or at least remembered (and hence fairly accurately registered in retrospective surveys). Likewise, mortality does not involve difficult debates over definition and operationalization. Third, mortality-based indices offer strong coverage across countries and through time. The ability to project mortality rates backward in time – based on a variety of sources but most especially surveys and censuses – is a useful feature.[51]

By contrast, education-based measures of human development are difficult to interpret, especially since education is hard to compare across contexts. Even the measurement of literacy, a seemingly straightforward topic, is subject to the incomparability of languages and literacy standards. Measures of health that add other features to mortality – e.g. disability-adjusted life-years – are more difficult to measure and therefore provide restricted temporal coverage. Policy-based measures of health such as vaccination rates also have limited temporal coverage and are not applied to highly developed countries. Composite measures such as the Human Development Index – combining health, education, and GDP/capita – involve the foregoing problems and aggregation formulas that are hard to defend and interpret.[52]

It is unsurprising, therefore, that global studies of human development often centre on mortality.[53] We choose infant mortality as our benchmark measure. Humans are most vulnerable in the first year of life, and a society's infant mortality rate (IMR),

the number of babies who expire prior to their first birthday per 1000 live births, is likely sensitive to changes in public policy and environmental disorders. As expected, it displays the highest variance among our three measures, both through time and across countries. While child mortality (CMR; child deaths prior to age 5 per 1000 live births) is sometimes preferred, the two indices are extremely highly correlated ($r = 0.99$), and IMR offers somewhat longer time-series. IMR is also, by construction, highly correlated with life expectancy (LE) ($r = 0.89$; 0.93 with our transformed LE index). Thus, we employ IMR as our primary measure, and CMR and LE as secondary measures.[54] Data is drawn from Gapminder with supplemental data from the World Development Indicators, as explained in Table B1.

Gapminder draws on a variety of sources for its combined estimates. These include the UN Inter-agency Group for Child Mortality Estimation,[55] the Human Mortality Database,[56] UNICEF,[57] and Mitchell.[58] Data sources are listed for each observation.[59] The sources listed previously above are also compilations of data based on a much wider variety of underlying sources, as specified for each source.

To account for the bounded nature of IMR and CMR, which makes it difficult to achieve improvements when a society has reached low mortality levels, we transform them by the natural logarithm (following convention). LE is also bounded, although in a less strict fashion. We thus recalculate the index by subtracting LE from the maximum sample value (85) and then taking the natural logarithm. This flips the scale so that, like IMR and CMR, low numbers signal better performance.

The long time-series (particularly for IMR) carries econometric benefits. Because change in mortality rates generally occurs slowly, a long period of observation is essential in discerning potential relationships. A long time-series also facilitates fixed-effect models, alleviating concerns about specification while avoiding so-called Nickell bias. We also employ generalized method of moments (GMM) models to assuage such concerns.

Granted, the broad coverage of our dataset also carries potential problems. Historically, many countries have not had effective systems for registering births and deaths, which means that mortality rates are estimated from surveys or censuses. If the resulting measurement errors are stochastic, regression coefficients will have larger standard errors, making it harder to identify effects. Robustness tests are carried out with samples restricted to the contemporary era, where data quality is higher, and fewer estimates of IMR are interpolated.

While a long time-series of annual measures minimizes information loss, it also exaggerates the issue of assuming independence of each observation, enhancing problems of serial autocorrelation. Our main analyses thus use data parsed at 10-year intervals – although results are robust to panels formed from annual data, as shown.

Despite the extensive coverage of our IMR data, some countries are still excluded from our analyses due to missing data. Ross[60] finds that this omission affects estimated relationships between democracy and human development. Accordingly, we conduct robustness tests with imputed datasets to mitigate selection biases.

Empirics

Prior to engaging the empirics, we recapitulate our arguments in the form of four hypotheses:

H_1: Democracy's impact on human development is registered primarily through contested elections. Consequently, the electoral components of democracy should have a more robust relationship to subsequent human development than other aspects of democracy.

H_2: Democracy's impact on human development is most plausible when democracy is measured in a finely graded fashion. Consequently, interval indices should have a more robust relationship to subsequent human development than binary or ordinal indices.

H_3: Electoral components of democracy interact with each other to improve electoral accountability. Consequently, a multiplicative method of aggregation should show a more robust relationship to subsequent human development than other methods of aggregation.

H_4: Democracy's impact on human development cumulates over time. Consequently, a stock measure of democracy that measures a country's entire regime history should show a more robust relationship to subsequent human development than "level" measures of democracy at t-1.

We proceed with empirical tests focused on (a) extant democracy indices, (b) the robustness of the relationship between MPI and human development, (c) modes of aggregation and disaggregation of the index, and (d) causal mechanisms.

Democracy indices

We begin by comparing our proposed index of democracy, MPI, with prominent alternatives. Variable definitions are in Table B1 in the Online Appendix, and descriptive statistics in Table B2. All variables are adjusted to the same (0–1) scale so coefficients can be directly compared.

The benchmark model regresses IMR (logged) on democracy along with per capita GDP (to control for level of economic development) and country and year fixed-effects. We regard country fixed-effects as crucial since right- and left-side variables may be affected by static country characteristics (e.g. colonial experience, culture, geography). We regard year fixed-effects as equally important since mortality reduction responds to global factors that affect all countries, including the diffusion of health-relevant information and technological developments. The unit of analysis is country-decades, with all right-side variables lagged one time-period behind the outcome. An ordinary least squares (OLS) estimator is employed, with standard errors clustered by country.

We begin with several composite democracy indices that offer extensive coverage and fine-grained distinctions between levels of democracy.[61] Polity2[62] uses a weighted additive aggregation procedure across five sub-components: competitiveness and openness of executive recruitment, competitiveness and regulation of political participation, and constraints on the chief executive. (The latter accounts for about 1/3 of the index's range). The Unified Democracy Scores (UDS), developed by Pemstein, Meserve, and Melton,[63] employ a Bayesian latent variable model to combine commonly used democracy measures. Finally, we include two summary indices – "Contestation" and "Inclusiveness" – developed by Miller,[64] following the conceptual model proposed by Coppedge, Alvarez, and Maldonado.[65] While intended to capture the two classic polyarchy dimensions of Dahl,[66] the measures also draw on indicators tapping other aspects of democracy. For instance, "Contestation" draws on Polity's Executive Constraints indicator and Banks' measure of Legislative Effectiveness.

A second set of indices, from V-Dem, focus on various features of citizen empowerment, which provides the main alternative theoretical account for why democracy

enhances human development (Section I). These indices include Participation, Deliberation, Female Empowerment, Civil Society, and Individual Liberty.

A third set of indices are binary democracy measures. The "BMR" index[67] resembles the Democracy-Dictatorship ("DD") index[68] insofar as it centres on existence of contested multi-party elections. Unlike DD, BMR adds a participation criterion, checks for reports of electoral irregularities, does not rely solely on post-electoral alternation of governments when coding elections as free and fair, and extends back to the nineteenth century. A second binary measure, "BNR", constructed by Bernhard, Nordstrom, and Reenock,[69] covers 124 countries from 1913 to 2010. Following Dahl,[70] BNR defines a country as democratic if contestation is high and ≥50% of the adult population can vote.

