In the Mood for Democracy?
Democratic Support as Thermostatic Opinion

Online Supplementary Materials

Contents

1. Measuring Democratic Mood

1.1. Survey Questions Used to Measure Democratic Mood

1. Three statements items

   1.1. Which of these three statements is closest to your own opinion? (AfroBarometer)
   - Democracy is preferable to any other kind of government
   - Under some circumstances, an authoritarian government can be preferable to a democratic one
   - For someone like me, it does not matter what kind of government we have.

   1.2. Which of the following statements comes closest to your own opinion? (AsianBarometer)
   - For people like me, it does not matter whether we have a democracy
   - Under some circumstances, an authoritarian government can be preferable
   - Democracy is always preferable to any other kind of government

   1.3. Which of these three statements is closest to your own opinion? (ArabBarometer)
   - Democracy is preferable to any other kind of government
   - Under some circumstances, an nondemocratic government can be preferable
   - For people like me, it does not matter what kind of government we have

   1.4. Here are three opinions about political systems. Which one comes closest to your own way of thinking? (EuroBarometer)
   - Democracy is the best political system in all circumstances
   - In certain circumstances a dictatorship could be a good thing
   - Whether we live in a democracy or under a dictatorship makes no difference to people like me

   1.5. With which of the following phrases are you in most agreement? (Latin American Public Opinion Project)
   - For people like me, it doesn't matter whether a regime is democratic or non-democratic
   - Democracy is preferable to any other type of government
   - Under some circumstances an authoritarian government can be preferable to a democratic one

   1.6. Which of the following statements do you agree with most? (LatinoBarometer)
   - Democracy is preferable to any other kind of government
   - In certain situations, an authoritarian government can be preferable to a democratic one
   - To people like me it doesn't matter whether we have a democratic government or a non-democratic government

   1.7. With which of the following statements do you agree most? (New Democracies Barometer)
   - Democracy is preferable to any other kind of government
   - Under some circumstances, an authoritarian government can be preferable to a democratic one
   - For people like me, it does not matter whether we have a democratic or a non- democratic regime

   1.8. Which of these three statements is closest to your own opinion? (Pew Global Attitudes)
   - Democracy is preferable to any other kind of government
   - Under some circumstances, an authoritarian government can be preferable to a democratic one
   - For someone like me, it does not matter what kind of government we have

   1.9. Which one of the following three statements do you agree with most? (South Asian Barometer)
   - Democracy is preferable to any other kind of government
   - In certain situations, a dictatorial government can be preferable to a democratic one
   - It doesn't matter to people like me whether we have democratic or non-democratic governance

2. "Churchill" items

    2.1. Democracy may have its problems, but it is better than any other form of government. To what extent do you agree or disagree? (ArabBarometer)

    2.2. Do you agree or disagree with the following statement: Democracy may have its problems, but it is still the best form of government (AsianBarometer).

    2.3. Please tell me how strongly you agree or disagree with the following statement: Democracy may have problems but it's better than any other form of government (Comparative Study of Electoral Systems)

    2.4. Democracy may have its problems, but it is better than any other form of government. To what extent do you agree or disagree? (European Values Survey)

    2.5. With which of the following phrases do you most agree: in general, despite its problems, democracy is the best form of government, there are other forms of government that can be just as good or even better than democracy, don't know (Latin American Public Opinion Project)

    2.6. Do you strongly agree, agree, disagree or strongly disagree with the following statements: Democracy may have problems but it is the best system of government (LatinoBarometer)

    2.7. Democracy may have its problems, but it is better than any other form of government. To what extent do you agree or disagree? (World Values Survey)

3. Strong leader items

    3.1. There are many ways to govern a country. Would you disapprove or approve of the following alternatives? Elections and Parliament are abolished so that the president can decide everything. (AfroBarometer)

    3.2. I will describe different political systems to you, and I want to ask you about your opinion of each one of them with regard to the countrys governance for each one would you say it is very good, good, bad, or very bad? (ArabBarometer)

    3.3. I will describe different political systems to you, and I want to ask you about your opinion of each one of them with regard to the countrys governance for each one would you say it is very good, good, bad, or very bad? (ArabBarometer)

    3.4. I'm going to describe various types of political systems. Please indicate for each system whether you think it would be very good, fairly good or bad for this country. Governance by a powerful leader without the restriction of parliament or elections (AsiaBarometer)

    3.5. Best to get rid of Parliament and elections and have a strong leader who can quickly decide everything. What do you think? (AsianBarometer)

    3.6. I'm going to describe various types of political systems and ask what you think about each as a way of governing this country. For each one, would you say it is a very good, fairly good, fairly bad or very bad way of governing this country? Having a strong leader who does not have to bother with parliament and elections (European Values Survey)

    3.7. On some occasions, democracy doesn't work. When that happens there are people that say we need a strong leader who doesn't have to be elected through voting. Others say that even if things don't function, democracy is always the best. What do you think? (Latin American Public Opinion Project)

    3.8. There are people who say that we need a strong leader that does not have to be elected. Others say

that although things may not work, electoral democracy, or the popular vote, is always best. What do you think? (Latin American Public Opinion Project)

3.9. Best to get rid of Parliament and elections and have a strong leader who can quickly decide everything. What do you think? (New Democracies Barometer)

3.10. Some feel that we should rely on a democratic form of government to solve our country's problems. Others feel that we should rely on a leader with a strong hand to solve our country's problems. Which comes closer to your opinion? (Pew Global Attitudes)

3.11. There are different ways in which a country may be governed. I will read out some suggestions. For each of these would you say that you strongly agree, agree, disagree or strongly disagree? We should have a strong leader who does not have to bother about elections (South Asian Barometer)

3.12. I'm going to describe various types of political systems and ask what you think about each as a way of governing this country. For each one, would you say it is a very good, fairly good, fairly bad or very bad way of governing this country? Having a strong leader who does not have to bother with parliament and elections (World Values Survey)

4. Military rule items

4.1. There are many ways to govern a country. Would you disapprove or approve of the following alternatives? The army comes in to govern the country (AfroBarometer)

4.2. I'm going to describe various types of political systems. Please indicate for each system whether you think it would be very good, fairly good or bad for this country – Military government (AsiaBarometer)

4.3. The army should govern the country. What do you think? (AsianBarometer)

4.4. I'm going to describe various types of political systems and ask what you think about each as a way of governing this country. For each one, would you say it is a very good, fairly good, fairly bad or very bad way of governing this country? Having the army rule (European Values Survey)

4.5. The army should govern the country. What do you think? (New Democracies Barometer)

4.6. I'm going to describe various types of political systems and ask what you think about each as a way of governing our country. For each one, would it be a very good, somewhat good, somewhat bad or very bad way of governing this country? The military rules the country (Pew Global Attitudes)

4.7. There are different ways in which a country may be governed. I will read out some suggestions. For each of these would you say that you strongly agree, agree, disagree or strongly disagree? The country should be governed by the Army (South Asian Barometer)

4.8. I'm going to describe various types of political systems and ask what you think about each as a way of governing this country. For each one, would you say it is a very good, fairly good, fairly bad or very bad way of governing this country? Having the army rule (World Values Survey)

5. One party rule items

5.1. There are many ways to govern a country. Would you disapprove or approve of the following alternatives? Only one political party is allowed to stand for election and hold office (AfroBarometer)

