OXFORD

National Perspectives and Quantitative Datasets: A Silver Lining?

Sarah Sunn Bush

Department of Political Science, Yale University

Abstract

As noted by other contributions to this special issue, an American perspective shapes many leading quantitative datasets used by international relations scholars. This tendency can lead to biased inferences, but it can also enhance scholarly accuracy under certain conditions. Precisely because some datasets reflect national perspectives, they are appropriate to use when seeking to test theories in which the actors of interest subscribe to the same national perspective. This argument is illustrated with the case of US democracy assistance. Using an appropriate measure of democracy reveals that—contrary to some claims in the literature—US policy-makers allocate democracy assistance in ways that reflect their perceptions of countries' regime types, giving less democracy assistance to countries that they perceive as more democratic.

Keywords: democracy, measurement, democracy assistance, and U.S. foreign policy

Elsewhere in this special issue, Colgan (2019b) shows that several leading datasets produced and used by international relations (IR) scholars reflect an American perspective.[1] As he shows, an American perspective can lead to faulty descriptive and even causal inferences about topics such as the resource curse, war-fighting effectiveness, and nuclear proliferation. Researchers should be aware of the potential for such biases and consider strategies for minimizing them.

Yet, some biases may remain. And, in fact, the American perspective reflected in some IR datasets makes them appropriate to use *when scholars seek to test theories in which the actors of interest subscribe to the same American perspective*. That is the case for many IR research questions, especially within the subfield of security studies (Hendrix and Vreede 2019). As such, this article argues that using datasets with an American perspective helps align certain IR theories and hypotheses with their

---

[1] For a discussion of what an "American perspective" entails, see the essays by Colgan (2019a) and Cheng and Brettle (2019).

empirical tests. More broadly, it calls for greater attention among empirical IR researchers to how datasets reflect conceptual decisions that make them more or less useful for testing theories.

I develop this argument in four steps. First, I argue that an American perspective can be reflected in datasets via two decisions: how to define concepts and how to code countries or other relevant units. Second, I propose that, precisely because the resulting datasets may capture American beliefs about important concepts or the state of the world, they are appropriate to use when answering research questions in which the actors of interest share those beliefs. Third, I show that using datasets that reflect an American perspective can lead to substantially different, and potentially more accurate, findings with an illustration from the case of US democracy assistance. Using a theoretically relevant measure of democracy reveals that—contrary to some claims in the literature—US policy-makers allocate democracy assistance in ways that reflect their perceptions of countries' regime types, giving less democracy assistance to countries that they perceive as more democratic. This

finding is significant both for our understanding of US foreign policy-making and also for researchers who seek to model the effectiveness of US aid. Finally, I provide examples from three other IR literatures—on foreign aid–targeting, naming and shaming by nongovernmental organizations (NGOs), and private investors' evaluations of countries' creditworthiness—where answering core research questions may benefit from using datasets that reflect an American (or other national) perspective.

Ultimately, IR scholars—especially as they embrace a "behavioral revolution" that emphasizes the importance of individual decision-making (Hafner-Burton et al. 2017)—should choose empirical measures that reflect the beliefs of the actors about which they theorize. In other words, sometimes researchers should prefer to use the actual data that decision-makers employ, even when those data reflect certain biases and subjectivities. This advice is particularly relevant for researchers who seek to measure concepts that are deeply contested and difficult to operationalize, such as democracy or creditworthiness.[2] When researchers are trying to learn about American decision-makers' strategies and behaviors in response to subjective conditions and events, they should use data reflecting an American perspective. More generally, using cross-national datasets in a self-conscious way allows researchers to uncover microfoundations that are useful for precisely and accurately testing hypotheses.

Ideology and Quantitative Datasets

With the IR field's growing interest in ratings, rankings, and other types of benchmarks (e.g., Cooley 2015; Kelley and Simmons 2015; Dolan 2017; Kelley 2017), more and more scholars have noted that quantitative datasets reflect their creators' ideological and normative commitments. Although such datasets often appear to be apolitical, they typically reflect their creators' subjective value judgements (Davis, Kingsbury, and Merry 2012, 9; Broome and Quirk 2015). Given the focus of the special issue, I concentrate on how datasets might reflect an American perspective and thus be useful for testing theories about American decision-makers, such as US government officials in the case of US democracy aid.

However, the argument developed here should extend to other types of national and ideological perspectives.

Dataset creators' ideological commitments—and thus, potentially, their national perspectives—can shape the datasets that they produce in at least two ways (Bush 2017, 715). First, raters' ideological commitments influence *concept definition*. Many IR datasets seek to measure states (or other units) in terms of contested concepts that can be and are defined in multiple ways. For example, Bhuta (2012, 134) finds that there has been a "proliferation of competing definitions and typologies" of "state fragility" in recent years, with indices relying variously on concepts such as state effectiveness, capability, legitimacy, and stability, which are themselves defined in diverse ways. There has also been a debate about whether human rights indicators emphasize civil and political rights too much in their definitions of "human rights," to the detriment of economic, cultural, and social rights (Rosga and Satterthwaite 2012, 299).

To understand the conceptual judgments inherent in producing datasets, consider democracy, a concept that has been defined in different ways by the datasets most commonly used by IR scholars (Gunitsky 2015). Democracy can be defined in minimalist ways (which typically emphasize competitive elections) and more maximalist ways (which may incorporate any number of additional political and even social traits). The Varieties of Democracy (V-Dem) Project argues that there are several distinct conceptual models of democracy: electoral, liberal, majoritarian, participatory, deliberative, and egalitarian (Coppedge et al. 2011, 253). Whereas, for example, the key question for liberal democracy is whether "political power [is] decentralized and constrained," the key question for majoritarian democracy is whether the majority rules (Coppedge et al. 2011, 254). Although a country may score well on both dimensions in practice, there is a classic tension between these conceptions of democracy since fostering individual rights often requires institutions that prevent majority tyranny. Given such conflicts, reasonable people can and do disagree about how to define democracy. Datasets' creators thus must make normative judgments about the meaning of democracy. Since meanings of democracy vary across cultures, including within advanced democracies (Diamond and Plattner 2008), it is plausible that raters' national perspectives will shape their normative judgments.

Dataset creators' ideological commitments also shape their *coding decisions*. Even the most detailed coding guidelines do not eliminate the need for human raters to make some subjective judgments about how to code specific countries, events, leaders, and so on. One challenge is that different information sources (e.g., newspapers or

---

2   Measuring concepts such as inflation or economic growth, which have less contested meanings and more objective indicators, involves less of a challenge. However, even ostensibly "objective" cross-national economic indicators have substantial measurement error. See Kerner, Jerven, and Beatty (2017), Dolan (2017), and Kerner and Crabtree (2018).

encyclopedias) may suggest different facts. Another challenge is that different coders may judge the same facts differently. In either case, raters' national perspectives may affect their coding decisions, whether by shaping the sources that they use for basic information about the world or by shaping how they interpret that information once the relevant sources have been gathered (Gunitsky and Tsygankov 2018). Although academic datasets are not immune from these problems, as Colgan (2019b) suggests, problems may be even more likely in the datasets that are widely used by IR researchers but produced by nonacademic, mission-driven organizations (e.g., Freedom House, Transparency International).[3]

National Perspectives in Quantitative Datasets: A Feature *and* a Bug?