The final set of indices is multiplicative. This includes the product of Miller's Contestation and Inclusiveness indices along with our MPI, described previously above.

In Table 1, we subject each index to several tests. First, we test "level" measures of each index. Second, we calculate each democracy index as a stock variable, extending back to 1900 with a 10% annual depreciation rate. Third, we introduce a lagged dependent variable to correct for possible trend effects or unmeasured confounders. Fourth, we calculate each index as a stock with a very slow 1% annual depreciation rate. Each column in Table 1 thus reports four regressions, with results inserted only for the variable of interest. Naturally, the interpretation of coefficients in each row is somewhat different. Here, we focus on statistical significance, taking standard *p*-value thresholds (90%, 95%, and 99%) as markers of success. This is arbitrary, to be sure, but imposes a uniform threshold and facilitates comparisons across multiple measures.

Results in Table 1 and in additional tests contained in Appendix C are consistent with each of our hypotheses. Composite and empowerment indices of democracy (Columns 1–9) are not as robustly associated with IMR as indices cantered on democracy's electoral dimension (Columns 10–13), corroborating H_1. Results suggest that binary, electoral measures (Columns 10–11) have a strong relationship to human development. However, further tests with annual data (Table C10) or with alternate estimators and control variables (Tables C14 and C15) reveal that these measures are not as robust as MPI, corroborating H_2. Multiplicative measures of democracy (Columns 12–13) are more strongly associated with improved IMR than continuous indices employing other aggregation techniques (Columns 1–9), corroborating H_3. The most direct comparison of this hypothesis is between Columns 3–4 (Contestation and Inclusiveness indices from Miller) and Column 12 (their product). Stock indices of democracy generally bear a stronger relationship to IMR than "level" measures, corroborating H_4. Here, each index can be compared to itself across various rows in Table 1. In only one case is a level measure stronger than a correspondent stock measure (Column 4).

A final comparison involves the two multiplicative indices, from Miller and MPI, shown in Columns 12–13. Both are robust across specifications in Table 1, as we would expect since the index we constructed from Miller's data achieves most of the goals of conceptualization and measurement outlined in Section II. However, the estimated coefficient for MPI is greater in three of four tests, and substantially so in two, although the means and standard deviations of these indices are similar (Table B2). Moreover, in horse-race tests MPI demonstrates greater predictive power (see Figure 1).

Not only is the relationship between IMR and MPI robust across different specifications; the predicted magnitude is substantial. To illustrate, using the estimate in Model 2, Column 13, a country with all other covariates at their means is predicted to

Table 1. Democracy indices.

Aggregation	Composite				Empowerment					Binary		Multiplicative	
Index	Polity2 (Marshall)	UDS (Pemstein)	Contestation (Miller)	Inclusiveness (Miller)	Participation (V–Dem)	Deliberation (V–Dem)	Female Power (V–Dem)	Civil Society (V–Dem)	Individual Liberty (V–Dem)	BMR (Boix)	BNR (Bernhard)	Contest.* Inclusive (Miller)	MPI (authors)
Measures	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Level	−0.160**	−0.450***	−0.156*	−0.142*	0.135	0.074	0.049	0.128	0.015	−0.100*	−0.117**	−0.300***	−0.320***
	(0.073)	(0.158)	(0.080)	(0.075)	(0.167)	(0.090)	(0.189)	(0.095)	(0.100)	(0.054)	(0.049)	(0.097)	(0.089)
2. Stock (10%)	−0.166*	−0.153	−0.145	0.110	0.049	0.039	−0.177	0.010	−0.133	−0.178**	−0.222***	−0.361***	−0.530***
	(0.100)	(0.174)	(0.116)	(0.099)	(0.182)	(0.137)	(0.167)	(0.134)	(0.136)	(0.083)	(0.075)	(0.131)	(0.119)
3. Stock (10%), Y^{t-1}	−0.097**	−0.073	−0.092*	0.046	−0.055	−0.035	−0.086	−0.092	−0.097*	−0.103***	−0.095**	−0.128**	−0.106*
	(0.047)	(0.097)	(0.054)	(0.055)	(0.076)	(0.053)	(0.085)	(0.057)	(0.056)	(0.037)	(0.042)	(0.064)	(0.056)
4. Stock (1%)	−0.290	−0.365	−0.329	0.259	−0.158	−0.428	−0.802**	−0.344	−0.904***	−0.377**	−0.300**	−0.392*	−0.741***
	(0.211)	(0.232)	(0.217)	(0.298)	(0.312)	(0.279)	(0.393)	(0.249)	(0.273)	(0.158)	(0.131)	(0.211)	(0.187)
GDPpc (ln)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Countries	152	154	150	150	154	154	154	154	154	154	152	150	154
Decades	11	7	11	11	11	11	11	11	11	11	7	11	11
Obs (approx.)	920	819	907	906	1005	1005	953	1001	1005	927	795	906	993

Note: *Outcome:* Infant mortality rate (ln). We test four measures of each index: 1. level, 2. stock (10% annual depreciation rate), 3. stock (10% annual depreciation rate) with lagged dependent variable, 4. stock (1% annual depreciation rate). *Units:* country-decades. *FE:* fixed effects. All right-side variables measured at t−1. All democracy indices measured on 0−1 scale. *Estimator:* OLS, standard errors clustered by country. ***$p < .01$ **$p < .05$ *$p < .10$.

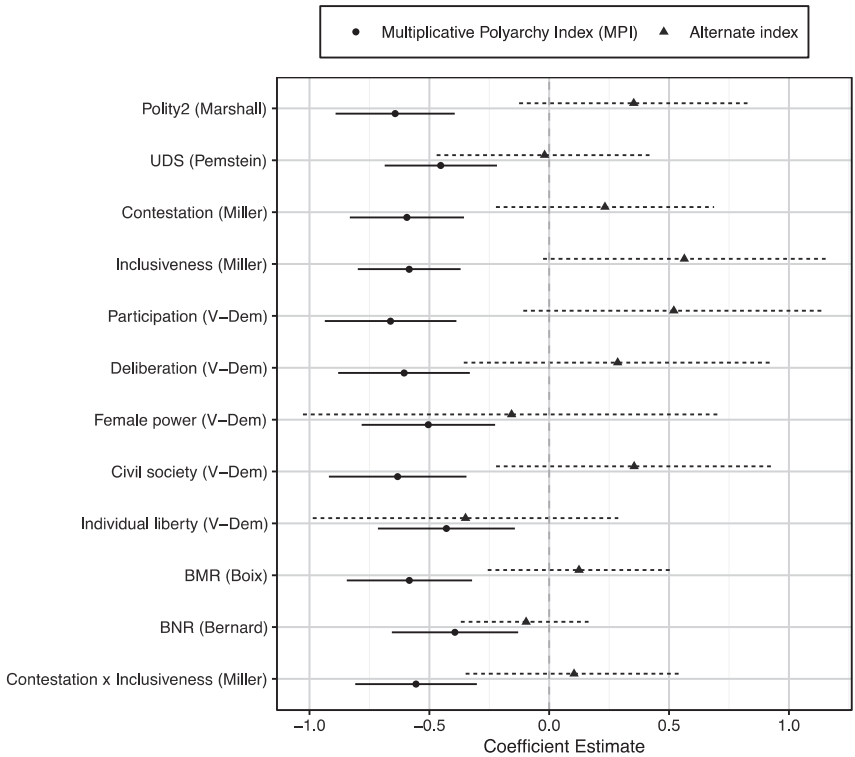

Figure 1. Horse-race tests.

