Revisiting the Evidence on Thermostatic Response to Democratic Change: Degrees of Democratic Support or Researcher Degrees of Freedom?*

Authors

- Yue Hu, ORCID: https://orcid.org/0000-0002-2829-3971, Associate Professor, Department of Political Science, Tsinghua University, yuehu@tsinghua.edu.cn

- Yuehong Cassandra Tai, ORCID: https://orcid.org/0000-0001-7303-7443, Postdoctoral Fellow, Center for Social Data Analytics, Pennsylvania State University, yhcasstai@psu.edu

- Frederick Solt, corresponding author, ORCID: https://orcid.org/0000-0002-3154-6132, Associate Professor, Department of Political Science, University of Iowa, frederick-solt@uiowa.edu

1

Abstract

Prominent recent work argues that support for democracy behaves thermostatically—that democratic erosion boosts democratic support while deepening democracy yields public backlash—and further contends that there is no evidence for the classic argument that democracy itself increases democratic support over time. Here, we document how these conclusions depend on subtle choices in measurement coding that constitute 'researcher degrees of freedom': analyses employing alternative reasonable choices provide little or no support for the original conclusions. The fragility of the statistical results demonstrates that researcher degrees of freedom in measurement must be taken seriously and that the question of the relationship between democratic institutions and democratic support remains unsettled.

With democracy under threat in many countries around the world, how the public reacts to democratic change is a crucial question. A long-established and indeed still-vibrant literature holds that experiencing more democracy boosts democratic support in the public—in short, that democracy creates its own demand (see, e.g., Lipset 1959; Welzel 2013; Wuttke, Gavras, and Schoen 2022). In contrast to this classic theory, one prominent recent article, Claassen (2020a), argues that democratic support behaves thermostatically: that increases in democracy yield an authoritarian backlash in the public, while democratic backsliding prompts the public to rally to democracy's cause.

In support of this thermostatic argument, Claassen (2020a) offers evidence based on recent advances in modeling public opinion as a latent variable. Drawing on aggregated responses to more than sixty distinct questions on attitudes toward democracy and its alternatives that were asked in thousands of national surveys, the article estimates a latent variable of democratic support in more than a hundred countries over as long as thirty years. These data constitute a broader evidentiary base than that employed in any earlier work on this topic. Analyses of these data support the article's conclusions that democratic support behaves thermostatically in the public, that "increases in democracy dampen public mood, while decreases cheer it," and that democracy "does not appear to create its own demand" as the classic theory has it (Claassen 2020a, 51).

But, inevitably, the aggregated survey data, the resulting estimates of democratic support, and ultimately the results of the article's analyses embody a particular set of choices. These choices, known as "researcher degrees of freedom," have attracted growing attention among political scientists in recent years (see, e.g., Wuttke 2019; Breznau et al. 2022). The concern extends beyond cases of 'p-hacking,' in which researchers sift through many different options to find a set of choices that yield statistically significant results. It turns out that no intentional fraud or misconduct is required to render results unreliable. Researchers may make only a single set of entirely reasonable choices, that is, they may take only a single walk through the metaphorical "garden of forking paths"; find a result that supports their expectations; and "get excited and believe it" without even considering that a different set of entirely reasonable choices would provide different results (Gelman and Loken 2014, 464).

Here, we illustrate the importance of taking seriously researcher degrees of freedom by

replicating the primary analyses of Claassen (2020a) while varying two aspects of the measurement of democratic support. The first is the manner in which ordinal survey responses are coded as supportive or not supportive of democracy. The second is the attitude ascribed to survey respondents who did not answer items regarding their support for democracy. We document a range of reasonable alternatives to the choices made in Claassen (2020a) on each of these issues, and demonstrate that analyses employing these alternatives provide little to no support for the conclusions drawn in that article. The fragility of the published results to alternate reasonable choices demonstrates the importance of taking researcher degree of freedom seriously and that the question of the relationship between democratic institutions and democratic support remains unsettled.

Researcher Degrees of Freedom and Democratic Support

The dependent variable considered in Claassen (2020a) is the public's support for democracy, a latent variable. This latent variable was estimated from aggregated responses to many survey questions asked in many countries in many years. Here we examine the consequences of two choices made early in the process of collecting and aggregating all of those survey data: one regarding how responses were coded as supporting democracy and a second involving how missing responses were treated.

First we consider the question of how to code survey responses. The latent variable model used in Claassen (2020a) requires responses to be dichotomous: either, in this case, democracy supporting or not. This is straightforward for the few survey items that gave respondents only two options, for example, in the Pew Global Attitudes surveys, "a democratic form of government" or "a leader with a strong hand." But the vast majority of the survey items employed are ordinal and so require a cutpoint to be chosen to split respondents between those who support democracy and those who do not.