5.2. There are many ways to govern a country. Would you disapprove or approve of the following alternatives? Only one political party is allowed to stand for election and hold office (AsianBarometer)

6. Evaluate democracy items

6.1. I will describe different political systems to you, and I want to ask you about your opinion of each one of them with regard to the country's governance – for each one would you say it is very good, good, bad, or very bad? A democratic political system (public freedoms, guarantees equality in political and civil rights, alternation of power, and accountability and transparency of the executive authority). (ArabBarometer)

6.2. I'm going to describe various types of political systems. Please indicate for each system whether you think it would be very good, fairly good or bad for this country – A democratic political system (AsiaBarometer)

6.3. I'm going to describe various types of political systems and ask what you think about each as a way of governing this country. For each one, would you say it is a very good, fairly good, fairly bad or very bad way of governing this country? Having a democratic political system (European Values Survey)

6.4. I'm going to describe various types of political systems and ask what you think about each as a way of governing our country. For each one, would it be a very good, somewhat good, somewhat bad or very bad way of governing this country? A democratic system where representatives elected by citizens decide what becomes law (Pew Global Attitudes)

6.5. I'm going to describe various types of political systems and ask what you think about each as a way of governing this country. For each one, would you say it is a very good, fairly good, fairly bad or very bad way of governing this country? Having a democratic political system (World Values Survey)

7. Elections items

7.1. There are different ways in which a country may be governed. I will read out some suggestions. For each of these would you say that you strongly agree, agree, disagree or strongly disagree? The country should be governed by those chosen by the people in a fair election (South Asian Barometer)

8. Suitability items

8.1. Suppose there was a scale from 0-10 measuring the extent to which democracy is suitable for your country, with 0 meaning that democracy is absolutely inappropriate for your country and 10 meaning that democracy is completely appropriate for your country. To what extent do you think democracy is appropriate for your country? (ArabBarometer)

8.2. Here is a similar scale of 1 to 10 measuring the extent to which people think democracy is suitable for our country. If 1 means that democracy is completely unsuitable for [name of country] today and 10 means that it is completely suitable, where would you place our country today? (AsianBarometer)

8.3. How suitable is democracy for our country - very suitable, suitable, not suitable or not at all suitable? (South Asian Barometer)

9. Importance items

9.1. How important for you to live in democratically governed country? (European Social Survey)

9.2. How important is it to you to live in a country where honest elections are held regularly with a choice of at least two political parties? Is it very important, somewhat important, not too important or not important at all? (Pew Global Attitudes)

9.3. How important is it for you to live in a country that is governed democratically? On this scale where 1 means it is "not at all important" and 10 means "absolutely important" what position would you choose? (World Values Survey)

10. Desire for democracy items

10.1. To what extent do you want our country to be democratic now? (AsianBarometer)

1.2. Excluded Survey Items

Kurzman (2014) notes that there appears to be serious translation errors in some of the questions relating to democracy in the 3rd and 4th waves of the World Values Survey. For example, as he describes, the Indonesian survey of 2001 asked respondents their opinion on having military *rules*, rather than military rule. The vast majority of Indonesians unsurprisingly favored having rules. Survey responses from the following items-year-country combinations from the World Values Survey were therefore excluded from the analysis due to evidence of, or suspicion of, poor translations and severe bias:

- Vietnam: Military rule 2001; Strong leader 2001
- Albania: Military rule 1998
- Indonesia: Military rule 2001 & 2006
- Iran: Military rule 2000; Strong leader 2000 & 2005
- India: Strong leader, all years.
- Pakistan: Military rule 1996 & 2001; Strong leader 1996 & 2001
- Kyrgyzstan: Strong leader 2003 & 2011
- Romania: Strong leader 1998, 2005 & 2012
- Egypt: Strong leader 2012

In addition, responses to the following sets of items were not included when measuring democratic mood:

- Items measuring satisfaction with the performance of democracy (e.g., "on the whole, how satisfied or dissatisfied are you with the way democracy works in [country]?")
- Items tapping evaluations of the political and economic performance of democracy (e.g., "Which of the following statements comes closer to your own view? Democracy is capable of solving the problems of our society; democracy can not solve our society's problems.")
- Items measuring respondents' understandings of the term "democracy" (e.g., "For each of the following things, how essential do you think it is as a characteristic of democracy? – Governments tax the rich and subsidize the poor.").
- Items gauging trust in national political institutions (e.g., "I'm going to name a number of institutions. For each one, please tell me how much trust you have in them. – Parliament").

1.3. Microlevel Coding of Survey Responses

The latent trait measurement model takes a binomial response, which requires two pieces of information: the number of "trials" and the number of these that were "successful." I therefore gathered two quantities for each survey question. First, the number of respondents asked each

relevant survey question (this was usually, but not always, the full sample size). Second, the number of respondents providing a response that was supportive of democracy. This may include the response "democracy is preferable to any other kind of government" in the three statements question, disagreeing that the military should rule, or offering a response above the median on a 0-10 scale for the question regarding the importance of democracy to the respondent. It follows that all other possible responses (i.e., the difference between the sample size and the number of supportive respondents) were treated, similarly, as not supportive of democracy. These non-supportive respondents may have actively opposed democracy, (e.g., "an authoritarian government can be preferable to a democratic one"), chosen an intermediate response (e.g., "for someone like me, it does not matter what kind of government we have"), responded with "don't know," or refused to provide any response.

1.4. Latent Variable Model

Democratic mood is measured using the dynamic Bayesian latent trait model developed by Claassen (2019). In particular, I use his "model 5," which performed best in validation tests using a held-out dataset. Although Claassen (2019) also applies this model to the topic of democratic support, I use an expanded dataset of survey responses here (see also **?**).

The observed number of respondents y supporting democracy for each country i, year t, and survey item k is modeled as a binomial distributed count:

$$y_{ikt} \sim \text{Binomial}(s_{ikt},\ \pi_{ikt}).$$

A beta prior is then used to model the probability parameter π. This produces a beta-binomial distribution, which allows for some additional dispersion in the observed survey responses beyond that induced by sampling alone:

$$\pi_{ikt} \sim \text{Beta}(\alpha_{ikt},\ \beta_{ikt}).$$

The two shape parameters of the beta distribution can be reparameterized to an expectation parameter η and a dispersion parameter ϕ:

$$\begin{aligned}\alpha_{ikt} &= \phi\eta_{ikt}\\ \beta_{ikt} &= \phi(1-\eta_{ikt}).\end{aligned}$$

The expectation parameter η is then modeled as a function of the latent country-year estimates θ, item parameters λ, and item-country parameters δ. The item parameters adjust for the effects of item-specific bias, while the item-country parameters adjust for "non-equivalence" bias.

$$\begin{aligned}\eta_{ikt} &= \text{logit}^{-1}(\lambda_k + \delta_{ik} + \theta_{it})\\ \lambda_k &\sim \text{N}(\mu_\lambda,\ \sigma^2_\lambda)\\ \delta_{ik} &\sim \text{N}(0,\ \sigma^2_\delta)\end{aligned}$$

Finally, the latent opinion estimates are allowed to evolve smoothly over time by adding a dynamic linear model:

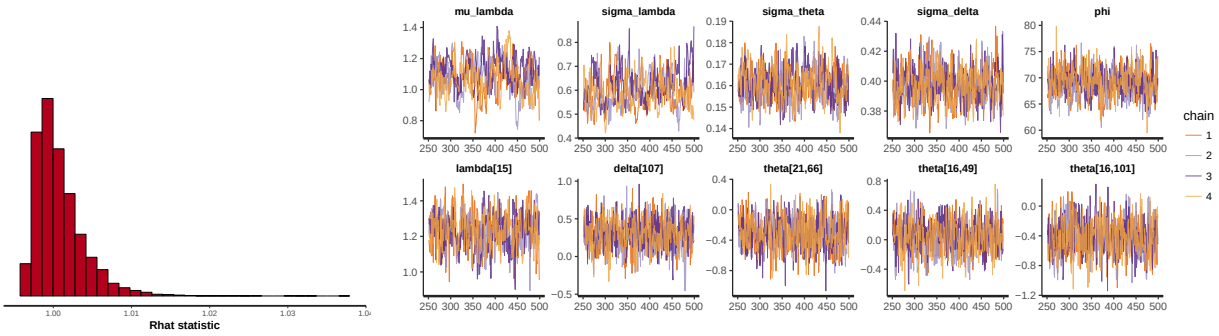
$$\theta_{it} \sim \text{N}(\theta_{i,t-1},\ \sigma^2_\theta).$$

The estimated variances are given weakly-informative half-Cauchy priors, e.g., $\sigma_\lambda \sim C^+(0,\ 2)$ (and similarly for σ_δ, σ_γ, and σ_θ). The expectation of the item intercepts μ_λ is given a N(1, 2) prior while the dispersion parameter ϕ, receives a gamma(4, 0.1) prior. The initial value of latent opinion for each country θ_{i1} receives a N(0, 1) prior. Finally, I identified the model by fixing the first item intercept λ_1 at a value of 1.

1.5. Estimating and Checking the Model

The model is estimated using Bayesian Markov-Chain Monte Carlo (MCMC) methods via Stan software, which implements Hamiltonian Monte Carlo sampling (Carpenter et al. 2017; Stan Development Team 2017). Four parallel chains were run for 500 samples each, with the first 250 samples in each chain used for warm up, and discarded. The remaining 1,000 samples of the posterior density were saved and analyzed further. This number of iterations proved to be more than sufficient for convergence, with the Gelman-Rubin R-hat diagnostic reaching a value close to one for all parameters (Figure S1). Traceplots (also Figure S1) of several parameters indicate the convergence of the four chains for several key parameters.

Figure S1. Model convergence

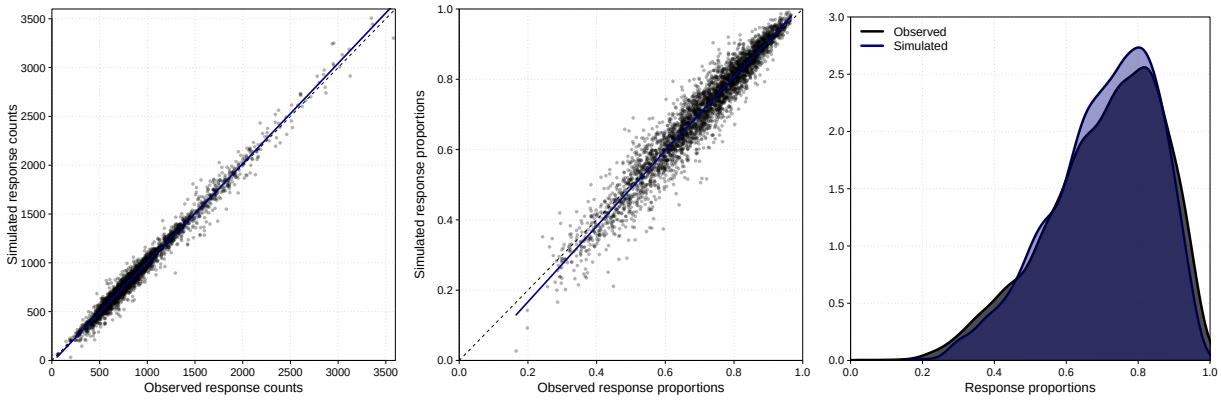


First plot shows the distribution of the Gelman-Rubin rhat statistic for all parameters. Second set of plots are traceplots of selected parameters

I further checked the model by using it to create a simulated dataset of responses (both as counts and as proportions), and plotted these against the actual survey responses. As the first two panels of Figure S2 show, there is a close correspondence between the simulated data $\tilde{y}_{ikt}$ and the actual responses y_{ikt}. The third panel then demonstrates that the distributions of the simulated response proportions follows that of the actual response proportions. These posterior predictive checks therefore suggest that the model fits the observed survey responses.

The model used to measure mood – Claassen's (2019) model 5 – includes item intercepts (or difficulty parameters in IRT terminology) but not slopes (i.e., discrimination parameters). Claassen's model 6, however also includes slopes / discrimination parameters. Although their inclusion does not improve the performance of the model, they do have a diagnostic use. They allow us to plot the item characteristic curves (ICCs). ICCs display the the relationship between the proportion of a national sample responding positively toward democracy (y-axis) and the latent estimates of mood (x-axis). The vertical alignment of the curves is governed by the item intercepts λ, while the steepness of the curves is governed by the item slopes γ. To aid in interpretation, I