The ideological commitments contained and revealed in datasets influence who uses them. Since datasets reflect subjective judgments, actors in world politics are more likely to adopt the datasets that reflect their shared values. For example, Bush (2017) finds that the *Freedom in the World* (FITW) dataset on countries' levels of democracy, which is produced by Freedom House, is used more widely by certain audiences—including American policymakers and democracy-promotion practitioners—than other datasets on democracy. As elaborated below, Bush argues that this pattern obtains because the FITW dataset is more consistent with elite American audiences' ideas about what democracy is and which countries are democratic than other datasets. In other words, that the FITW ratings capture certain beliefs about democracy encourages some audiences to use them and discourages others from doing so.[4]

Such dynamics are relevant as IR scholars increasingly seek to develop and test "middle-range theories" that use microfoundations.[5] In many cases, such theories advance

claims about how individuals—whether citizens, political leaders, investors, human rights activists, or some other group of people—respond to political conditions and events. Of course, these individuals hold many national perspectives. However, perhaps because of the way that the American national perspective affects which research questions IR scholars tend to ask (Hendrix and Vreede 2019), many of the individuals under study in contemporary IR have an American perspective. If we seek to understand how individuals respond to political conditions and events, then it is important to measure those political conditions and events in the way that the individuals in our studies do so. In other words, what testing certain theories requires is not simply an "objective" dataset that is without any type of bias but rather data that reflect how the individuals in question conceptualize the relevant topic and how *they* are likely to evaluate countries or other units.

I illustrate this argument with an example from the literature on democracy assistance. In recent years, many scholars working at the intersection of international and comparative politics have studied the effectiveness of democracy assistance using large-N, quantitative methods (e.g., Finkel, Pérez-Liñán, and Seligson 2007; Scott and Steele 2011; Cornell 2013; Dietrich and Wright 2015; Gibson, Hoffman, and Jablonski 2015; Savage 2017; Heinrich and Loftis 2019). Numerous studies have explicitly focused on aid from the United States, which is the largest democracy assistance donor. These studies generally suggest that democracy aid has a positive average effect on countries' levels of democracy and differ mainly in their conclusions about the conditions under which democracy aid is most likely to have a positive effect. These studies' relatively optimistic conclusions about democracy aid contrast with the more pessimistic tenor of qualitative studies on the same topic (e.g., Carothers 1999; Zeeuw 2005; Brown 2006; Burnell 2011; Carapico 2013).

A challenge for any study of the effectiveness of democracy assistance is that the researcher must control for the factors that make some countries more likely to receive aid in the first place. As such, many studies of the effects of democracy assistance on democracy have sought to empirically model the process by which countries get selected to receive aid (e.g., Finkel et al. 2007; Scott and Steele 2011; Dietrich and Wright 2015). Indeed, some researchers have turned their attention more fully to questions such as why countries are targeted with certain types or amounts of US democracy assistance (e.g., Azpuru et al. 2008; Bush 2015, 2016; Peterson and Scott 2018), often focusing on the effect of countries' pre-existing levels of democracy. Within this literature, there is a debate about whether and how donors tailor aid to

---

3   There are several possible explanations. First, datasets produced by academics may have more explicit coding guidelines and thus permit fewer subjective coding decisions. Second, datasets produced by mission-driven organizations may be more deliberately ideological in orientation.

4   Audiences may consciously or subconsciously select indicators that align with their national perspectives or ideological commitments. My argument applies in either case.

5   Kertzer (2017, 83) defines the use of microfoundations as "an analytic strategy where one explains outcomes at the aggregate level via dynamics at a lower level." Developing a microlevel theory is not synonymous with focusing on individuals, though the two often go together.

countries' political systems. Some scholars suggest that democracy assistance donors tend to operate on "strategic autopilot, carrying out many types of programs in any one setting with little careful thought about which among them offer the most fruitful avenues for change" (Carothers 2015, 64). Others have posited that the people making decisions about where to fund democracy assistance do so in more careful a way that reflects countries' domestic conditions (Scott and Steele 2011, 53–54).

To assess whether a country's preexisting democratic trajectory influences the allocation of democracy assistance, the researcher must choose an indicator of how democratic countries are. As noted above, that choice is complicated: democracy can be and is defined in myriad ways. Researchers can use a dichotomous measure (e.g., Przeworski et al. 2000) or a continuous one. Within the category of continuous measures, there are also many options, including Freedom House, V-Dem (which itself contains multiple measurement options), Polity IV (Marshall, Gurr, and Jaggers 2010), the Universal Democracy Score (Pemstein, Meserve, and Melton 2010), and the Economist Democracy Index (Kekic 2007). These indicators define democracy differently (Gunitsky 2015).[6] In fact, their substantial differences in method have lead researchers to different conclusions about the effects of democracy (e.g., Elkins 2000; Munck and Verkuilen 2002). To test hypotheses about whether donors' strategies of democracy promotion depend on target countries' levels of democracy, *it is appropriate for researchers to choose a democracy indicator that aligns with how donors think about democracy.*

Within the literature on democracy assistance, several studies posit theories that explicitly state that donors care how democratic target countries are, although only some of these studies make reference to how democracy is understood by the relevant set of actors.[7] In a study of US Agency for International Development (USAID) democracy assistance, for example, Peterson and Scott (2018,

276) write, "an important cue for aid allocators comes from the regime conditions most relevant to the prospects for democratic progress and consolidation: the regime's democratic condition as reflected by its level of and trend toward democracy." They use the Polity measure to assess a country's "democratic condition" but do not explain why Polity best reflects how aid allocators think about democracy. By way of comparison, in a study of US assistance to civil society groups, Bush (2016, 375) posits that donor officials will give less aid to more democratic countries and explicitly justifies using the Freedom House measure to test that hypothesis with reference to how democracy is understood by aid allocators. She writes, "I use these scores because donor officials use them" (Bush 2016, 375).