Note: Coefficients plot with 95% confidence intervals. Outcome: infant mortality rate (ln). See Table D6 for complete results.

experience a 10% decrease in ln(IMR) when moving from the 10th to the 90th percentile observation on MPI. To further illustrate, imagine a very poor country with per capita GDP of $1000 and no, or an extremely autocratic, regime history (giving 0 on MPI stock), while all other variables are at their means. This approximates contemporary North Korea and many African countries upon independence. Our benchmark predicts that this country – with country and year fixed-effects at their means – should have an IMR of about 73 (per 1000 live births). Now, suppose that this hypothetical country quickly transitions to high-quality democracy (1 on MPI) and maintains that democracy level for a decade without *any* increase in wealth. Our model predicts that this change will result in a 32% drop in IMR during those ten years. For similar countries starting with per capita GDP of $500 and $5000, a similar change in regime history would reduce IMR from 94 to 64 and 41 to 28, respectively. Thus, the onset of high-quality democratic rule may have a dramatic effect on mortality rates.

Next, we enlist the same indices in "horse-race" tests with our preferred measure, MPI stock (10%). Since alternate indices perform best in a stock format with slow (1%) depreciation rate (Table 1), we employ this version for alternate indices. These tests, summarized in Figure 1, show that the relationship of MPI to IMR is scarcely affected by including other democracy indices. In some instances, the estimated coefficient is slightly stronger and in others slightly weaker, but in *all* instances, MPI remains

statistically significant (at 1%). By contrast, alternate indices of democracy are generally attenuated, and in every case, the confidence interval encompasses zero.

Together, the tests displayed in Table 1 and Figure 1 corroborate our hypotheses that democracy's impact on human development is most robust when the conceptualization of democracy centres on elections (rather than the "empowerment" dimension of democracy), when the components of electoral democracy are aggregated multiplicatively, when the resulting index is continuous, and when democracy is measured as a historical stock.

Robustness

Table 2 displays a battery of tests with different specifications and functional forms. A key feature of all these tests is the incorporation of measurement error from the V-Dem measurement model, where multiple ratings are combined into a single point estimate along with a confidence interval. All models employ 10-year panels; replications with 1-year panels are shown in the bottom row. To save space, we present the following tests without commentary; a full discussion is available in Appendix C.

Model 1 is the benchmark, where MPI is measured as a stock variable with 10% annual depreciation. Model 2 offers the conventional "level" measure of MPI. Model 3 returns to the 10% depreciation rate, this time using linear, squared and cubic trend variables to control for temporal effects. The next two models are intended to deal with the highly trended nature of our variables. Model 4 includes a lagged dependent variable, which mitigates but does not entirely resolve the issue of first-order serial correlation. Model 5 adopts a first-difference specification, measuring IMR as a change from t to t+1 and MPI stock and GDP per capita from t-1 to t; lagged MPI stock (undifferenced) is also included to capture long-term effects. Model 6 removes per capita GDP from the model, on the assumption that it may be post-treatment. Model 7 adds several covariates that might possibly affect infant mortality and MPI. Model 8 imputes missing data using the Amelia II software. Model 9 lags MPI by three decades (t-30), offering more assurance against circularity and simultaneity. Model 10 enlists a dynamic panel model, system-generalized method of moments (GMM), developed explicitly for studying sluggish variables. The AR(2) test on residual autocorrelation returns a p-value of .16, whereas the Hansen test p-value is .20; thus, MPI should be consistently estimated. (The result is even stronger when including a second lag on the DV; there, AR(2) and AR(3) tests yield p-values of .50 and .39, respectively.)

The final models in Table 2 focus on alternate mortality-based outcomes. Model 11 features IMR drawn from the World Development Indicators. This model, restricted to the post-1960 period, also mitigates concerns about poor data quality earlier in the twentieth century. Model 12 features the child mortality rate. Model 13 adopts (the transformed measure of) life expectancy.

All of these robustness tests are passed at conventional thresholds, except for Model 10 where MPI is weakly significant (10%-level). In the bottom row of Table 2 we show results for the same set of tests when estimated with annual (rather than decadal) panels. Results are robust in all cases, and more precisely estimated – though with greater risk of serial autocorrelation.

In Appendix C, we conduct additional robustness tests focused on various threats to inference, e.g. excluding various regions (Table C2), functional form (Table C3), non-

Table 2. MPI and mortality.

Outcome	IMR	IMR	IMR	IMR	IMR	IMR	IMR	IMR	IMR	IMR	IMR(WDI)	CMR	LE
Estimator	OLS	OLS	OLS	OLS	FD, RE	OLS	OLS	OLS	OLS	Sys. GMM	OLS	OLS	OLS
Sample	Full	Full	Full	Full	Full	Full	Full	Imputed	Full	Full	1960–	Full	Full
Model	1	2	3	4	5	6	7	8	9	10	11	12	13
MPI level		−0.314***											
		(0.090)											
MPI stock (10%)	−0.537***		−0.539***	−0.107*	−0.071***	−0.804***	−0.394***	−0.964***		−0.286*	−0.480***	−0.465***	−0.197***
	(0.125)		(0.120)	(0.060)	(0.023)	(0.136)	(0.122)	(0.127)		(0.161)	(0.150)	(0.146)	(0.073)
MPI stock (10%), FD					−0.267***								
					(0.082)								
MPI stock (10%), *T*−30									−0.316**				
									(0.125)				
GDPpc (ln)	−0.369***	−0.397***	−0.357***	−0.047			−0.350***	−0.456***	−0.373***	−0.217	−0.312***	−0.378***	−0.199***
	(0.078)	(0.079)	(0.075)	(0.030)			(0.055)	(0.043)	(0.084)	(0.149)	(0.063)	(0.082)	(0.041)
GDPpc (ln), FD					−0.141***								
					(0.040)								
Urbanization							0.235						
							(0.311)						
Fertility (ln)							0.500***						
							(0.087)						
Growth							0.002						
							(0.004)						
Internal conflict							0.057						
							(0.040)						
External conflict							0.022						
							(0.060)						
Corruption stock (10%)							0.057***						
							(0.019)						
IMR^{t-1}				0.805***						0.690***			
				(0.029)						(0.164)			

(*Continued*)

Table 2. Continued.