There are at least three reasonable options for recoding these ordinal responses into a dichotomous variable of democratic support. The first coding rule is the most demanding. It stipulates that only those who supply only the most democracy-supporting response of those available can be considered unflinching supporters of democracy; anything but the highest response indicates a lack of support. A second coding rule takes the opposite tack, with all

but the lowest response considered as indicating at least some support for democracy and only the lowest value considered as not supportive. The last splits the difference, considering answers above the median, e.g., the two highest values on a five-point scale, to indicate that the respondent supports democracy, with the median and below being unsupportive.

In the event, Claassen (2020a) resorts to a mix of the above three rules. Most survey items were dichotomized following the 'above the median' rule (see Claassen 2020a, Appendix 1.3). Although not documented in the article's appendix on "Microlevel Coding of Survey Responses," the other two rules were employed as well. The most demanding rule was applied to the four-point item asked in Pew Global Attitudes surveys, "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?" For this question, rather than including respondents who gave both responses above the median—"very important" and "somewhat important"—only those respondents who answered "very important" were entered as supporting democracy. And for the three-point item in the Asia Barometer asking whether respondents thought "a democratic political system" would be very good, fairly good, or bad for their country, the 'all but the lowest' rule was applied.[1] Adding this mixed rule to those described above gives us a total of four coding rules.

The second choice involves how to treat missing responses, that is, when survey respondents either indicate that they "don't know" or simply refuse to answer the question. Here, too, there is a range of reasonable choices. We identify four possibilities. One might reasonably assume that a missing response is equivalent to answering with a lack of support for democracy since the respondent did not provide a democracy-supporting response. Claassen (2020a, Appendix 1.3) takes this first approach. But one might conversely and equally reasonably consider a missing response to suggest *support* for democracy; the respondent, after all, did not provide a democracy-opposing response. A third possibility is that non-responses

---

[1]One charitable reading, offered by Hu, Tai, and Solt (2022), is that these deviations from the 'above the median' rule documented in the paper's appendix are simply unintentional mistakes. A corrigendum to the article written in response to Hu, Tai, and Solt (2022), however, rejects this interpretation (see Claassen 2023).

indicate opposition to the current regime. That is, their meaning depends on the context. In democracies missing responses indicate an unwillingness to supply an honest but socially unacceptable rejection of democracy, but in autocracies missing responses evince support for democracy coupled with a fear of reprisal. One final possibility is that non-responses occur at random. In that case, a non-response tells nothing about the respondent's views toward democracy and can simply be dropped from the sample. There are, then, at least four reasonable treatments of survey non-responses. Together, the four coding rules and four missing-data treatments described above give us sixteen combinations of plausible researcher choices.

Consequences for Inference

How does the selection among these sixteen garden paths influence our conclusions regarding the public's response to changes in democracy? To find out, we collected all of the survey data used in Claassen (2020a) from their original sources. We then applied each combination of coding rule and non-response treatment and generated sixteen sets of cross-national time-series estimates of democratic support using the latent-variable model employed in the article. Finally, we used each of these sets of estimates to replicate the article's analyses.

Figure 1 is a "small multiple" plot (see Solt and Hu 2015) showing the results of replicating Model 1.1, the principal model of Claassen (2020a, 47), with each of the sixteen combinations of coding rule and treatment of survey non-response. In the top panel, the dots represent point estimates for the coefficients for change in liberal democracy; according to the thermostatic theory, these coefficients should be negative. In the bottom panel, the dots depict point estimates for the coefficients for lagged level of liberal democracy; according to the classic theory, these coefficients should be positive. In both panels, the whiskers show the associated 95% confidence intervals. Each coding rule is represented by a different color, while the four non-response treatments are shown in separate grouped estimates from left to right.[2]

The darkest, left-most coefficients in the left-most group of estimates represent the combination of mixed coding rule and treatment of non-responses as unsupportive of democracy,

---

[2]The full tabular results can be found in the online Supplementary Materials.