Indeed, in the United States, the decision-makers involved in democracy assistance use the FITW ratings as their primary indicator of democracy. Since the US State Department began assessing countries' human rights records in the 1970s, the FITW reports have been an important source of information for the American government (Bush 2017, 718). When USAID began devoting more effort to evaluating its democracy and governance programming in the 1990s, FITW was adopted as the main indicator (McMahon 2001, 463). Since then, USAID and more generally the US government have continued to use the FITW ratings to evaluate the effectiveness of democracy assistance (Bush 2015, 59). For example, the Millennium Challenge Corporation, which is a US initiative that seeks to give economic aid to countries that are governed justly and democratically, chose the FITW ratings in 2004 as one of its primary indicators (Girod, Krasner, and Stoner-Weiss 2009, 71).

There are likely several reasons why US policy-makers gravitate toward the FITW ratings over other potential democracy indicators. One reason is that the FITW coding criteria reflect a broadly liberal conception of democracy (Giannone 2010; Bush 2017, 720–21). Interviews, archival materials, practitioner writings, and other forms of qualitative evidence indicate that American elites engaged in democracy promotion likewise tend to think about democracy in liberal terms (e.g., Monten 2005; Kurki, 2010, 2013; Bunce and Wolchik 2012; Green 2012; Mitchell 2016, 111–12). Another reason is that the FITW ratings tend to code countries that are aligned with the US government more favorably than other countries (Steiner 2016). Examples of this trend that have been cited by previous scholars include relatively negative scores for Russia during the 1990s and 2000s and, during the Cold War, the scores for El Salvador and Nicaragua (Goldstein 1986, 48; Mainwaring, Brinks, and Pérez-Liñán 2001, 54; Gunitsky 2015).

---

6   For further discussion of measurement bias in democracy and governance measures, see Bollen and Paxton (2000); McHenry (2000); Bowman, Lehoucq, and Mahoney (2005); Munck (2009); and Thomas (2010).

7   See also Scott and Steele (2011) and Carothers (2015). In contrast, other studies in this literature do not make their microfoundations explicit by identifying the relevant actors (e.g., aid allocators), let alone those actors' preferences. For example, Finkel et al. (2007) seek to empirically model the selection process without fully theorizing how allocation occurs. As I discuss below, making these steps in a theory clear can open new avenues for testing causal mechanisms.

In these ways, FITW reflects an American perspective on democracy that is dominant within the US government, though of course it is not the *only* American perspective on democracy.[8] One explanation for why the FITW ratings conceptualize democracy and code countries in a way that is consistent with the view of the US government is that they are produced by an American NGO, Freedom House. That Americans are involved with the creation of a dataset does not necessarily mean that the dataset will result in an "American perspective" (Cheng and Brettle 2019). However, researchers have argued that there is substantial overlap between the types of people working for Freedom House and people involved in US democracy assistance (Giannone 2010; Steiner 2016; Bush 2017). Since the people who work in US democracy assistance rely heavily on the FITW ratings, researchers should use the FITW ratings when evaluating whether the US government allocates democracy assistance according to target countries' levels of democracy.

Below, I examine the relationship between democracy assistance and prior democracy using FITW, a theoretically relevant indicator of democracy. Many scholars working interested in understanding the drivers of US democracy assistance policy have explored the *linear* relationship between democracy and democracy assistance, and I follow suit.[9] The implicit or explicit

theoretical logic being tested in such studies is that, as countries become more democratic, they are less likely to receive democracy assistance. When a country successfully transitions to democracy or becomes a more consolidated democracy, it arguably has less need for democracy assistance. If donors are making strategic decisions about how to allocate democracy assistance, one would expect donors to respond to such developments. In fact, aid donors—and USAID specifically—have been described as "obsessed with the imperative of 'graduation,'" or the idea that some countries may no longer need democracy support (Diamond 2009, 38). Some countries, such as Poland, have not only ceased to receive democracy assistance because of their successful democratic transitions but have also become democracy assistance donors themselves (Petrova 2014).

My analysis shows that using FITW instead of alternative indicators can result in substantially different conclusions about the nature of the relationship between democracy and democracy assistance. To support this claim, I compare FITW to its most commonly used academic alternative, Polity IV.[10] Polity IV is generally thought to have a more maximalist definition of democracy than FITW and sometimes codes countries differently (Steiner 2016). FITW measures countries' political rights and civil liberties on seven-point scales, which I invert (so that 1 represents the least free countries and 7 represents the freest countries) and then add together. Polity IV measures countries' levels of democracy on a scale that ranges from −10 (most autocratic) to 10 (most democratic). To aid comparison, I standardize both variables so that they have means of zero and standard deviations of one. A graph with the nonstandardized variables can be found in the supplementary appendix.

Figure 1 compares the relationship between democracy in the previous year and the amount of US democracy aid that a country received for all potential recipient countries between 2002 and 2013.[11] I log the democracy assistance variable since it is highly skewed.

---

8   Of course, Americans are also involved in the production of other democracy indicators (e.g., Polity). However, the affinity between how FITW approaches democracy and how the US government approaches democracy distinguishes the Freedom House ratings from other prominent democracy indicators. See Bush (2017).

9   E.g., see Finkel, Pérez-Liñán, and Seligson (2007, 422); Scott and Steele (2011, 62); Dietrich and Wright (2015, 223); and Bush (2016, 377) for examples of these linear tests. Alternatively, one might hypothesize that there is a curvilinear relationship between democracy and democracy assistance, with donors expending more effort in partial democracies than in full autocracies or democracies. The supplementary appendix contains locally weighted smoothing scatterplots (LOWESS plots), which involve locally weighted regressions of democracy assistance on democracy. As expected, the LOWESS plots suggest somewhat different functional forms in terms of the relationship between democracy and democracy assistance depending on which democracy measure is used. When FITW is used, the relationship between democracy and democracy assistance is monotonically decreasing (as suggested by the linear

regression models), whereas the relationship is non-monotonic when Policy is used. As such, this further analysis supports my general point that it matters which indicator of democracy is used.