Outcome	IMR	IMR	IMR	IMR	IMR	IMR	IMR	IMR	IMR	IMR	IMR(WDI)	CMR	LE
Estimator	OLS	OLS	OLS	OLS	FD, RE	OLS	OLS	OLS	OLS	Sys. GMM	OLS	OLS	OLS
Sample	Full	Full	Full	Full	Full	Full	Full	Imputed	Full	Full	1960–	Full	Full
Model	1	2	3	4	5	6	7	8	9	10	11	12	13
Year cubed			✓										
Decade FE	✓	✓		✓	✓	✓	✓	✓	✓		✓	✓	✓
Country FE	✓	✓	✓	✓		✓	✓	✓	✓		✓	✓	✓
Time trend										✓			
Countries	154	154	154	154	153	168	108	170	133	154	154	155	154
Decades	11	11	11	11	10	11	11	11	9	11	5	9	11
Obs	993	993	993	923	839	1132	751	1578	885	800	659	841	1090
R2 (within)	(0.907)	(0.904)	(.905)	(0.965)	(0.215)	(0.885)	(0.942)		(0.894)		(0.837)	(0.878)	(0.887)
MPI, as above	−0.571***	−0.233***	−0.582***	−0.015**	−0.010***	−0.730***	−0.384***	−1.657***	−0.338***	−0.080**	−0.456***	−0.460***	−0.184***
(annual data)	(0.127)	(0.072)	(0.124)	(0.007)	(0.003)	(0.129)	(0.135)	(0.110)	(0.124)	(0.040)	(0.136)	(0.131)	(0.066)

Note: *Outcomes*: IMR (infant mortality rate, logged), CMR (child mortality rate, logged), LE (life expectancy, reverse scale, logged). *Units of analysis:* country-decades (above) or country-years (bottom, except Column 11, which uses 5-year panels). Right-side variables measured at $T-1$ unless otherwise noted. *Estimators:* OLS (ordinary least squares), FD (first-difference), RE (random effects). All models incorporate measurement error for MPI based on posteriors produced by the V-Dem measurement model. Robust standard errors clustered by country. ***$p < .01$ **$p < .05$ *$p < .10$.

linearity in MPI (Table C4), instrumental variables (Table C5), non-linearities in GDP (Table C6), "level" measures of MPI (Table C7), Driscoll-Kraay Standard Errors (Table C8), GDP calculated as historical stock (Table C9), full results for tests with annual panels (Table C10), controlling for political durability (Table C11), no measurement error in MPI (Table C12), and panels restricted to the post-1960 period (Table C13).

Serial correlation is a serious concern given that the left- and right-side variables of theoretical interest are both highly trended. To mitigate this issue, our benchmark model adopts 10-year (rather than annual) panels and standard errors clustered by country to generate panel-correlated errors. In robustness tests shown in Table 2, we include a lagged dependent variable model, a first-difference model, and a GMM model. We also construct a model with multiple lags of the dependent variable (t-1, t-2, and t-3), shown in Model 5, Table C12. Several additional attempts to remedy the issue can also be found in the Appendix B. Although the main finding is robust in all of these models, a Wooldridge test for autocorrelation in panel data suggests that we cannot dismiss this possible threat to inference. Only the GMM model passes tests of serial autocorrelation, as noted in our previous discussion above. This reservation should be borne in mind.

Aggregation and disaggregation

In this section, we investigate alternative modes of aggregation and disaggregation. First, working with the five components of MPI and the same stock depreciation rate (10%), we generate alternate indices by combining these components in different ways, with results displayed in Figure 2. Model 1 replicates our benchmark where MPI is constructed by multiplication. Model 2 tests an alternative that is similar in conception to MPI, namely the minimum, or weakest-link, rule, where the index takes its value from the lowest-scoring indicator.[71] Conveniently, V-Dem indices are arranged across a 0–1 scale, assuring scale equivalence. Model 3 tests V-Dem's Electoral Democracy Index (EDI). This aggregation scheme, described in Coppedge et al.,[72] combines multiplicative and additive rules. Model 4 tests an index constructed by the first component of a principal component analysis. Model 5 tests a measure averaging across the indicators.

Results displayed in Figure 2 show that the multiplication rule outperforms other aggregation schemes. Among the alternatives, those closest to MPI in their construction – namely, the weakest-link index and V-Dem EDI – also predict lower infant mortality, though only at 5%. Other alternatives, which dispense entirely with multiplicative logic, show no relationship to IMR. Figure D4 shows that the substantive size of the predicted effect is also much stronger for MPI than for the alternatively aggregated indices. For instance, a change in MPI from the 10th to 90th percentile predicts a 10% decrease in ln(IMR), compared to a 6% decrease when using V-Dem's EDI.

This finding corroborates our hypothesis that with respect to human development, political institutions pertaining to electoral democracy are not substitutable. Rather, they reinforce one another. Aggregation schemes that average across these components, or observe only the common dimension (as identified by factor analysis), do not capture these interactions and show no clear relationship to human development. These results resonate with studies showing that there are many ways to subvert electoral democracy

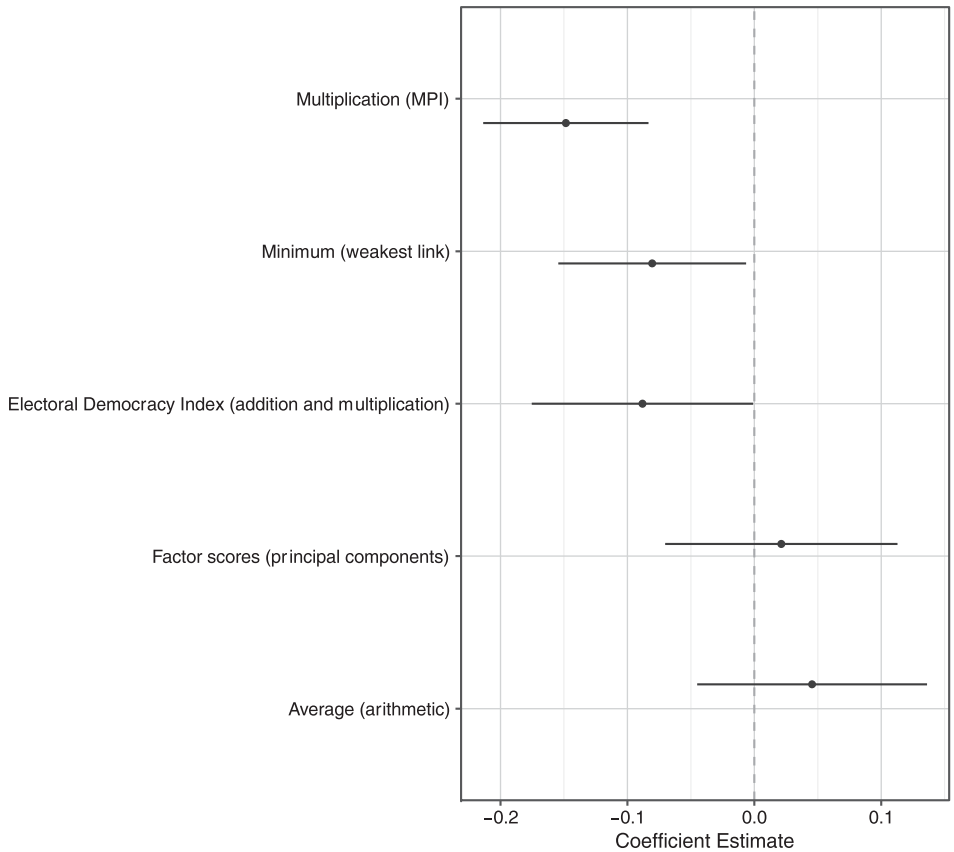

Figure 2. Varying aggregation rules.