Figure S2. Posterior predictive checks



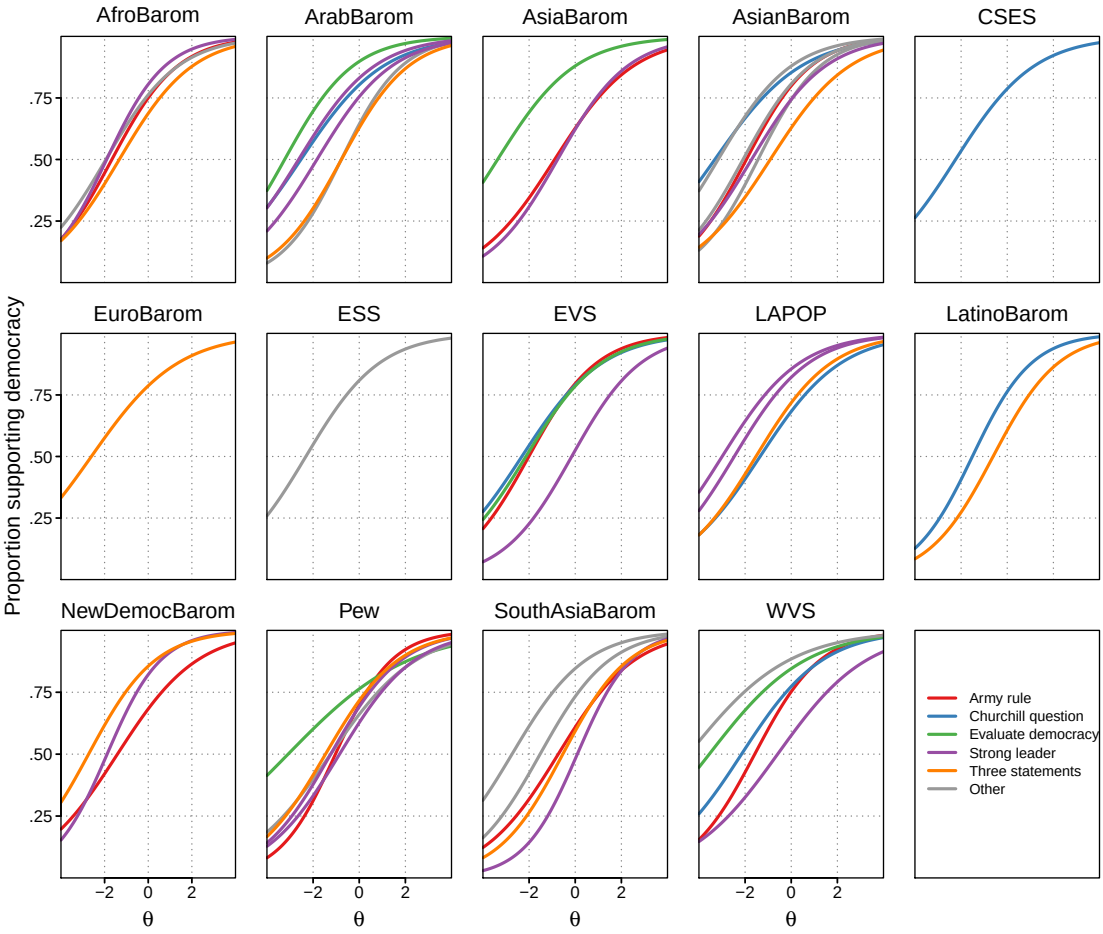
group the items by their survey project, and use varying colors to identify the main question wording approach. These ICCs allows us to verify the performance of particular items. In particular, I can evaluate whether items fit the latent construct (in factor analytic terms: whether they "load" on the factor). These item characteristic curves are plotted in Figure S3.

Turning to Figure S3, one can see that all the items display a positive relationship between the latent quantity and the observed responses. All items, in other words, have positive slopes. In addition, most items have slopes of similar magnitude. These are welcome findings, as they indicate that the included survey items do indeed measure the latent construct. It is not a particularly surprising finding, however, as items were selected based on the results of previous analyses of microlevel survey data (e.g. Klingemann 1999; Mattes and Bratton 2007; Rose, Mishler, and Haerpfer 1998). Items that bore some superficial resemblance to democratic mood, but which did not display a deeper empirical relationship with this latent variable were not included in the analysis in the first place.

There are nonetheless a few items with weaker slopes and therefore more tenuous relationships with latent mood. First is the "army rule" item from the New Democracies Barometer and second is the "evaluate democratic political system" item from the AsiaBarometer. Another three come from the World Values Surveys: the items asking respondents to rate the "importance of living in a democracy," and to evaluate a "strong leader" and a "democratic political system." This latter survey question has previously been criticized as offering only "lip service" to democracy, rather than deeply-rooted support (Inglehart 2003). In two out of the three survey projects in which it is employed, this type of question does indeed show a weaker relationship with latent democratic mood.

In contrast, two widely-used approaches for measuring support for democracy – the "three statements" and "evaluate army rule" survey questions – perform well across regions and survey projects. Both have pronounced positive slopes, indicating that such questions allow researchers to discriminate between respondents who favor democracy and those who do not. Indeed, national samples show widely-varying levels of agreement with these items as their underlying democratic mood increases: at low levels of mood (two standard deviations below the mean), 25% to 40% of respondents tend to offer the democratic responses to the three statements questions; at high levels (two standard deviations above the mean), around 85% do so.

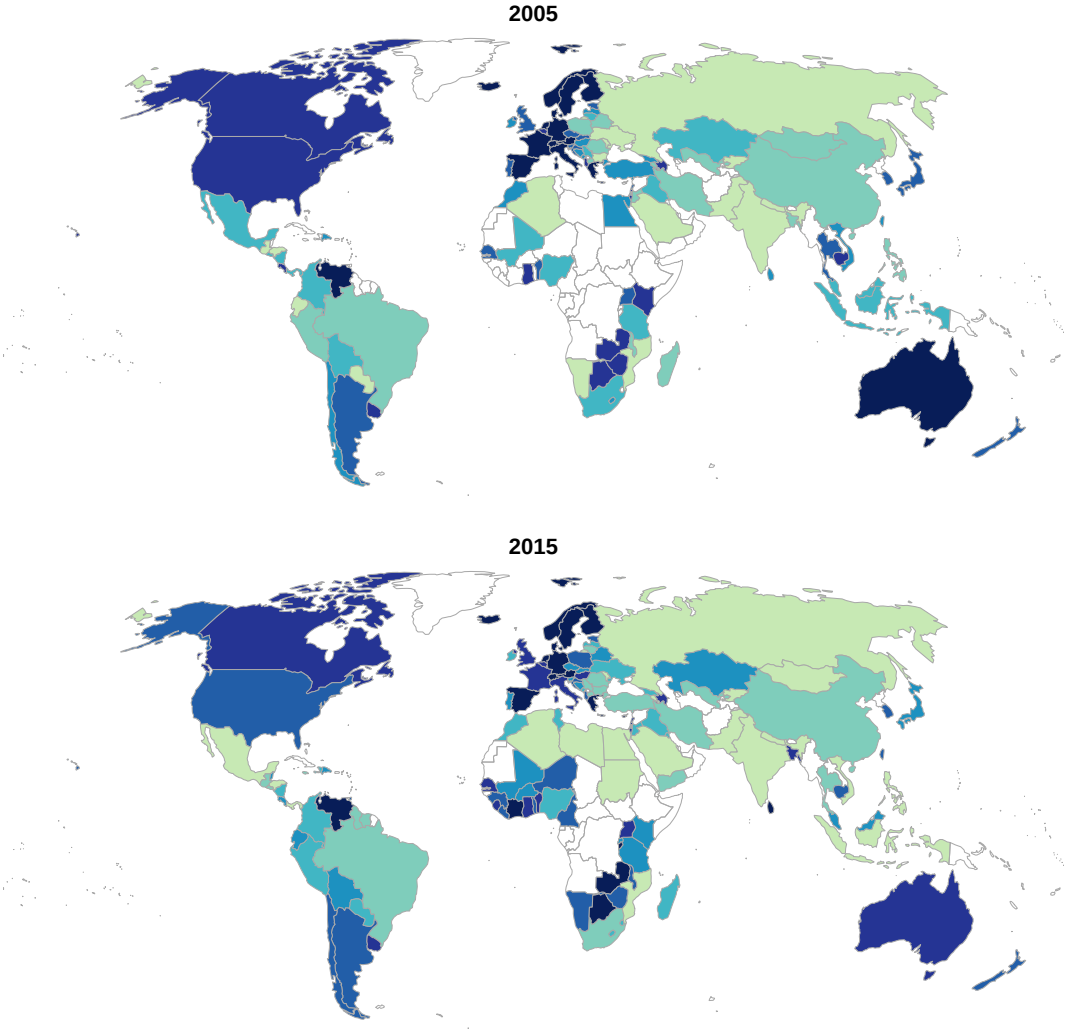
Figure S3. Item characteristic curves



1.6. Validating the Estimates

Are the estimates of latent mood valid? Do they measure what they purport to measure? This section describes various tests of validation that have been conducted. First, Claassen (2019) develops six versions of the latent variable model and chooses among these using cross-validation. In particular, he fits the six models to a 75% training dataset of cross-national aggregate opinion and estimates the error and credible interval coverage of these six models when applied to the 25% held-out, test dataset. He furthermore compares the six models to an alternative that has been proposed for estimating smooth opinion panels – Caughey and Warshaw's (2015) "Dynamic Group-level Item Response Theory (DGIRT)" model – but also two naive, baseline estimates – the country and items averages. The most accurate model ("model 5" – the model used in this article) has 45% less bias than using the simple country average to predict held-out opinion, 31% less bias than the DGIRT model, and 26% less bias than the simplest model which Claassen develops and tests ("model 1"). In absolute terms, the predicted responses produced by model 5 deviate from the observed responses in the held-out dataset by six percentage points. This is a satisfactory level of bias when one considers that cross-national public opinion data are gathered by different survey projects using a variety of methodologies and in dozens of languages and countries, and are