10 Figures in this article were made using Bischof (2017).

11 Democracy aid is operationalized using data on US aid commitments in constant 2011 US dollars from the Organization for Economic Co-operation and Development (OECD). I focus on aid that falls into OECD purpose code 15150 ("democratic participation and civil society"), which encompasses aid for elections, legislatures and political parties, and free media, among other things. This category excludes aid that is
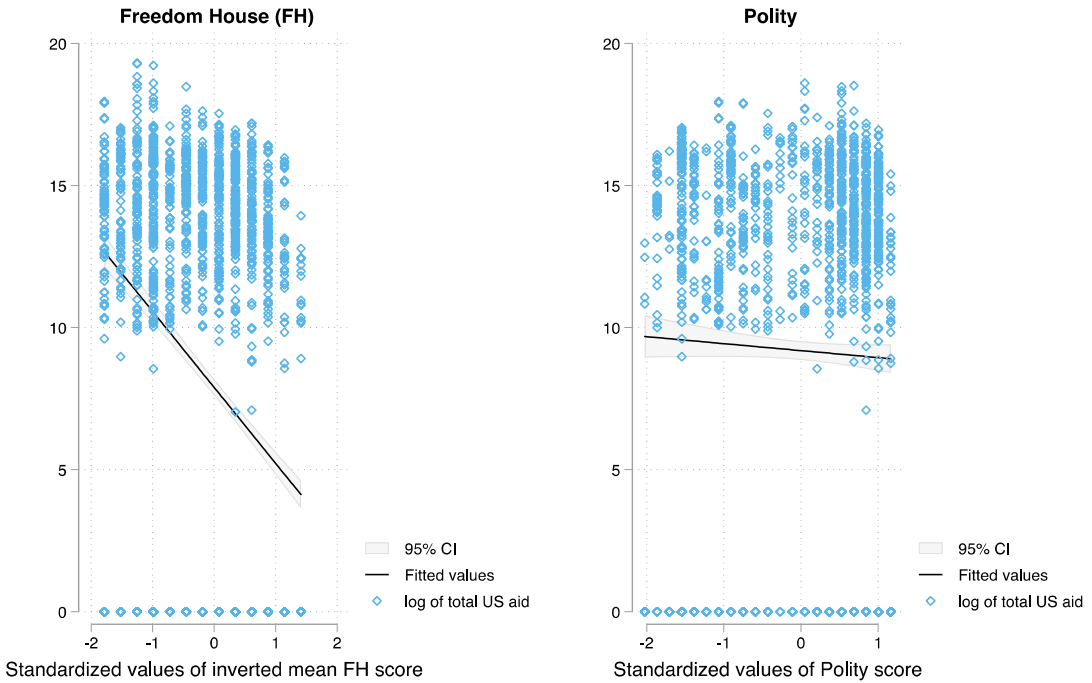
Figure 1. Freedom House vs. Polity comparison

As the figure suggests, the observed relationship between democracy and democracy aid varies significantly depending on which indicator of democracy is used. Prior democracy is a significant predictor of democracy aid when the FITW indicator is used; an improvement in the FITW score of one standard deviation is associated with around $2.3 million less in democracy aid on average, and this difference is highly statistically significant ($p < 0.001$). In contrast, prior democracy is not a clear predictor of democracy aid when the Polity IV indicator is used. Indeed, we cannot confidently reject the null hypothesis that there is no relationship between democracy and democracy aid at conventional levels of statistical significance ($p < 0.10$). Moreover, and although the relationship trends negative, the substantive effect is much smaller, with an improvement in the

primarily targeted at governance or the rule of law and therefore less squarely focused on promoting democracy. OECD reporting requirements have changed over time, and so data are available only between 2002 and 2013. OECD donor countries were excluded from the analysis of potential recipient countries since they do not receive democracy aid. This approach to operationalizing democracy aid follows Lührmann, McMann, and van Ham (2016).

Freedom House score of one standard deviation being associated with around $140,000 less in democracy aid on average. These general patterns hold when I construct a more fully specified model that controls for other factors that are correlated with democracy, including the recipient country's region, experience with past civil conflict, and similarity in United Nations voting with the United States.

One possibility is that the factors that determine whether countries receive any democracy assistance may differ from the factors that determine how much democracy assistance countries receive. To explore this possibility, I replicate the analysis on the sample of those countries that received at least some of democracy assistance.[12] As Figure 2 shows, I again find a significant negative relationship between democracy and democracy assistance when I use the Freedom House indicator.

12 Separately, I estimated the effect of democracy on the probability of receiving any democracy assistance using probit models. Again, the choice of democracy indicator matters. Whereas improvements in the Freedom House score are associated with significantly lower probabilities of receiving aid ($p < 0.001$), the same is not true for improvements in the Polity score. This analysis is contained in the supplementary appendix.
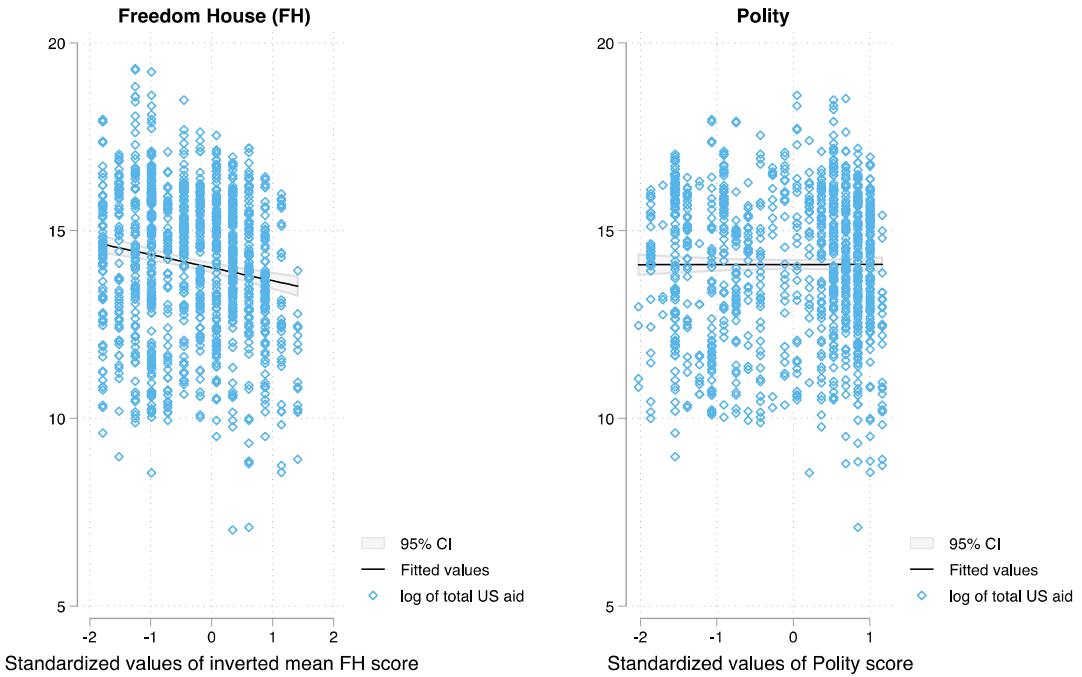
Figure 2. Freedom House vs. Polity comparison, no-aid cases excluded

Specifically, an improvement in the Freedom House score of one standard deviation is associated with around $2.9 million less in democracy aid on average. Again, this difference is highly statistically significant ($p < 0.001$). And also, similar to the earlier analysis, when I use the Polity measure, I do not find that democracy is clearly related to democracy assistance.