Note: Coefficients plots with 95% confidence intervals, using standardized coefficients. *Outcome:* Infant mortality rate (ln). See Table D7 for complete results.

even while maintaining a pretence of democracy by satisfying some elements of that ideal.[73]

In Appendix D, we probe MPI's components further. First, the benchmark model is replicated with versions of the MPI that exclude each component, seriatim (Table D3). None of these exclusions compromise the core relationship between electoral democracy and infant mortality, though the estimated impact of indices with excluded components is lower than the full index, as expected. Second, the benchmark is replicated with stock variables constructed for each ingredient of MPI, seriatim (Table D4), where we theoretically expect weaker results. Only the purely electoral component, Clean Elections, predicts lowered infant mortality.

Conclusion

The question of democracy and human development has attracted a sizeable body of scholarly work. Despite this attention, there is no consensus about whether democracies outperform autocracies. Most of the debate centres on issues of research design and analysis. We propose that a key factor underlying the disparate findings in published research concerns issues of conceptualization and measurement, which are often neglected.

We argue that democracy's relationship to human development is most robust when measures are focused on the electoral components of democracy (H_1), when democracy is measured in a finely graded fashion (H_2), when components are aggregated in a multiplicative fashion (H_3), and when a country's historical experience is incorporated into the resulting index (H_4).

This explains why extant studies that use (a) composite indices that include disparate elements, (b) truncated measures, (c) indices composed through factor analysis, IRT models, addition, or other non-interactive methods of aggregation, and/or (d) "level" measures of democracy do not always demonstrate a robust association to human development.

These alternative ways of conceptualizing and measuring democracy are not wrong-headed. They are, however, less than ideal for this particular theoretical question, for reasons already elaborated above. Granted, if samples were larger – e.g. if there were 1000 countries rather than 180, or if the time-series available for analysis stretched back for 500 years rather than 100 – then, we suspect, many measures of democracy would clear the bar. That is, they would show a statistically significant relationship to improved human development, proxied by infant mortality or some other indicator. Even so, issues of conceptualization and measurement would presumably affect the magnitude of the estimated effect and hence its practical significance.

In our analyses, we compared different democracy indices to each other in separate tests, and we conducted "horse-race" tests in which alternate indices are tested against our preferred MPI. Moreover, we conducted a wide range of robustness tests focused on the MPI. Finally, we examined the impact of different aggregation techniques. These tests support our claims that the democracy-human development relationship is contingent upon several distinct choices pertaining to the conceptualization and measurement of democracy. More specifically, the relationship with human development is most robust when we focus on electoral features, aggregating these features in a (multiplicative) fashion that recognizes their complementarities, and accounting for a country's historical experience.

Despite our best efforts, we cannot rule out the possibility of unobserved confounding or autocorrelation issues that disturb estimates of uncertainty. Another caveat concerns the outcome: infant mortality, supplemented by child mortality and life expectancy. We regard mortality as the best available measure of human development. But it is by no means the only measure. Further work must determine whether the relationships discovered here are applicable to other measures of human development.

Before concluding, we turn our attention to the mechanisms at work in this story. Our analyses have focused on the main effect, for which we have posited several potential explanations. We cannot test all these factors in a rigorous fashion due to problems of measurement and problems generic to mechanismic inference.[74] Nonetheless, some mechanisms – particularly those associated with public policies focused on public health – are measurable, opening the prospect of a preliminary mediation analysis. These tests are reported and discussed in Appendix E. Although limited data coverage and problems of causal inference associated with observational data suggest caution in reaching conclusions, these analyses corroborate our theory. Public health spending seems to serve as an important pathway connecting electoral democracy to lower mortality. That said, the question of causal mechanisms deserves more attention from scholars, as this result is subject to multiple assumptions and is undoubtedly incomplete.

Notes

1. Costanza et al., "Beyond the GDP."
2. ul Haq, *Reflections*.
3. Blaydes and Kayser, "Counting Calories"; Brown, "Reading, Writing and Regime Type"; Deacon, "Public Good Provision"; Eterovic and Sweet, "Democracy and Education"; Gerring, Thacker, and Alfaro, "Democracy and Human Development"; Ghobarah, Huth, and Russett, "Comparative Public Health"; Hanson, "Democracy and State Capacity"; Kudamatsu, "Has Democratization Reduced Infant Mortality?"; Lake and Baum, "The Invisible Hand"; Miller, "Women's Suffrage"; Przeworski et al., *Democracy and Development*; Stasavage, "Democracy and Education Spending"; Zweifel and Navia, "Democracy, Dictatorship"; Wang, Mechkova, and Andersson, "Does Democracy Democracy Enhance Health?"
4. Gauri and Khaleghian, "Immunization"; Hallerod et al., "Bad Governance"; Jacobsen. "Revising the Modernization Hypothesis"; Miller, "Electoral Authoritarianism"; Nelson, "Elections, Democracy, and Social Services"; Ross, "Is Democracy Good for the Poor?"; Rothstein, "Guilty as Charged?"; and Truex, "The Myth of the Democracy Advantage."
5. Harding and Stasavage, "What Democracy Does"; Dahlum and Knutsen, "Do Democracies Provide Better Education?"
6. Przeworski, *Democracy and the Limits of Self-government*.
7. Lipton, *Why Poor People Stay Poor*.
8. Mansfield and Snyder, *Electing to Fight*.
9. Keefer, "Clientelism, Credibility."
10. Rothstein, *Quality of Government*.
11. Casper and Tufis, "Correlation versus Interchangeability."
12. Wang, Mechkova, and Andersson, "Does Democracy Democracy Enhance Health?"
13. Coppedge et al., *V-Dem Methodology v7.1*.
14. Inglehart and Welzel, *Modernization*.
15. Wang, Mechkova, and Andersson, "Does Democracy Democracy Enhance Health?"
16. Christoforou, "Social Capital and Human Development"; Fung and Wright, *Deepening Democracy*.
17. Sen, *Poverty and Famines*.
18. Wigley and Akkoyunlu-Wigley, "The Impact of Regime Type on Health."
19. Ibid., 653.
20. Ibid., 651.