Figure S4. Mapping Mood



Estimates of democratic mood in 2005 and 2015. Dark blue shades indicate countries with the most democratic mood; light green, those with the most autocratic mood. Countries for which no estimates are available in the given year are shaded white.

furthermore fragmented over space and time and fractured across the numerous survey items that are used to measure any given opinion construct.

Figure S4 then provides evidence of the construct validity of the mood estimates (see also Claassen 2019). These choropleths show the estimates of mood for two years, 2005 and 2015. They support three conclusions. First, mood is more democratic where democracy has a long history – Western Europe, North America, and Australasia – and is lower in Asia and North Africa, where the heritage is authoritarian. Latin America, with a chequered history of democracy, has a correspondingly ambivalent attitude to democracy (Venezuela is the major exception, which is discussed in the article). This geographic pattern of support is consistent with existing research (e.g., Fuchs-Schündeln and Schündeln 2015; Klingemann 1999; Rose, Mishler, and Haerpfer 1998;

Norris 2011). Second, a comparison of the two choropleths reveals that mood turned somewhat against democracy between 2005 and 2015 in some Western bastions of democracy, notably the USA, France, and Australia. This trend is also consistent with other recent analyses (Foa and Mounk 2016; 2017). Third, quite the opposite trend is revealed in Africa, particularly in the South, East, and West: high – and indeed, rising – levels of democratic support in 2015. This pattern is yet another that echoes the findings of previous research (e.g., Bratton and Houessou 2014). In sum, the estimates of mood are generally consistent with extant research on the geographic and temporal variation in support for democracy, which demonstrates their construct validity.

2. Measuring Democracy

2.1. The V-Dem Measurement Approach

With at least ten different cross-national measures in existence (Pemstein, Meserve, and Melton 2010), democracy is perhaps the most-measured concept in social science. The sheer number of measures indicates the fragmentation of the market, with no single measure of democracy having won universal acceptance in the discipline. Indeed, even the three most widely-used measures – the continuous Polity and Freedom House (FH) indices and Przeworski and colleagues' dichotomous Democracy-Dictatorship (DD) indicator – have been subject to trenchant criticisms.

In particular, the minimalist Polity and DD measures have been criticized for neglecting important features of democracy, such as the universal right to vote and more generally participation in politics (Munck and Verkuilen 2002; Teorell et al. 2018). In addition, although the dichotomous nature of the DD indicator affords a conceptual clarity, Elkins (2000) show that it leads to lower validity and reliability. Continuous measures suffer from their own problems. Pemstein et al's (2010) latent variable model reveals that both the Polity and FH scales suffer from a fairly high degree of measurement error. Both scales also cluster at extreme values, raising the concern that they are not finely-grained enough to pick up variation in highly democratic, or highly autocratic countries (Alexander and Welzel 2011; Teorell et al. 2018). Indeed, as Pemstein, Meserve, and Melton (2010) show, the Polity scale exhibits a pronounced nonlinear relationship with a latent measure of democracy (see also Figure S5 below). Criticism has furthermore been directed at the method by which the various aggregate scales are assembled out of indicators of specific aspects of democracy. Munck and Verkuilen (2002) claim that FH assigns scores to indicators in an opaque fashion. Pemstein, Meserve, and Melton (2010) demonstrate that the usage of small numbers of judges who rate democracy for all countries results in lower measurement validity than regionally-specific measures scored by regional experts.

Such issues of conceptualization and measurement have concerned scholars for some time. Without a clear solution to this impasse, empirical scholars have stuck to using one or two of the big three measures. Fortunately we are now in a position to address these criticisms because of new measures of democracy provided by the Varieties of Democracy (V-Dem) project (Lindberg et al. 2014; Teorell et al. 2018; V-Dem Institute 2018*b*). The advantages of the V-Dem measures are fourfold.

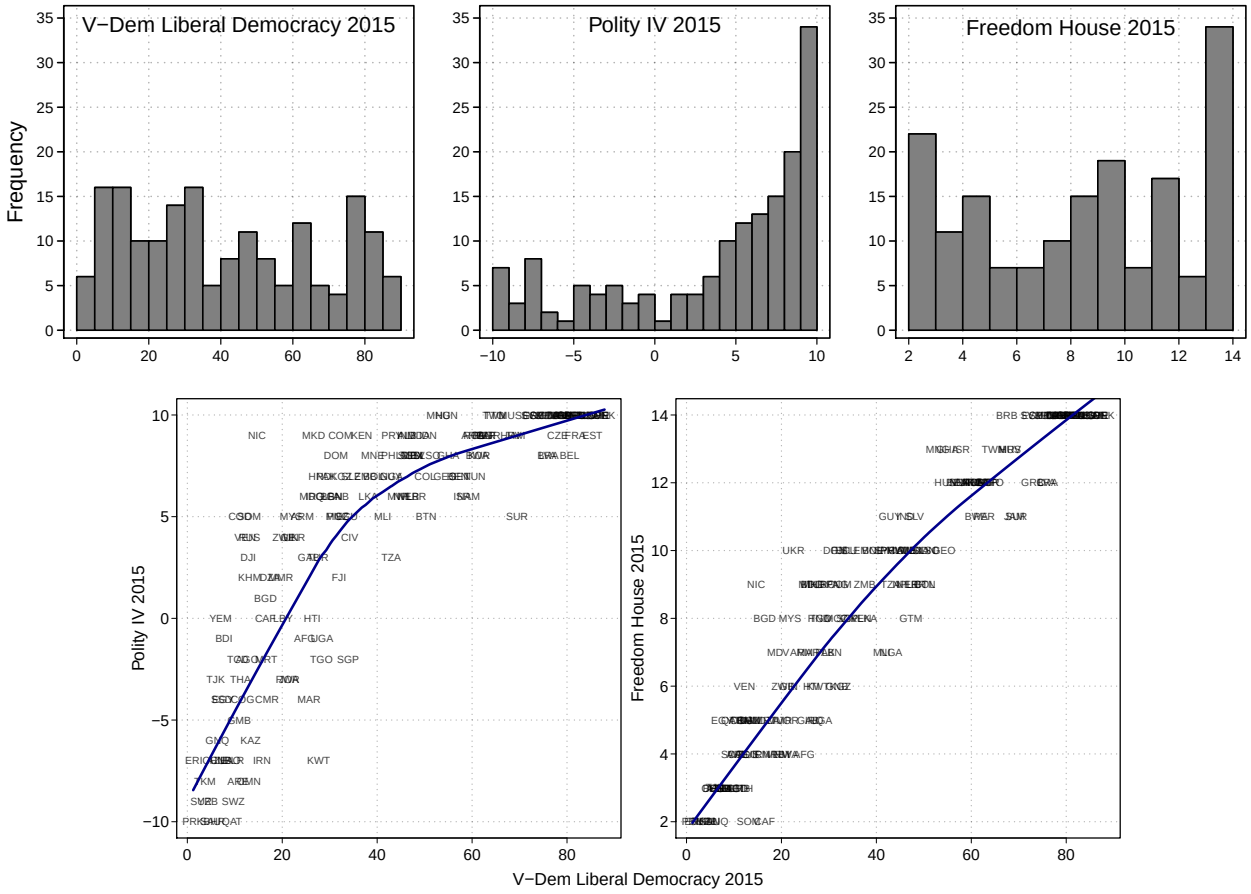
First, V-Dem derive five conceptualizations of democracy from the extensive political theory literature on this topic: electoral, liberal, participatory, egalitarian, and deliberative democracy (Lindberg et al. 2014). Other, extant measures of democracy are based far more loosely, if at all, on extant theoretical conceptualizations. In particular, V-Dem's measure of electoral democracy – which is used as a stand-alone measure in this paper but also forms the main component of their liberal democracy index – is alone among democracy measures in being explicitly derived from Dahl's (1989) concept of polyarchy (Teorell et al. 2018).