For a further comparison, I examined the relationship between democracy assistance and democracy using the indicator of liberal democracy (v2x_liberal) produced by V-Dem (Coppedge et al. 2017; Pemstein et al. 2017). Similar to FITW but unlike Polity, the V-Dem indicator has a statistically significant relationship with democracy ($p < 0.001$). The relevant graph is in the supplementary appendix. Since the V-Dem liberal democracy component is similar to the FITW ratings in terms of how it conceptualizes democracy, this finding suggests that US democracy aid allocations may be as much guided by the *idea* of democracy that US policy-makers have in mind—which is shared by both FITW and the V-Dem liberal democracy component score—as the specific judgements of countries that were presented in the previous year's FITW ratings.

The analysis demonstrates that the choice of democracy indicator matters for how we understand the relationship between democracy and democracy aid. Indeed,

the analysis suggests that the US government, rather than allocating aid on autopilot, *does* attempt to target its approach to countries' levels of democracy, with the United States giving less aid to countries that aid officials perceive as democratizing. This is an important substantive conclusion for scholars who want to understand US foreign policy. It also suggests that researchers who want to identify the effect of democracy aid on democracy must control for the prior level of democracy and, more generally, take selection issues seriously. If countries that receive democracy assistance are less democratic, on average, then we may be underestimating the effect of democracy assistance, which tends to go to the more challenging cases.

Finally, the divergence in the findings between the FITW and Polity indicators suggests that it remains a somewhat open question whether democracy assistance actually *is* tailored to countries' levels of democracy, even though the evidence suggests that the US government is attempting to give more aid to countries that it views as being less democratic. Answering that question depends on how one thinks it is best to define and operationalize democracy. For example, the further analysis using the V-Dem measure supports an interpretation that the US government is taking countries' levels of *liberal*

democracy into account when making decisions about how to promote democracy. However, scholars may disagree about the extent to which that is the appropriate way to conceptualize democracy in the first place.

Given these complexities, the analysis suggests that answering even a seemingly straightforward empirical question such as "is democracy assistance targeted according to countries' regimes types?" can benefit from some attention to who is doing the targeting and how they might think about regime types. A further implication that follows is that scholars must take great care when drawing conclusions about null results, such as that aid allocators are not targeting aid according to how democratic countries are. To draw such a conclusion, researchers must be confident that they are using indicators of democracy that reflect how the relevant actors think about democracy.

Extending the Argument to Other Research Questions

Arguably, democracy is a concept that is unusually difficult to measure. Yet, it is essential to many contemporary research questions in IR and security studies, including those about the democratic peace or war-fighting effectiveness. Moreover, my argument has broader applications since democracy indicators are not the only data for which researchers should prefer to use sources that reflect the perspectives of decision-makers. I briefly illustrate this point using examples of research questions in three other literatures. The examples show that it may be both possible and worthwhile to learn about decision-makers' uses of quantitative data in domains beyond democracy assistance. Such information can help researchers select appropriate indicators for testing midrange IR theories.

First, do countries target foreign economic aid according to countries' domestic conditions, such as levels of corruption, development, or rule of law? Many studies have shown that strategic interests influence development aid allocation, with bilateral and multilateral donors giving aid to reward their allies even when they do not demonstrate the characteristics that are thought to ensure that aid is used effectively (e.g., Alesina and Dollar 2000; Bueno de Mesquita and Smith 2009). Yet, some studies find that donors give aid more selectively; targeting aid (either in its amounts or its types) in a manner that reflects donors' thinking about it is most likely to have a positive impact on development (e.g., Winters 2010; Bermeo 2017).

This debate centers on whether the individuals in charge of aid allocation—such as bureaucrats and policy-

makers—tailor aid according to their understandings of country performance. To answer that question, it is appropriate for scholars to use the datasets that are closest to decision-makers' understandings of country performance, potentially comparing how perceived performance relates to "actual" performance as captured by other indicators. Even if donors are not themselves relying on quantitative IR datasets to select countries for aid,[13] it is desirable for scholars to approximate decision-makers' beliefs about country performance when testing theories about aid targeting. In the case of foreign aid provided by the US government or by international institutions in which the United States plays an important role, it makes sense to use datasets that reflect an American perspective to understand how decision-makers' perceptions influence aid allocation.

Second, do NGOs such as Amnesty International and Human Rights Watch focus on the countries that are the worst human rights abusers when they engage in naming and shaming? Scholars disagree about how effective campaigns seeking to publicize human rights abuses are and under what conditions they are most successful (e.g., Hafner-Burton 2008; Murdie and Davis 2012; Hendrix and Wong 2013). One pattern that this debate has revealed is that human rights NGOs are biased, reporting more on abuses committed in powerful states and in states that are covered more often by the US media (Ron, Ramos, and Rodgers 2005; Hafner-Burton and Ron 2013). When seeking to understand variation in NGOs' decisions about naming and shaming, researchers should use datasets that match how such NGO decision-makers think about human rights (e.g., by emphasizing civil and political rights as opposed to other types of rights; see Rosga and Satterthwaite 2012). Given the importance of national environments for how human rights NGOs do their work (Stroup 2012), it may be sensible for scholars to use datasets influenced by an American perspective when they seek to explain the behaviors of American human rights NGOs, such as Human Rights Watch.

Finally, how do private investors evaluate countries' creditworthiness? Sovereign risk is traditionally thought to be correlated with macroeconomic indicators such as inflation, current and capital account balances, and government consumption. However, investors often rely on other heuristics to infer how risky it is to lend a country money, such as that country's participation in international institutions (Gray 2013; Gray and Hicks 2014) or the performance of its peers

---

13   Though in some cases, donors *are* relying on such datasets. See Girod et al. (2009).

(Brooks, Cunha, and Mosley 2015). Investors' ideas about what makes a country a risky investment may originate in their educational experiences (e.g., Nelson 2014) or through some other experiences that relate to their national perspectives. Regardless of the mechanism, US-based credit-rating agencies tend to give higher ratings to the United States and other democracies than do other credit-rating agencies, suggesting the importance of an American perspective and the possibility of American bias (Fuchs, Gehring, and McDowell 2016; Fuchs and Gehring 2017). Once again, it makes sense for scholars seeking to understand how investors evaluate countries' creditworthiness to rely on the actual credit ratings produced by private agencies that are used by the investors of interest (e.g., Fitch, Moody's, and Standard and Poor's), as well as the more subjective indicators that the investors in question are known to value.

Returning to the issue of democracy indicators, there is a large literature on the idea of a "democratic advantage" in sovereign debt markets that is germane to this discussion (Schultz and Weingast 2003; Saiegh 2005; Archer, Biglaiser, and DeRouen 2007; Beaulieu, Cox, and Saiegh 2012; Biglaiser and Staats 2012). Some of the questions raised earlier about conceptualizing and measuring democracy are also relevant here. For example, the goal in Archer et al. (2007) is to understand if credit-rating agencies are considering regime type, and the key finding (using Polity as the indicator of democracy) is that they do not. Yet, before researchers conclude that credit-rating agencies are not considering regime type, the present study suggests that it would be important for them to confirm that the credit-rating agencies under study do conceptualize and code democracy in a way that is consistent with the Polity approach. In much the same way that we may be drawing inferences about the relationships between democracy and democracy assistance that are contingent on the measures of democracy that are used, some of our conclusions about a "democratic advantage" may also benefit from further attention to measurement of democracy.