21. Inglehart and Welzel, *Modernization*.
22. Bhargava et al., "Modeling the Effects of Health."
23. Mattila et al., "Healthy Voting."
24. Ferejohn, "Incumbent Performance."
25. Autocrats, under certain conditions, could also be incentivized to pursue policies that improve human development. Specific institutional features (e.g. Wright, "Do Authoritarian Institutions Constrain?"), characteristics of the autocrat's core supporters (e.g. Bueno de Mesquita et al., *The Logic of Political Survival*), or objectives of the ruling elite may induce leaders to pursue good policies even without contested elections. Still, in most relevant contexts, improvements in electoral democracy should, *ceteris paribus*, strengthen leaders' incentives to improve nation-wide human development outcomes.
26. Meltzer and Richards, "A Rational Theory."
27. Bueno de Mesquita et al., *The Logic of Political Survival*; Lake and Baum, "The Invisible Hand."
28. E.g. Eterovic and Sweet, "Democracy and Education."
29. Brown, "Reading, Writing and Regime Type."
30. Besley and Kudamatsu, "Health and Democracy."
31. Boix, "Democracy, Development and the Public Sector"; Huber, Mustillo, and Stephens, "Politics and Social Spending."
32. Chaudhury et al., "Missing in Action."
33. van de Walle, "The Distribution of Subsidies."
34. Filmer and Pritchett, "The Impact of Public Spending on Health"; Ross, "Is Democracy Good for the Poor?"
35. Rawlings and Rubio, "Evaluating the Impact."
36. Boix, Miller, and Rosato, "A Complete Data Set"; Przeworski et al., *Democracy and Development*.
37. Cf. Miller, "Electoral Authoritarianism."; see also Wang, Mechkova, and Andersson, "Does Democracy Democracy Enhance Health?"
38. Dahl, *Polyarchy*; *Democracy and its Critics*, 221.
39. Coppedge et al., *V-Dem [Country-Year/Country-Date]*; see Appendix A.
40. See, e.g. Jensen and Justesen, "Poverty and Vote Buying"; Knutsen et al., "Economic Development and Democracy."
41. Ross, "Is Democracy Good for the Poor?"
42. Bueno de Mesquita et al., *The Logic of Political Survival*.
43. Knutsen and Rasmussen, "The Autocratic Welfare State."
44. Brown and Mobarak, "The Transforming Power of Democracy."
45. Schedler, "The Menu of Manipulation."
46. Munck, *Measuring Democracy*.
47. Teorell et al., "Measuring Polyarchy."
48. Marshall, Gurr, and Jaggers, *Polity IV Project*.
49. For further discussion of the stock concept, see Gerring, Thacker, and Alfaro, "Democracy and Human Development."
50. McGillivray, "Measuring Non-economic Well-being Achievement"; McGuire, *Wealth, Health, and Democracy*, 17–21.
51. Riley, "Estimates of Regional and Global Life Expectancy."
52. Acharya and Wall, "An Evaluation"; Raworth and Stewart, "Critiques of the Human Development Index."
53. Sen, "Mortality as and Indicator."
54. Potential problems in the measurement of IMR – largely related to under-counting in poor countries – are discussed in Anthopolos and Becker, "Global Infant Mortality." Yet, this problem may be regarded as orthogonal to regime type and hence part of the error term. Insofar as per capita GDP (logged) explains under-reporting across the sample, as Anthopolos and Becker (ibid.) suggest, bias from measurement error is conditioned in our analyses.
55. www.childmortality.org
56. www.mortality.org
57. www.childinfo.org
58. Mitchell, *International Historical Statistics. Africa, Asia & Oceania*; *International Historical Statistics. The Americas*; *International Historical Statistics. Europe*.
59. www.gapminder.org/data/documentation/gd002/

60. Ross, "Is Democracy Good for the Poor?"
61. We exclude indices with shorter time-series, e.g. those produced by Freedom House and Bertelsmann Transformation Index, because they are less appropriate for estimators privileging within-unit change over time.
62. Marshall, Gurr, and Jaggers, *Polity IV Project*.
63. Pemstein, Meserve, and Melton, "Democratic Compromise."
64. Miller, "Democracy Pieces."
65. Coppedge, Alvarez, and Maldonado, "Two Persistent Dimensions."
66. Dahl, *Polyarhcy*.
67. Boix, Miller, and Rosato, "A Complete Data Set."
68. Cheibub, Gandhi, and Vreeland, "Democracy and Dictatorship Revisited"; Przeworski et al., *Democracy and Development*.
69. Bernhard, Nordstrom, and Reenock, "Economic Performance."
70. Dahl, *Polyarhcy*.
71. Goertz, *Social Science Concepts*.
72. Coppedge et al. *V-Dem Methodology v7.1*.
73. Levitsky and Way, *Competitive Authorianism*; Gandhi and Lust-Okar, "Elections under Authoritarianism"; Schedler, "The Menu of Manipulation."
74. Imai et al., "Unpacking the Black Box."

Acknowledgments

This research project was supported by Riksbankens Jubileumsfond, Grant M13- 0559:1, PI: Staffan I. Lindberg, V-Dem Institute, University of Gothenburg, Sweden; and by Knut and Alice Wallenberg Foundation to Wallenberg Academy Fellow Staffan I. Lindberg, Grant 2013.0166, V-Dem Institute, University of Gothenburg, Sweden. We are thankful for valuable inputs from the reviewers and for comments received in connection to presentations at the annual V-Dem conference in 2015, the annual APSA conference in 2015 and the yearly meeting of the Danish Political Science Association in 2016. Moreover, we thank Staffan I. Lindberg for several rounds of feedback.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research project was supported by Riksbankens Jubileumsfond, Grant M13-0559, by the Swedish Research Council, Grant C0556201, and by the Knut and Alice Wallenberg Foundation, Grant 2013.0166.

Notes on contributors

John Gerring (PhD, University of California at Berkeley, 1993) is Professor of Government at University of Texas at Austin, where he teaches and conducts research on methodology and comparative politics. He is co-editor of *Strategies for Social Inquiry*, a book series at Cambridge University Press, and serves as co-PI of *Varieties of Democracy* (*V-Dem*) and the *Global Leadership Project* (*GLP*).

Carl Henrik Knutsen is Professor of Political Science at the University of Oslo. His research interests include the economic effects of political institutions, the determinants of regime change, and the measurement of democracy. He has published his work in journals such as the *American Journal of Political Science*, *British Journal of Political Science*, *Comparative Political Studies*, *Journal of Politics*, *World Development*, and *World Politics*. He currently leads an ERC Consolidator Grant project on the "Emergence, Life, and Demise of Autocratic Regimes" and is co-PI of V-Dem

Matthew Maguire (PhD, Boston University, 2017) is Assistant Professor of Management in the Lucas College and Graduate School of Business at San José State University. His teaching and research

interests lie in the fields of comparative politics, business-government relations, and corporate social responsibility. He has held visiting positions at the Max Planck Institute for the Study of Societies in Cologne and the Institute for Human Sciences in Vienna.

Svend-Erik Skaaning is professor of political science at Aarhus University, Denmark. His research interests include the conceptualization, measurement, and correlates of democracy. He has recently published three books on these issues: the single-authored *Democracy* (Aarhus University Press) and the co-authored *Democratic Stability in an Age of Crisis: Reassessing the Interwar Period* (Oxford University Press) and *Varieties of Democracy: Measuring Two Centuries of Political Change* (Cambridge University Press).

Jan Teorell is professor of political science at Lund University. He has twice won the Lijphart, Przeworski, and Verba Award for Best Dataset by the APSA Comparative Politics Section, and he is the author of *Determinants of Democratization* (Cambridge University Press, 2010) and co-author of *Varieties of Democracy* (Cambridge University Press, 2020). His work has appeared in journals such as the *American Political Science Review, Comparative Political Studies, European Journal of Political Research, Journal of Democracy, Governance, Political Research Quarterly, and Studies in Comparative International Development*. His research interests include political methodology, history and comparative politics, comparative democratization, corruption, and state making.

Michael Coppedge is Professor of Political Science and a Faculty Fellow of the Kellogg Institute at the University of Notre Dame and one of the Principal Investigators of the Varieties of Democracy project (V-Dem). He is the author of *Democratization and Research Methods* (Cambridge, 2012), *Strong Parties and Lame Ducks: Presidential Partyarchy and Factionalism in Venezuela* (Stanford, 1994) and co-author *of Varieties of Democracy: Measuring Two Centuries of Political Change* (Cambridge, 2020).

