

The Uses and Abuses of Democracy Ratings

By Sarah Bush

Few concepts are more important to the study of politics than democracy. Despite—or perhaps because of—democracy’s importance, measuring it is highly controversial. In this essay, I draw on my past research examining a particularly influential and contentious measure of democracy: the annual *Freedom in the World* (FITW) country ratings.¹ The research finds that the FITW ratings have authority in real-world politics not because of their epistemic quality or independence but because of their affinity with how U.S. policymakers think about democracy. An important take-away from my research is that scholars should approach their choice of democracy indicators for empirical analysis carefully; for certain research questions, a measure of democracy that is close to how policymakers understand this concept may make sense, but in other cases it might bias the analysis in worrying ways. I close the essay by suggesting directions for future research related to democracy ratings on which the insights of scholars of the Middle East would be especially welcome.

The Politics of Measuring Democracy

An American non-governmental organization, Freedom House, created the FITW ratings in 1972. FITW is an annual report that assigns countries overall ratings—“free,” “partly free,” and “not free”—as well as numerical scores for their political rights and civil liberties. Initially, the ratings were compiled by social scientist Raymond Gastil with assistance from his wife, Jeanette Gastil. Over time, the number of experts consulted for the creation of the report has expanded considerably, and the FITW methodology has become much more detailed and rigorous.

Although numerous democracy indicators exist, FITW enjoys special prominence, especially in the United States. FITW has shaped the U.S. State Department’s human rights reports, been used to assess the effectiveness of the U.S. Agency for International Development’s assistance programs, and provided qualification criteria for the Millennium Challenge Corporation. Given that FITW is used in this way, it is also influential in countries that care about their international reputations for democracy, who sometimes seek to affect their ratings. Outside of the policy world, academics also commonly use FITW to study the causes and consequences of democratization.²

Because of the ratings’ authority, FITW has been subject to much scrutiny and criticism. One set of critiques pertains to the ratings’ lack of transparency historically, when the organization did not provide much information about its coding process.³ Since then, the methodology has improved, although some methodological critiques still remain. Another line of criticism points to the ratings’ tendency, again especially historically, to score U.S. allies more positively than

¹ Sarah Sunn Bush, “The Politics of Rating Freedom: Ideological Affinity, Private Authority, and the Freedom in the World Ratings,” *Perspectives on Politics* 15, no. 3 (2017): 711-731.

² For more details on usage of FITW, see Bush, “The Politics of Rating Freedom,” p. 718.

³ Geraldo L. Munck and Jay Verkuilen, “Conceptualizing and Measuring Democracy,” *Comparative Political Studies* 35, no. 1 (2002): 5-34.

other ratings; for example, the FITW scores for El Salvador vs. Nicaragua during the Cold War were correlated with the two countries' positions vis-à-vis the United States.⁴ Although FITW is usually compared unfavorably to the Polity measure of democracy in such exercises, it is worth noting that criticisms of U.S. bias have also been leveled against Polity.⁵ A final critique is that FITW—despite revising its coding criteria over time—remains overly committed to a liberal conceptualization of democracy, similar to U.S. policymakers, whereas other conceptualizations are possible.⁶ The Varieties of Democracy or V-Dem project offers one solution to this problem for researchers as it provides indicators of different conceptualizations of democracy (e.g., electoral, participatory, liberal) that can be used depending on the research question.⁷

I have argued in past work that applied researchers would benefit from greater recognition of these debates surrounding FITW. At the same time, my message is not that awareness of criticisms of FITW should prompt researchers to abandon it. For example, their documented historical biases may ironically make the FITW ratings *useful* democracy ratings to use for answering certain research questions. Some research questions—such as about whether the U.S. government is targeting and tailoring aid to countries' regime types—might be answered most effectively through the use of an indicator that best captures how policymakers are likely to think about democracy.⁸

Directions for Future Research

Although there is a growing literature on benchmarks in global politics, it has not yet benefited fully from the expertise of scholars of Middle East politics. I suggest two promising directions for future research on the politics of democracy ratings with a focus on the Middle East and North Africa: one concerning the effects of democracy ratings and one concerning their production.