Conclusions

There is a silver lining that emerges when we recognize the national perspectives that shape some of the important datasets used in IR: when we seek to test theories about actors with specific national perspectives, the availability of datasets that reflect those national perspectives is helpful for understanding actors' behavior. Of course, such datasets are not always available. Moreover, some concepts can be defined and measured in a straightforward manner, without significant scholarly disagreement.

Yet, any dataset on a complex topic of interest to IR scholars will likely be subject to some debate about concept definition and coding, as ideas such as "democracy," "state fragility," or "prosperity" can and do mean different things to different people. Recognizing that possibility is a way that positivist IR scholars can learn from IR scholars working in a more critical or conceptual research tradition. National perspectives and biases in quantitative data are not always or merely "bugs" that can easily be identified and fixed; instead, given the complexity of the concepts IR scholars seek to measure, such perspectives and biases are to some degree inevitable. As such, a good starting point is for researchers to be aware of them and take them into account when deciding which data to use.

Moreover, choices about which dataset to use should be informed by the theory that researchers seek to test. For example, the analysis presented above suggests that using a dataset that reflects an American perspective on democracy is appropriate and even necessary when seeking to explain US decision-makers' choices about how to give democracy aid. Indeed, and despite prominent criticisms of the Freedom House indices (e.g., Munck and Verkuilen 2002), IR researchers can and should still use the FITW ratings to measure democracy under certain conditions. The keys when doing so are that the researcher should make the potential biases—including but not limited to those related to national perspective—explicit and justify the choice of measurement with respect to the argument and research design.

In some cases, navigating these decisions will require researchers to think more deeply about the actors and concepts about which they are theorizing. As Coppedge et al. (2011, 260) point out, for example, when democracy is used as an explanatory variable (e.g., to explain economic outcomes or conflict outcomes), "we need to know, as specifically as possible, which elements of democracy are related to which results. This is helpful from the policy perspective as well as from the analytic perspective, so that we can gain insight into causal mechanisms." Of course, not every cross-national research study can or should include a deep conceptual treatment of each variable included in its analysis. Still, further attention is often warranted, especially for key explanatory and outcome variables that involve contested concepts.

This type of nuanced approach to selecting quantitative datasets may be increasingly necessary. As IR scholars move toward specifying and testing microfoundational arguments, they are essentially advancing arguments that rest on the decisions of individuals with unique national perspectives. These individuals may include government leaders, aid bureaucrats, investors,

and NGO directors. To the extent possible, researchers should attempt to use indicators that reflect those individuals' unique national and other perspectives. Doing so matters for how we understand the behaviors of the individuals we seek to study.

Acknowledgements

Supplementary Information

Supplementary information is available at the *Journal of Global Security Studies* data archive.

References

Alesina, Alberto, and David Dollar. 2000. "Who Gives Foreign Aid to Whom and Why?" *Journal of Economic Growth* 5 (1): 33–63.

Archer, Candace C., Glen Biglaiser, and Karl DeRouen. 2007. "Sovereign Bonds and the 'Democratic Advantage': Does Regime Type Affect Credit Rating Agency Ratings in the Developing World?" *International Organization* 61 (2): 341–65.

Azpuru, Dinorah, Steven E. Finkel, Aníbal Pérez-Liñán, and Mitchell A. Seligson. 2008. "What Has the United States Been Doing?" *Journal of Democracy* 19 (2): 150–59.

Beaulieu, Emily, Gary W. Cox, and Sebastian Saiegh. 2012. "Sovereign Debt and Regime Type: Reconsidering the Democratic Advantage." *International Organization* 66 (4): 709–38.

Bermeo, Sarah Blodgett. 2017. "Aid Allocation and Targeted Development in an Increasingly Connected World." *International Organization* 71 (4): 735–66.

Bhuta, Nehal. 2012. "State Failure: The US Fund for Peace Failed States Index." In *Governance by Indicators: Global Power through Classification and Rankings*, edited by Kevin E. Davis, Angelina Fisher, Benedict Kingsbury, and Sally Engle Merry, 132–64. Oxford: Oxford University Press.

Biglaiser, Glen, and Joseph L. Staats. 2012. "Finding the Democratic Advantage in Sovereign Bond Ratings: The Importance of Strong Courts, Property Rights Protection, and the Rule of Law." *International Organization* 66 (3): 515–35.

Bischof, Daniel. 2017. "New Graphic Schemes for Stata: Plotplain and Plotting." *Stata Journal* 17 (3): 748–59.

Bollen, Kenneth A., and Pamela Paxton. 2000. "Subjective Measures of Liberal Democracy." *Comparative Political Studies* 33 (1): 58–86.

Bowman, Kirk, Fabrice Lehoucq, and James Mahoney. 2005. "Measuring Political Democracy: Case Expertise, Data Adequacy, and Central America." *Comparative Political Studies* 38 (8): 939–70.

Brooks, Sarah M., Raphael Cunha, and Layna Mosley. 2015. "Categories, Creditworthiness, and Contagion: How Investors' Shortcuts Affect Sovereign Debt Markets." *International Studies Quarterly* 59 (3): 587–601.

Broome, André, and Joel Quirk. 2015. "The Politics of Numbers: The Normative Agendas of Global Benchmarking." *Review of International Studies* 41 (5): 813–18.

Brown, Keith, ed. 2006. *Transacting Transition: The Micro-Politics of Democracy Assistance in the Former Yugoslavia.* Bloomfield, CT: Kumarian Press.

Bueno de Mesquita, Bruce, and Alastair Smith. 2009. "A Political Economy of Aid." *International Organization* 63 (2): 309–40.

Bunce, Valerie J., and Sharon L. Wolchik. 2012. "Concepts of Democracy Among Donors and Recipients of Democracy Promotion: An Empirical Pilot Study." In *The Conceptual Politics of Democracy Promotion*, edited by Christopher Hobson and Milja Kurki, 151–70. London: Routledge.

Burnell, Peter J. 2011. *Promoting Democracy Abroad: Policy and Performance.* New Brunswick, NJ: Transaction Publishers.

Bush, Sarah Sunn. 2015. *The Taming of Democracy Assistance: Why Democracy Promotion Does Not Confront Dictators.* Cambridge: Cambridge University Press.

———. 2016. "When and Why is Civil Society Support 'Made-in-America'? Delegation to Non-State Actors in American Democracy Promotion." *The Review of International Organizations* 11 (3): 361–85.