ORCID

John Gerring http://orcid.org/0000-0001-9858-2050
Carl Henrik Knutsen http://orcid.org/0000-0002-5470-9045
Svend-Erik Skaaning http://orcid.org/0000-0002-8587-8183
Matthew Maguire http://orcid.org/0000-0001-8964-9137
Jan Teorell http://orcid.org/0000-0002-0628-0096
Michael Coppedge http://orcid.org/0000-0002-4865-6901

Bibliography

Acharya, A., and H. J. Wall. "An Evaluation of the United Nations' Human Development Index." *Journal of Economic and Social Measurement* 20, no. 1 (1994): 51–65.

Anthopolos, R., and C. Becker. "Global Infant Mortality: Correcting for Undercounting." *World Development* 38, no. 4 (2010): 467–481.

Bernhard, M., T. Nordstrom, and C. Reenock. "Economic Performance, Institutional Intermediation, and Democratic Survival." *Journal of Politics* 63, no. 3 (2001): 775–803.

Besley, T., and M. Kudamatsu. "Health and Democracy." *American Economic Review* 96, no. 2 (2006): 313–318.

Bhargava, A., D. T. Jamison, L. J. Lau, and C. J. L. Murray. "Modeling the Effects of Health on Economic Growth." *Journal of Health Economics* 20, no. 3 (2001): 423–440.

Blaydes, L., and M. A. Kayser. "Counting Calories: Democracy and Distribution in the Developing World." *International Studies Quarterly* 55, no. 4 (2011): 887–908.

Boix, C. "Democracy, Development and the Public Sector." *American Journal of Political Science* 45, no. 1 (2001): 1–17.

Boix, C., M. K. Miller, and S. Rosato. "A Complete Data Set of Political Regimes, 1800–2007." *Comparative Political Studies* 46, no. 12 (2013): 1523–1554.

Brown, D. "Reading, Writing and Regime Type: Democracy's Impact on Primary School Enrollment." *Political Research Quarterly* 52, no. 4 (1999): 681–707.

Brown, D., and A. Mobarak. "The Transforming Power of Democracy: Regime Type and the Distribution of Electricity." *American Political Science Review* 103, no. 2 (2009): 193–213.

Bueno de Mesquita, B., A. Smith, R. M. Siverson, and J. M. Morrow. *The Logic of Political Survival*. Cambridge: MIT Press, 2003.
Casper, G., and C. Tufis. "Correlation versus Interchangeability: The Limited Robustness of Empirical Findings on Democracy Using Highly Correlated Data Sets." *Political Analysis* 11, no. 2 (2003): 196–203.
Chaudhury, N., J. Hammer, M. Kremer, K. Muralidharan, and F. H. Rogers. "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspectives* 20, no. 1 (2006): 91–116.
Cheibub, J. A., J. Gandhi, and J. R. Vreeland. "Democracy and Dictatorship Revisited." *Public Choice* 143, no. 1–2 (2010): 67–101.
Christoforou, A. "Social Capital and Human Development: An Empirical Investigation across European Countries." *Journal of Institutional Economics* 6, no. 2 (2010): 191–214.
Coppedge, M., A. Alvarez, and C. Maldonado. "Two Persistent Dimensions of Democracy: Contestation and Inclusiveness." *Journal of Politics* 70, no. 3 (2008): 632–647.
Coppedge, M., J. Gerring, S. I. Lindberg, S. Skaaning, J. Teorell, D. Altman, M. Bernhard, et al. *V-Dem [Country-Year/Country-Date] Dataset v7.1*. Varieties of Democracy (V-Dem) Project, 2017.
Coppedge, M., J. Gerring, S. I. Lindberg, S. Skaaning, J. Teorell, J. Krusell, K. L. Marquardt, et al. *V-Dem Methodology v7.1*. Varieties of Democracy (V-Dem) Project, 2017.
Costanza, R., M. Hart, S. Posner, and J. Talberth. "Beyond GDP: The Need for New Measures of Progress. " Pardee Papers #4. The Frederick S. Pardee Center for the Study of the Longer-Range Future at Boston University, 2009.
Dahl, R. A. *Democracy and its Critics*. New Haven: Yale University Press, 1989.
Dahl, R. A. *Polyarchy: Participation and Opposition*. New Haven: Yale University Press, 1971.
Dahlum, S., and C. H. Knutsen. "Do Democracies Provide Better Education? Revisiting the Democracy-Human Capital Link." *World Development* 94, no. 1 (2017): 186–199.
Deacon, R. T. "Public Good Provision under Dictatorship and Democracy." *Public Choice* 139, no. 1–2 (2009): 241–262.
Eterovic, D. S., and C. M. Sweet. "Democracy and Education in Twentieth-century Latin America." *Economics & Politics* 26, no. 2 (2014): 237–262.
Ferejohn, J. "Incumbent Performance and Electoral Control." *Public Choice* 50, no. 1–3 (1986): 5–25.
Filmer, D., and L. Pritchett. "The Impact of Public Spending on Health: Does Money Matter?" *Social Science and Medicine* 49, no. 10 (1999): 1309–1323.
Freedom House. *Freedom in the World* 2012. Accessed 5 January 2018. https://freedomhouse.org/report/freedom-world-2012/methodology#.VVe-b3Y4VaQ.
Fung, A., and E. O. Wright, eds. *Deepening Democracy: Institutional Innovations in Empowered Participatory Governance*. London: Verso, 2003.
Gandhi, J., and E. Lust-Okar. "Elections under Authoritarianism." *Annual Review of Political Science* 12 (2009): 403–422.
Gauri, V., and P. Khaleghian. "Immunization in Developing Countries: Its Political and Organizational Determinants." *World Development* 30, no. 12 (2002): 2109–2132.
Gerring, J., S. C. Thacker, and R. Alfaro. "Democracy and Human Development." *Journal of Politics* 74, no. 1 (2012): 1–17.
Ghobarah, H. A., P. Huth, and B. Russett. "Comparative Public Health: The Political Economy of Human Misery and Well-being." *International Studies Quarterly* 48, no. 1 (2004): 73–94.
Goertz, G. *Social Science Concepts: A User's Guide*. Princeton: Princeton University Press, 2006.
Hallerod, B., B. Rothstein, A. Daoud, and S. Nandy. "Bad Governance and Poor Children: A Comparative Analysis of Government Efficiency and Severe Child Deprivation in 68 Low and Middle-income Countries." *World Development* 48, no. 1 (2013): 19–31.
Hanson, J. K. "Democracy and State Capacity: Complements or Substitutes?" *Studies in Comparative International Development* 50, no. 3 (2015): 304–330.
Harding, R., and D. Stasavage. "What Democracy Does (and Doesn't do) for Basic Services: School Fees, School Inputs, and African Elections." *Journal of Politics* 76, no. 1 (2014): 229–245.
Huber, E., T. Mustillo, and J. D. Stephens. "Politics and Social Spending in Latin America." *Journal of Politics* 70, no. 2 (2008): 420–436.
Imai, K., L. Keele, D. Tingley, and T. Yamamoto. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105, no. 4 (2011): 765–789.