Second, V-Dem then measures each of these conceptualizations in a far more finely-grained fashion than existing democracy indices. They break each conceptualization into its major components, e.g., electoral democracy has five: (1) the extent to which the chief executive and legislature are appointed through popular elections; (2) the integrity of those elections; (3) the share of adult citizens with the right to vote in those elections; (4) freedom of association; and (5) freedom of expression (V-Dem Institute 2018*a*). Each of these components is then disggregated further into

sub-components, most measured with multiple indicators. Ultimately V-Dem's liberal democracy index is measured with 69 indicators, compared with the 22 used for the combined Freedom House scale and 4 used for Przeworski and colleagues' Democracy-Dictatorship indicator.

Third, considerable thought is put into the way in which these indicators are combined into subcomponents, components, and ultimately each index. Items are combined using Bayesian factor analyses, which allows V-Dem to actually test if the items load together as the theory suggests (V-Dem Institute 2018*c*). The sub-components are then combined into indices of democracy in either a multiplicative or additive fashion, as demanded by the particular conceptualization of democracy that is of interest.

Figure S5. Comparisons of V-Dem, Polity, and Freedom House Democracy Measures in 2015



Finally, V-Dem use country experts to code their indicators, rather than the in-house coders used by the Polity and Freedom House projects. Country experts are well placed to rate various aspects of democracy in that country with which they are most familiar. Democratic institutions and rights, and the extent to which these are protected or violated, are often a matter of public record (e.g., academic studies, news reports). Such records will be readily accessible to country experts. Indeed, the Pemstein, Meserve, and Melton (2010) analysis of ten democracy measures suggests that those relying on country experts (e.g., the Bowman, Lehoucq, and Mahoney (2005) measures of democracy in Central America) have lower error variance than those relying on generalists (e.g., Polity and Freedom House) as well as those using objective data (Vanhanen's index).

In addition, V-Dem uses, on average, five independent ratings from different country experts for all indicators. Errors – which are inevitable – are therefore reduced. Indeed, their Bayesian latent variable models allows them to identify (and adjust for) the reliability of particular raters.

2.2. Comparing Freedom House, Polity and V-Dem Democracy Indices

Figure S5 compares the V-Dem index with the combined and reversed FH and the combined Polity IV indices, all for 2015. As the histograms show, both FH and Polity exhibit clustering at the top of their respective scales: both are unable to discriminate between countries with very high and merely moderately high levels of democratic rights and institutions. The V-Dem measure, in contrast, does not suffer from this problem.

The scatterplots in the lower part of Figure S5 then demonstrate that the Polity measure, in particular, has a nonlinear relationship with V-Dem liberal democracy. Although Polity provides fairly fine-grained measures of autocracy, it fails to distinguish between countries that are fully democratic and those that are partial democracies (see also Munck and Verkuilen 2002; Pemstein, Meserve, and Melton 2010; Teorell et al. 2018). The scatterplot also reveals that the FH measure is ordinal rather than truly continuous.

2.3. Indicators Used by V-Dem to Measure Liberal Democracy

See V-Dem Institute (2018*a*) for further details on these indicators, including their wording, response sets, and methods by which they are aggregated into indices.

1. Electoral democracy index (i.e., majoritarian democracy)

   1.1. Share of population with suffrage

   1.2. Clean elections index
   - election management body autonomy
   - election management body capacity
   - election voter registry
   - election vote buying
   - election other voting irregularities
   - election government intimidation
   - election other electoral violence
   - election free and fair

   1.3. Elected officials index
   - head of state directly elected
   - legislature is popularly elected
   - head of state appointed by the legislature, or approval of legislature necessary
   - head of government appointed by legislature, or approval of legislature necessary
   - head of government appointed by head of state
   - head of government directly elected

   1.4. Freedom of Expression and Alternative Sources of Information index
   - media censorship effort
   - harassment of journalists
   - media bias
   - media self-censorship

- print/broadcast media critical
- print/broadcast media perspectives
- freedom of discussion for men/women
- freedom of academic and cultural expression

1.5. Freedom of association thick index
- party ban
- barriers to parties
- opposition parties autonomy
- CSO entry and exit
- CSO repression
- elections multiparty

2. Liberal component index (i.e., minoritarian democracy)

2.1. Judicial constraints on the executive index
- executive respects constitution
- compliance with judiciary
- compliance with high court
- high court independence
- lower court independence

2.2. Legislative constraints on the executive index
- legislature questions officials in practice
- executive oversight
- legislature investigates in practice
- legislature opposition parties

2.3. Equality before the law and individual liberty index
- rigorous and impartial public administration
- transparent laws with predictable enforcement
- access to justice for men/women
- property rights for men/women
- freedom from torture
- freedom from forced labor for men/women
- freedom of religion
- freedom of foreign movement
- freedom of domestic movement for men/women

3. Liberal democracy index

3.1. Electoral democracy index

3.2. Liberal component index

3. Calculating Long-Run Effects

To estimate the long run effects of democracy on mood using the error-correction models, I use a method of simulation inspired by that of Williams and Whitten (2012). First, I set all independent variables to some moderate value and allow the system of equations to run for 100 years, until predicted values of change in democracy stabilize. The level of democracy is then increased from half a standard deviation below to half a standard deviation above the mean, and the system of equations is allowed to run for 30 more years. These 30 simulated years of data are those depicted in the plots in the article. To capture the uncertainty inherent in the model, I take 10,000 draws from a multivariate normal distribution with expectation being the vector of model coefficients and variance being the robust covariance matrix: $\tilde{\Theta} \sim MVN(\Theta, \Sigma)$. When predicting effects using each of the i =10,000 simulated vectors of k coefficients, I add additional noise as estimated by the regression standard error: $\tilde{Y}_i \sim N(X_k\tilde{\Theta}_{ki}, \sigma)$. Finally, the mean value of $\tilde{Y}_i$ is used as the point estimate for that year, and the 0.025 and 0.975 quantiles of $\tilde{Y}_i$ as the lower and upper confidence bounds.