———. 2017. "The Politics of Rating Freedom: Ideological Affinity, Private Authority, and the Freedom in the World Ratings." *Perspectives on Politics* 15 (3): 711–31.

Carapico, Sheila. 2013. *Political Aid and Arab Activism: Democracy Promotion, Justice, and Representation.* New York: Cambridge University Press.

Carothers, Thomas. 1999. *Aiding Democracy Abroad: The Learning Curve.* Washington, DC: Carnegie Endowment for International Peace.

———. 2015. "Democracy Aid At 25: Time to Choose." *Journal of Democracy* 26 (1): 59–73.

Cheng, Christine, and Alison Brettle. 2019. "How Cognitive Frameworks Shape the American approach to International Relations and Security Studies." *Journal of Global Security Studies* 4 (3): 321–44.

Colgan, Jeff. 2019a. "American Perspectives and Blind Spots on World Politics." *Journal of Global Security Studies* 4 (3): 300–9.

———. 2019b. "American Bias in Global Security Studies Data." *Journal of Global Security Studies* 4 (3): 358–71.

Cooley, Alexander. 2015. "The Emerging Politics of International Rankings and Ratings: A Framework for Analysis." In *Ranking the World: The Politics of International Rankings and Ratings*, edited by Alexander Cooley and Jack Snyder, 1–38. New York: Cambridge University Press.

Coppedge, Michael, John Gerring, David Altman, Michael Bernhard, Steven Fish, Allen Hicken, and Staffan I. Lindberg et al. 2011. "Conceptualizing and Measuring Democracy: A New Approach." *Perspectives on Politics* 9 (2): 247–67.

Coppedge, Michael, John Gerring, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, and Michael

Bernhard et al. 2017. "V-Dem [Country-Year/Country-Date] Dataset v7.1." Varieties of Democracy (V-Dem) Project. Accessed May 24, 2019, https://www.v-dem.net/en/data/data-version-7-1/.

Cornell, Agnes. 2013. "Does Regime Type Matter for the Impact of Democracy Aid on Democracy?" *Democratization* 20 (4): 642–67.

Davis, Kevin E., Benedict Kingsbury, and Sally Engle Merry. 2012. "Introduction: Governance By Indicators." In *Governance by Indicators: Global Power through Classification and Rankings*, edited by Kevin E. Davis, Angelina Fisher, Benedict Kingsbury, and Sally Engle Merry, 3–28. Oxford: Oxford University Press.

Diamond, Larry. 2009. "Supporting Democracy: Refashioning US Global Strategy." In *Democracy in US Security Strategy: From Promotion to Support*, edited by Alexander T. J. Lennon, 29–55. Washington, DC: Center for Strategic and International Studies.

Diamond, Larry and Marc F. Plattner, eds. 2008. *How People View Democracy*. Baltimore: Johns Hopkins University Press and the National Endowment for Democracy.

Dietrich, Simone, and Joseph Wright. 2015. "Foreign Aid Allocation Tactics and Democratic Change in Africa." *Journal of Politics* 77 (1): 216–34.

Dolan, Lindsay R. 2017. "The Politics of Classification in Global Development." PhD dissertation, Columbia University. Accessed May 24, 2019, https://academiccommons.columbia.edu/doi/10.7916/D8NZ9R20.

Elkins, Zachary. 2000. "Gradations of Democracy? Empirical Tests of Alternative Conceptualizations." *American Journal of Political Science* 44 (2): 287–94.

Finkel, Steven E., Aníbal Pérez-Liñán, and Mitchell A. Seligson. 2007. "The Effects of US Foreign Assistance on Democracy Building, 1990–2003." *World Politics* 59 (3): 404–39.

Fuchs, Andreas, and Kai Gehring. 2017. "The Home Bias in Sovereign Ratings." *Journal of the European Economic Association* 15 (6): 1386–423.

Fuchs, Andreas, Kai Gehring, and Daniel McDowell. 2016. "Sovereign Debt, Regime Type, and the Rise of China: Re-Reconsidering the Democratic Advantage." Paper presented at the Annual Meeting of the American Political Science Association, September 1–4, 2016, Philadelphia.

Giannone, Diego. 2010. "Political and Ideological Aspects in the Measurement of Democracy: The Freedom House Case." *Democratization* 17 (1): 68–97.

Gibson, Clark C., Barak D. Hoffman, and Ryan S. Jablonski. 2015. "Did Aid Promote Democracy in Africa? The Role of Technical Assistance in Africa's Transitions." *World Development* 68: 323–35.

Girod, Desha M., Stephen D. Krasner, and Kathryn Stoner-Weiss. 2009. "Governance and Foreign Assistance: The Imperfect Translation of Ideas into Outcomes." In *Promoting Democracy and the Rule of Law: American and European Strategies*, edited by Amichai Magen, Thomas Risse, and Michael A. McFaul, 61–92. London: Palgrave Macmillan.

Goldstein, Robert Justin. 1986. "The Limitations of Using Quantitative Data in Studying Human Rights Abuses." *Human Rights Quarterly* 8 (4): 607–27.

Gray, Julia. 2013. *The Company States Keep: International Economic Organizations and Investor Perceptions*. Cambridge: Cambridge University Press.

Gray, Julia, and Raymond P. Hicks. 2014. "Reputations, Perceptions, and International Economic Agreements." *International Interactions* 40 (3): 325–49.

Green, Brendan Rittenhouse. 2012. "Two Concepts of Liberty: US Cold War Grand Strategies and the Liberal Tradition." *International Security* 37 (2): 9–43.

Gunitsky, Seva. 2015. "Competing Measures of Democracy in the Former Soviet Republics." In *Ranking the World: The Politics of International Rankings and Ratings*, edited by Alexander Cooley and Jack Snyder, 112–50. New York: Cambridge University Press.

Gunitsky, Seva, and Andrei P. Tsygankov. 2018. "The Wilsonian Bias in the Study of Russian Foreign Policy." *Problems of Post-Communism* 65 (6): 385–93.

Hafner-Burton, Emilie M. 2008. "Sticks and Stones: Naming and Shaming the Human Rights Enforcement Problem." *International Organization* 62 (4): 689–716.

Hafner-Burton, Emilie M., and James Ron. 2013. "The Latin Bias: Regions, the Anglo-American Media, and Human Rights." *International Studies Quarterly* 57 (3): 474–91.

Hafner-Burton, Emilie, Stephen Haggard, David Lake, and David A. Victor. 2017. "The Behavioral Revolution and International Relations." *International Organization* 71 (S1): S1–S31.

Heinrich, Tobias, and Matt W. Loftis. 2019. "Democracy Aid and Electoral Accountability." *Journal of Conflict Resolution* 63 (1): 139–66.