Inglehart, R., and C. Welzel. *Modernization, Cultural Change, and Democracy*. Cambridge: Cambridge University Press, 2005.
Jacobsen, J. "Revisiting the Modernization Hypothesis: Longevity and Democracy." *World Development* 67 (2015): 174–185.
Jensen, P. S., and M. K. Justesen. "Poverty and Vote Buying: Survey-based Evidence from Africa." *Electoral Studies* 33, no. 1 (2014): 220–232.
Keefer, P. "Clientelism, Credibility and the Policy Choices of Young Democracies." *American Journal of Political Science* 51, no. 4 (2006): 804–821.
Knutsen, C. H., J. Gerring, S.-E. Skaaning, J. Teorell, M. Maguire, M. Coppedge, and S. I. Lindberg. "Economic Development and Democracy: An Electoral Connection." *European Journal of Political Research* (2018). OnlineFirst.
Knutsen, C. H., and M. B. Rasmussen. "The Autocratic Welfare State: Old-age Pensions, Credible Commitments, and Regime Survival." *Comparative Political Studies* 51, no. 5 (2018): 659–695.
Kudamatsu, M. "Has Democratization Reduced Infant Mortality in sub-Saharan Africa? Evidence from Micro Data." *Journal of the European Economic Association* 10, no. 6 (2012): 1294–1317.
Lake, D. A., and M. A. Baum. "The Invisible Hand of Democracy: Political Control and the Provision of Public Services." *Comparative Political Studies* 34, no. 6 (2001): 587–621.
Levitsky, S., and L. Way. *Competitive Authoritarianism*. Cambridge: Cambridge University Press, 2010.
Lipton, M. *Why Poor People Stay Poor: Urban Bias in World Development*. Cambridge: Harvard University Press, 1977.
Mansfield, E. D., and J. Snyder. *Electing to Fight: Why Emerging Democracies Go to War*. Cambridge: MIT Press, 2005.
Marshall, M., T. Gurr, and K. Jaggers. *Polity IV Project: Dataset Users' Manual*. 2014. http://www.systemicpeace.org/inscr/p4manualv2013.pdf.
Mattila, M., P. Söderlund, H. Wass, and L. Rapeli. "Healthy Voting: The Effect of Self-reported Health on Turnout in 30 Countries." *Electoral Studies* 32, no. 4 (2013): 886–891.
McGillivray, M. "Measuring Non-economic Well-being Achievement." *Review of Income and Wealth* 51, no. 2 (2005): 337–364.
McGuire, J. W. *Wealth, Health, and Democracy in East Asia and Latin America*. Cambridge: Cambridge University Press, 2010.
Meltzer, A. H., and S. F. Richards. "A Rational Theory of the Size of Government." *Journal of Political Economy* 89, no. 5 (1981): 914–927.
Miller, M. K. "Democratic Pieces: Autocratic Elections and Democratic Development since 1815." *British Journal of Political Science* 45, no. 3 (2015): 501–530.
Miller, M. K. "Electoral Authoritarianism and Human Development." *Comparative Political Studies* 48, no. 12 (2015): 1526–1562.
Miller, G. "Women's Suffrage, Political Responsiveness, and Child Survival in American History." *Quarterly Journal of Economics* 123, no. 3 (2008): 1287–1327.
Mitchell, B. R. *International Historical Statistics. Africa, Asia & Oceania 1750–1993*. 3rd ed. Basingstoke: Macmillan, 1998.
Mitchell, B. R. *International Historical Statistics. Europe 1750–1993*. 4th ed. Basingstoke: Macmillan, 1998.
Mitchell, B. R. *International Historical Statistics. The Americas 1750–1993*. 4th ed. Basingstoke: Macmillan, 1998.
Munck, G. L. *Measuring Democracy: A Bridge between Scholarship and Politics*. Baltimore: Johns Hopkins University Press, 2009.
Nelson, J. M. "Elections, Democracy, and Social Services." *Studies in Comparative International Development* 41, no. 4 (2007): 79–97.
Pemstein, D., S. A. Meserve, and J. Melton. "Democratic Compromise: A Latent Variable Analysis of Ten Measures of Regime Type." *Political Analysis* 18, no. 4 (2010): 426–449.
Przeworski, A. *Democracy and the Limits of Self-government*. Cambridge: Cambridge University Press, 2010.
Przeworski, A., M. E. Alvarez, J. A. Cheibub, and F. Limongi. *Democracy and Development*. Cambridge: Cambridge University Press, 2000.
Rawlings, L. B., and G. M. Rubio. "Evaluating the Impact of Conditional Cash Transfer Programs." *World Bank Research Observer* 20, no. 1 (2005): 29–55.

Raworth, K., and D. Stewart. "Critiques of the Human Development Index: A Review." In *Readings in Human Development*, edited by S. Fukuda-Parr, and A. K. Kumar, 164–176. New Delhi: Oxford University Press, 2003.

Riley, J. C. "Estimates of Regional and Global Life Expectancy, 1800–2001." *Population and Development Review* 31, no. 3 (2005): 537–543.

Ross, M. "Is Democracy Good for the Poor?" *American Journal of Political Science* 50, no. 4 (2006): 860–874.

Rothstein, B. "Guilty as Charged? Human Well-being and the Unsung Relevance of Political Science." In *The Relevance of Political Science*, edited by G. Stoker, B. G. Peters, and J. Pierre, 84–103. London: Palgrave Macmillan, 2015.

Rothstein, B. *Quality of Government: Corruption, Social Trust, and Inequality in International Perspective*. Chicago: University of Chicago Press, 2011.

Schedler, A. "The Menu of Manipulation." *Journal of Democracy* 13, no. 2 (2002): 36–50.

Sen, A. "Mortality as an Indicator of Economic Success and Failure." *Economic Journal* 108, no. 446 (1998): 1–25.

Sen, A. *Poverty and Famines: An Essay on Entitlement and Deprivation*. Oxford: Clarendon Press, 1982.

Stasavage, D. "Democracy and Education Spending in Africa." *American Journal of Political Science* 49, no. 2 (2005): 343–358.

Teorell, J., M. Coppedge, S. Skaaning, and S. I. Lindberg. "Measuring Polyarchy across the Globe, 1900–2017." *Studies in Comparative International Development* 54 (2016): 71–95.

Truex, R. "The Myth of the Democratic Advantage." *Studies in Comparative International Development* 52, no. 3 (2017): 261–277.

ul Haq, M. *Reflections on Human Development*. New Dehli: Oxford University Press, 1995.

van de Walle, D. "The Distribution of Subsidies through Public Health Services in Indonesia, 1978–87." *World Bank Economic Review* 8, no. 2 (1994): 279–309.

Wang, Y., V. Mechkova, and F. Andersson. "Does Democracy Democracy Enhance Health? New Empirical Evidence 1900–2012." *Political Research Quarterly* 72, no. 3 (2019): 554–569.

Wigley, S., and A. Akkoyunlu-Wigley. "The Impact of Regime Type on Health: Does Redistribution Explain Everything?" *World Politics* 63, no. 4 (2011): 647–677.

Wright, J. "Do Authoritarian Institutions Constrain? How Legislatures Affect Economic Growth and Investment." *American Journal of Political Science* 52, no. 2 (2008): 322–343.

Zweifel, T. D., and P. Navia. "Democracy, Dictatorship, and Infant Mortality Revisited." *Journal of Democracy* 14, no. 3 (2003): 90–103.