4. Additional Tables

Table S1. Descriptive Statistics

Variable	Mean	Std. Dev.	Min.	Max.	*N* Miss.
Year			1988	2017	0
Democratic mood	0.03	0.89	−2.10	2.74	0
Liberal democracy	0.39	0.92	−1.36	1.74	0
log GDP per capita	9.22	1.07	6.27	11.55	0
Corruption	−0.04	1.00	−2.51	1.84	331
Gini index	36.73	8.65	18.80	58.90	357
Education	8.53	2.87	1.10	14.10	95

Year refers to the year in which the dependent variable, democratic mood, is measured. All other variables are lagged one year. Gini data are from the Standardized World Income Inequality Dataset (SWIID), version 6.2 (Solt 2016). Education is measured using data on average years of education from the United Nations Development Programme (UNDP). The sources and measurement of all other variables are described in the main text.

Table S2. Tests of Stationarity

	Im-Pesharan-Shin Test		Levin-Lin-Chu Test	
Variable	Statistic	p-value	Statistic	p-value
Mood (69 countries, 20 years)	−4.960	<.001	−4.151	<.001
Mood (121 countries, 10 years)	−40.716	<.001	−109.138	<.001
Δ Mood (65 countries, 20 years)	−6.777	<.001	−5.147	<.001
Δ Mood (120 countries, 10 years)	−33.452	<.001	−78.482	<.001
Liberal democracy	−10.932	<.001	−16.234	<.001
Change in democracy	−37.535	<.001	−32.031	<.001
log GDP per capita	−16.709	<.001	−18.443	<.001
Minoritarian democracy	−15.383	<.001	−29.577	<.001
Electoral democracy	−13.373	<.001	−16.534	<.001
Corruption (99 countries, 15 years)	−12.777	<.001	−7.535	<.001

Because these tests require square panels, the mood panels were trimmed as described. All other panels comprise 134 countries and 30 years. With the alternative hypothesis for both tests being stationarity, the null can be rejected in all cases.

Table S3. Including Education

	General Error-Correction Models		First-Difference Models	
Intercept	.019	.022	−.009*	−.008*
	(.029)	(.028)	(.004)	(.004)
Democratic mood$_{t-1}$	.432*	.431*		
	(.028)	(.028)		
Democratic mood$_{t-2}$	−.450*	−.449*		
	(.028)	(.027)		
Δ liberal democracy	−.060		−.073*	
	(.032)		(.035)	
Liberal democracy$_{t-1}$	.002			
	(.004)			
Δ minoritarian democracy		−.067*		−.089*
		(.030)		(.030)
Minoritarian democracy$_{t-1}$		−.004		
		(.006)		
Δ electoral democracy		.038		.032
		(.041)		(.041)
Electoral democracy$_{t-1}$		.005		
		(.006)		
Δ log GDP per capita	.044	.040	.094	.086
	(.046)	(.047)	(.052)	(.052)
log GDP per capita$_{t-1}$	−.002	−.003		
	(.004)	(.003)		
Δ corruption	−.007	−.006	−.022	−.021
	(.016)	(.016)	(.017)	(.017)
Corruption$_{t-1}$	−.012*	−.013*		
	(.004)	(.004)		
Δ education	−.018	−.019	−.020	−.021
	(.012)	(.012)	(.014)	(.014)
Education$_{t-1}$	−.000	−.000		
	(.001)	(.001)		
N	1920	1920	2008	2008

* $p < .05$. Arellano-White standard errors in parentheses. Education is measured using data on average years of education from the UNDP.

Table S4. Including Income Inequality

	General Error-Correction Models		First-Difference Models	
Intercept	.041	.042	−.010*	−.010*
	(.046)	(.045)	(.004)	(.004)
Democratic mood$_{t-1}$	.431*	.431*		
	(.032)	(.032)		
Democratic mood$_{t-2}$	−.451*	−.451*		
	(.031)	(.031)		
Δ Liberal democracy	−.079*		−.102*	
	(.037)		(.040)	
Liberal democracy$_{t-1}$	.003			
	(.005)			
Δ Minoritarian democracy		−.062		−.089*
		(.033)		(.032)
Minoritarian democracy$_{t-1}$		.000		
		(.009)		
Δ Electoral democracy		−.001		−.012
		(.037)		(.038)
Electoral democracy$_{t-1}$		.003		
		(.008)		
Δ log GDP per capita	.027	.026	.078	.080
	(.054)	(.054)	(.066)	(.066)
log GDP per capita$_{t-1}$	−.003	−.004		
	(.004)	(.004)		
Δ Corruption	−.011	−.011	−.019	−.018
	(.017)	(.017)	(.018)	(.017)
Corruption$_{t-1}$	−.012*	−.012*		
	(.005)	(.004)		
Δ Gini	−.024*	−.024*	−.036*	−.037*
	(.010)	(.010)	(.012)	(.012)
Gini$_{t-1}$	−.000	−.000		
	(.000)	(.000)		
N	1644	1644	1723	1723

*$p < .05$. Arellano-White standard errors in parentheses. Gini index data are from SWIID (version 6.2).

Table S5. One-Way (Country) Fixed Effects models

Liberal democracy$_{t-1}$	−.307*	−.246*		
	(.090)	(.104)		
Electoral democracy$_{t-1}$			.019	.036
			(.090)	(.088)
Minoritarian democracy$_{t-1}$			−.309*	−.262*
			(.092)	(.120)
log GDP per capita$_{t-1}$	−.054	−.042	−.051	−.035
	(.062)	(.075)	(.061)	(.075)
Corruption$_{t-1}$		.004		.002
		(.063)		(.063)
N	2570	2175	2570	2175

*$p < .05$. Arellano-White standard errors in parentheses.

Table S6. Two-Way (Country and Year) Fixed Effects models

Liberal democracy$_{t-1}$	−.293*	−.240*		
	(.093)	(.108)		
Electoral democracy$_{t-1}$			.012	.036
			(.090)	(.090)
Minoritarian democracy$_{t-1}$			−.290*	−.259*
			(.094)	(.122)
log GDP per capita$_{t-1}$	−.038	−.018	−.030	−.007
	(.112)	(.117)	(.109)	(.117)
Corruption$_{t-1}$		.012		.012
		(.059)		(.060)
N	2570	2175	2570	2175

*$p < .05$. Arellano-White standard errors in parentheses.