Hendrix, Cullen S., and Jon Vreede. 2019. "US Dominance in International Relations and Security Scholarship in Leading Journals." *Journal of Global Security Studies* 4 (3): 310–20.

Hendrix, Cullen S., and Wendy H. Wong. 2013. "When Is the Pen Truly Mighty? Regime Type and the Efficacy of Naming and Shaming in Curbing Human Rights Abuses." *British Journal of Political Science* 43 (3): 651–72.

Kekic, Laza. 2007. "The Economist Intelligence Unit's Index of Democracy." In *The World in 2007*, 1–11. London: Economist Intelligence Unit.

Kelley, Judith G. 2017. *Scorecard Diplomacy: Grading States to Influence Their Reputation and Behavior*. Cambridge: Cambridge University Press.

Kelley, Judith G., and Beth A. Simmons. 2015. "Politics by Number: Indicators As Social Pressure in International Relations." *American Journal of Political Science* 59 (1): 55–70.

Kerner, Andrew, and Charles Crabtree. 2018. "The IMF and the Political Economy of Data Production." Working Paper, Michigan State University and University of Michigan. Accessed May 24, 2019, https://osf.io/preprints/socarxiv/qsxae/.

Kerner, Andrew, Morten Jerven, and Alison Beatty. 2017. "Does It Pay to Be Poor? Testing for Systematically Underreported GNI Estimates." *Review of International Organizations* 12 (1): 1–38.

Kertzer, Joshua D. 2017. "Microfoundations in International Relations." *Conflict Management and Peace Science* 34 (1): 81–97.

Kurki, Milja. 2010. "Democracy and Conceptual Contestability: Reconsidering Conceptions of Democracy in Democracy Promotion." *International Studies Review* 12 (3): 362–86.

———. 2013. *Democratic Futures: Revisioning Democracy Promotion*. Abingdon: Routledge.

Lührmann, Anna, Kelley McMann, and Carolien van Hamm. 2016. "The Effectiveness of Democracy Aid to Different Regime Types and Democracy Sectors." Working Paper, Series 2017:40, The Varieties of Democracy Institute. Accessed May 24, 2019, https://www.v-dem.net/files/25/V-Dem%20Working%20Paper%202017_40.pdf.

Mainwaring, Scott, Daniel Brinks, and Aníbal Pérez-Liñán. 2001. "Classifying Political Regimes in Latin America." *Studies in Comparative International Development* 36 (1): 37–65.

Marshall, Monty G., Ted Robert Gurr, and Keith Jaggers. 2010. "Polity IV Project Dataset Users' Manual." Accessed May 24, 2019, http://www.systemicpeace.org/polity/polity4.htm.

McHenry, Dean E. 2000. "Quantitative Measures of Democracy in Africa: An Assessment." *Democratization* 7 (2): 168–85.

McMahon, Edward R. 2001. "Assessing USAID's Assistance for Democratic Development: Is It Quantity Versus Quality?" *Evaluation* 7 (4): 453–67.

Mitchell, Lincoln. 2016. *The Democracy Promotion Paradox*. Washington, DC: Brookings Institution Press.

Monten, Jonathan. 2005. "The Roots of the Bush Doctrine: Power, Nationalism, and Democracy Promotion in US Strategy." *International Security* 29 (4): 112–56.

Munck, Gerardo L. 2009. *Measuring Democracy: A Bridge between Scholarship and Politics*. Baltimore: John Hopkins University Press.

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

Murdie, Amanda M., and David R. Davis. 2012. "Shaming and Blaming: Using Events Data to Assess the Impact of Human Rights INGO." *International Studies Quarterly* 56 (1): 1–16.

Nelson, Stephen C. 2014. "Playing Favorites: How Shared Beliefs Shape the IMF's Lending Decisions." *International Organization* 68 (2): 297–328.

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

Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi-ting Wang, Joshua Krusell, and Farhad Miri et al. 2017. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data." University of Gothenburg, Varieties of Democracy Institute: Working Paper No. 21, 2nd edition.

Peterson, Timothy M., and James M. Scott. 2018. "The Democracy Aid Calculus: Regimes, Political Opponents, and the Allocation of US Democracy Assistance, 1981–2009." *International Interactions* 44 (2): 268–93.

Petrova, Tsveta. 2014. *From Solidarity to Geopolitics: Support for Democracy Among Postcommunist States*. Cambridge: Cambridge University Press.

Przeworski, Adam, Michael E. Alvarez, Jose Antonio Cheibub, and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Well-Being in the World, 1950–1990*. New York: Cambridge University Press.

Ron, James, Howard Ramos, and Kathleen Rodgers. 2005. "Transnational Information Politics: NGO Human Rights Reporting, 1986–2000." *International Studies Quarterly* 49 (3): 557–88.

Rosga, AnnJanette, and Margaret L. Satterthwaite. 2012. "Measuring Human Rights: UN Indicators in Critical Perspective." In *Governance by Indicators: Global Power through Classification and Rankings*, edited by Kevin E. Davis, Angelina Fisher, Benedict Kingsbury, and Sally Engle Merry, 297–316. Oxford: Oxford University Press.

Saiegh, Sebastian. 2005. "Do Countries Have a 'Democratic Advantage?' Political Institutions, Multilateral Agencies, and Sovereign Borrowing." *Comparative Political Studies* 38 (4): 366–87.

Savage, Jesse Dillon. 2017. "Military Size and the Effectiveness of Democracy Assistance." *Journal of Conflict Resolution* 61 (4): 839–68.

Schultz, Kenneth A., and Barry R. Weingast. 2003. "The Democratic Advantage: Institutional Foundations of Financial Power in International Competition." *International Organization* 57 (1): 3–42.

Scott, James M., and Carie A. Steele. 2011. "Sponsoring Democracy: The United States and Democracy Aid to the Developing World, 1988–2001." *International Studies Quarterly* 55 (1): 47–69.

Steiner, Nils D. 2016. "Comparing Freedom House Democracy Scores to Alternative Indices and Testing for Political Bias: Are US Allies Rated As More Democratic By Freedom House?" *Journal of Comparative Policy Analysis* 18 (4): 329–49.

Stroup, Sarah S. 2012. *Borders Among Activists: International NGOs in the United States, Britain, and France*. Ithaca, NY: Cornell University Press.

Thomas, Melissa A. 2010. "What Do the Worldwide Governance Indicators Measure?" *European Journal of Development Research* 22 (1): 31–54.

Winters, Matthew S. 2010. "Choosing to Target: What Types of Countries Get Different Types of World Bank Projects." *World Politics* 62 (3): 422–58.

Zeeuw, Jeroen de. 2005. "Projects Do Not Create Institutions: The Record of Democracy Assistance in Post-Conflict Societies." *Democratization* 12 (4): 481–504.