First, researchers might study how Middle East governments respond to their Freedom House scores. International ratings have been shown to influence countries' policies and behavior through mechanisms such as domestic mobilization, elite shaming, and transnational pressure.⁹

⁴ Scott Mainwaring, Daniel Brinks, and Aníbal Pérez-Liñán, "Classifying Political Regimes in Latin America," *Studies in Comparative International Development* 36, no. 1 (2001): 37-65; and Nils D. Steiner, "Comparing Freedom House Democracy Scores to Alternative Indices and Testing for Political Bias: Are U.S. Allies Rated as More Democratic by Freedom House?" *Journal of Comparative Policy Analysis* 18, no. 4 (2016): 329-349.

⁵ Jeff D. Colgan, "American Bias in Global Security Studies," *Journal of Global Security Studies* 4, no. 3 (2019): 358-371.

⁶ Diego Giannone, "Political and Ideological Aspects in the Measurement of Democracy: The Freedom House Case," *Democratization* 17, no. 1 (2010): 68-97.

⁷ Michael Coppedge et al., "Conceptualizing and Measuring Democracy: A New Approach," *Perspectives on Politics* 9, no. 2 (2011): 247-267.

⁸ Sarah Sunn Bush, "National Bias in Quantitative Datasets: A Silver Lining?" *Journal of Global Security Studies* 4, no. 3 (2019): 372-383.

⁹ Judith G. Kelley and Beth A. Simmons, "Politics by Number: Indicators as Social Pressure in International Politics," *American Journal of Political Science* 59, no. 1 (2015): 55-70.

Many countries care about their FITW ratings, believing that the scores will influence their access to foreign aid that is conditional on democracy, among other benefits. Yet is the same true in the Middle East? Arab countries have been relatively insulated from pressure related international democracy promotion, although they have attempted to coopt some of these efforts and game others, such as through a selective embrace of certain liberal norms (e.g., related to women's rights) in order to present a modernizing image.¹⁰ It would be informative to study whether and how FITW ratings are used domestically by civil society and governments in the Middle East and North Africa to advocate for or deflect pressure to reform.

Second, because the Middle East and North Africa is the world region that is widely thought to be least democratic, it would be intriguing to study the dynamics of peer effects on experts' perceptions of democracy in the region. Countries are sometimes graded on a curve in international assessments, such as when international election observers are more lenient when countries experiencing a transitional election or a history of election violence have some malpractice.¹¹ It is possible that when a country in the Middle East shows signs of democratic opening, international experts may be overly generous given the overall state of the region. Alternatively, and more similar to the positive and negative boosts that countries get from their perceived peer groups, Arab countries' ratings may not reflect when they make real democratic progress given their association with a region that is known to be undemocratic.¹²

¹⁰ On cooptation, see Sarah Sunn Bush, *The Taming of Democracy Assistance: Why Democracy Promotion Does Not Confront Dictators* (Cambridge: Cambridge University Press, 2015). On gaming, especially in the area of women's rights, see Aili Mari Tripp, *Seeking Legitimacy: Why Arab Autocracies Adopt Women's Rights* (Cambridge: Cambridge University Press, 2019); Sarah Sunn Bush and Pär Zetterberg, "Gender Quotas and International Reputation," *American Journal of Political Science* 65, no. 2 (2021): 326-341; Daniela Donno, Sara Fox, and Joshua Kaasik, "International Incentives for Women's Rights in Dictatorships," *Comparative Political Studies* <https://doi.org/10.1177/00104140211024306>.

¹¹ Judith G. Kelley, "D-Minus Elections: The Politics and Norms of International Election Observation," *International Organization* 63, no. 4 (2009): 765-787.

¹² Julia Gray, *The Company States Keep: International Economic Organizations and Investor Perceptions* (Cambridge: Cambridge University Press, 2013).