Table S7. Error-Correction Models with One Lag of Mood

Intercept	−.015	.039	−.025	.049
	(.035)	(.044)	(.033)	(.042)
Democratic mood$_{t-1}$	−.012*	−.013*	−.011*	−.014*
	(.004)	(.005)	(.004)	(.004)
Δ Liberal democracy	−.070*	−.078*		
	(.028)	(.034)		
Liberal democracy$_{t-1}$	.008	−.001		
	(.005)	(.007)		
Δ Electoral democracy			.014	.028
			(.034)	(.041)
Electoral democracy$_{t-1}$			.007	.013
			(.010)	(.011)
Δ Minoritarian democracy			−.073*	−.083*
			(.026)	(.032)
Minoritarian democracy$_{t-1}$			−.000	−.015
			(.010)	(.010)
Δ log GDP per capita	.105*	.090	.100	.080
	(.052)	(.051)	(.053)	(.052)
log GDP per capita$_{t-1}$	.000	−.005	.001	−.006
	(.004)	(.005)	(.004)	(.004)
Δ Corruption		−.031		−.032
		(.017)		(.017)
Corruption$_{t-1}$		−.015*		−.017*
		(.007)		(.006)
N	2435	2026	2435	2026

* $p < .05$. Arellano-White standard errors in parentheses.

Table S8. Within-Between Multilevel Models

Intercept	.528	1.620*	−.001	1.533*
	(.553)	(.617)	(.538)	(.628)
Liberal democracy (within)	−.326*	−.221*		
	(.026)	(.031)		
Liberal democracy (between)	.567*	.333*		
	(.080)	(.100)		
Minoritarian democracy (within)			−.312*	−.236*
			(.037)	(.042)
Minoritarian democracy (between)			.265	−.032
			(.170)	(.174)
Electoral democracy (within)			.002	.039
			(.040)	(.045)
Electoral democracy (between)			.269	.304
			(.170)	(.160)
log GDP per capita (within)	−.045*	−.024	−.044*	−.017
	(.018)	(.023)	(.018)	(.023)
log GDP per capita (between)	−.062	−.181*	−.003	−.172*
	(.063)	(.069)	(.061)	(.070)
Corruption (within)		.025		.024
		(.029)		(.029)
Corruption (between)		−.404*		−.474*
		(.109)		(.108)
N	2570	2239	2570	2239
Variance of country intercepts	.433	.399	.455	.404
Residual variance	.082	.072	.081	.072

*$p < .05$.

References

Alexander, Amy C., and Christian Welzel. 2011. "Measuring Effective Democracy: The Human Empowerment Approach." *Comparative Politics* 43(3): 271–289.

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.

Bratton, Michael, and Richard Houessou. 2014. "Demand for Democracy Is Rising in Africa, But Most Political Leaders Fail to Deliver." Afrobarometer Policy Paper No. 11. http://afrobarometer.org/sites/default/files/publications/Policy%20paper/ab_r5_policypaperno11.pdf.

Carpenter, Bob, Andrew Gelman, Matthew Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software* 76(1): 1–32.

Caughey, Devin, and Christopher Warshaw. 2015. "Dynamic Estimation of Latent Opinion Using a Hierarchical Group-Level IRT Model." *Political Analysis* 23(2): 197–211.

Claassen, Christopher. 2019. "Estimating Smooth Country-Year Panels of Public Opinion." *Political Analysis* 27(1): 1–20.

Dahl, Robert A. 1989. *Democracy and Its Critics*. New Haven, CT: Yale University Press.

Elkins, Zachary. 2000. "Gradations of Democracy? Empirical Tests of Alternative Conceptualizations." *American Journal of Political Science* 44(2): 293–300.

Foa, Roberto Stefan, and Yascha Mounk. 2016. "The Danger of Deconsolidation: The Democratic Disconnect." *The Journal of Democracy* 27(3): 5–17.

Foa, Roberto Stefan, and Yascha Mounk. 2017. "The Signs of Deconsolidation." *Journal of Democracy* 28(1): 5–16.

Fuchs-Schündeln, Nicola, and Matthias Schündeln. 2015. "On the Endogeneity of Political Preferences: Evidence from Individual Experience with Democracy." *Science* 347(6226): 1145–1148.

Inglehart, Ronald. 2003. "How Solid is Mass Support for Democracy – and How Can We Measure It?" *PS: Political Science & Politics* 36(1): 51–57.

Klingemann, Hans-Dieter. 1999. "Mapping Political Support in the 1990s: A Global Analysis." In *Critical Citizens: Global Support for Democratic Governance*, ed. Pippa Norris. Oxford, UK: Oxford University Press.

Kurzman, Charles. 2014. "World Values Lost in Translation." *Washington Post* September 2, https://www.washingtonpost.com/news/monkey-cage/wp/2014/09/02/world-values-lost-in-translation/.

Lindberg, Staffan I., Michael Coppedge, John Gerring, Jan Teorell, Daniel Pemstein, Eitan Tzelgov, Yi ting Wang, Adam Glynn, David Altman, Michael Bernhard, Steven Fish, Alan Hicken, Matthew Kroenig, Kelly McMann, Pamela Paxton, Megan Reif, Svend-Erik Skaaning, and Jeffrey Staton. 2014. "V-Dem: A New Way to Measure Democracy." *Journal of Democracy* 25(3): 159–169.

Mattes, Robert, and Michael Bratton. 2007. "Learning about Democracy in Africa: Awareness, Performance, and Experience." *American Journal of Political Science* 51(1): 192–217.

Munck, Gerardo L., and Jay Verkuilen. 2002. "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices." *Comparative Political Studies* 35(1): 5–34.

Norris, Pippa. 2011. *Democratic Deficit: Critical Citizens Revisited*. Cambridge: Cambridge University Press.

Pemstein, Daniel, Stephen Meserve, and James Melton. 2010. "Democratic Compromise: A Latent Variable Analysis of Ten Measures of Regime Type." *Political Analysis* 18(4): 426–49.

Rose, Richard, William Mishler, and Christian Haerpfer. 1998. *Democracy and Its Alternatives: Understanding Post-Communist Societies*. Baltimore: Johns Hopkins University Press.

Solt, Frederick. 2016. "The Standardized World Income Inequality Database." *Social Science Quarterly* 97(5): 1267–1281.

Stan Development Team. 2017. *Stan Modeling Language: User's Guide and Reference Manual: Stan 2.15.0*. Stan Development Team.

Teorell, Jan, Michael Coppedge, Staffan Lindberg, and Svend-Erik Skaaning. 2018. "Measuring Polyarchy Across the Globe, 1900–2017." *Studies in Comparative International Development* forthcoming, https://doi.org/10.1007/s12116-018-9268-z.

V-Dem Institute. 2018*a*. "Codebook: v8 – April 2018." V-Dem Institute, University of Gothenburg, https://www.v-dem.net/media/filer_public/64/ad/64ad9308-45fa-473e-8e2b-e1c0c4e421e6/v-dem_codebook_v8.pdf.

V-Dem Institute. 2018*b*. "Democracy for All? V-Dem Annual Democracy Report 2018." V-Dem Institute, University of Gothenburg, https://www.v-dem.net/media/filer_public/3f/19/3f19efc9-e25f-4356-b159-b5c0ec894115/v-dem_democracy_report_2018.pdf.

V-Dem Institute. 2018*c*. "Methodology, v8." V-Dem Institute, University of Gothenburg, https://www.v-dem.net/media/filer_public/5a/f1/5af198e9-f3e8-4619-b9fd-a8387fdc22a5/v-dem_methodology_v8.pdf.

Williams, Laron K., and Guy D. Whitten. 2012. "But Wait, There's More! Maximizing Substantive Inferences from TSCS Models." *Journal of Politics* 74(3): 685–693.