Predicting the Change in Public Democratic Support
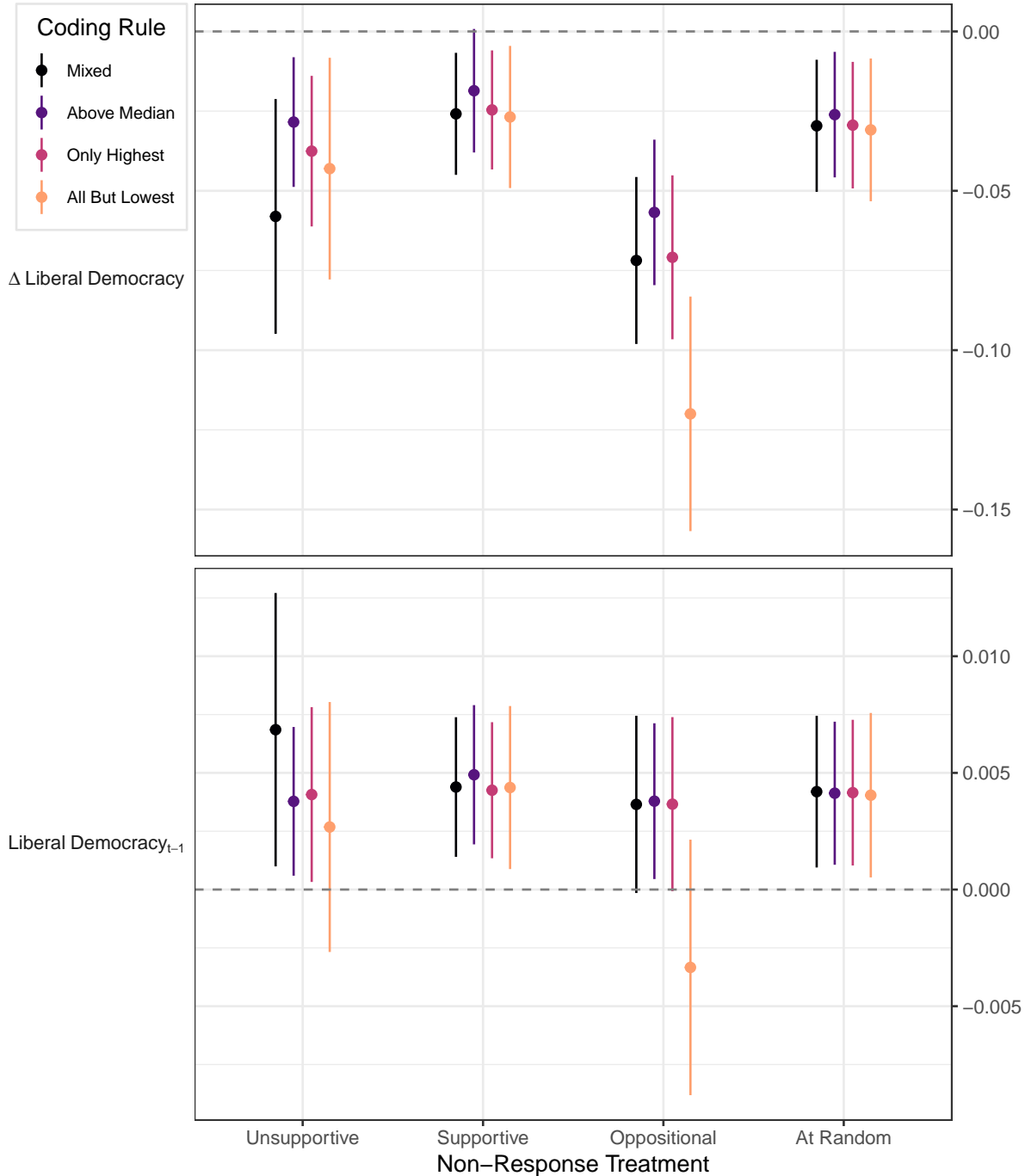
Notes: Replications of Claassen (2020, 47), Table 1, Model 1.1. The mixed coding rule employed in Claassen (2020) along with that work's assumption that non−responses indicate a lack of support for democracy yield a larger negative point estimate of the coefficient for change in liberal democracy than most other combinations and a larger point estimate of the coefficient for the lagged level of liberal democracy than all other combinations. In error−correction models like these, both coefficients must be interpreted together; see Figure 2.

Figure 1: The Effects of Democracy on the Change in Public Support

that is, the combination employed in Claassen (2020a). It replicates the results reported in that article exactly. More importantly for our purposes, this combination yields a larger negative point estimate of the coefficient for change in liberal democracy than most other combinations. It also yields a larger point estimate of the coefficient for the lagged level of liberal democracy than all of the other combinations. But like the coefficients of constitutive terms of multiplicative interactions (see, e.g., Brambor, Clark, and Golder 2006), the coefficients of error-correction models like Model 1.1 are properly interpreted together rather than separately (see Williams and Whitten 2012).

Figure 2 is similar to Claassen's (2020a, 48) Figure 5. It depicts simulated effects, in differences rather than levels for ease of interpretation, of a one standard deviation increase in democracy on the public's support for democracy using the sixteen sets of regression coefficients presented in Figure 1 with the four different coding rules appearing the columns and the four different non-response treatments in the rows.[3]

The upper left pane simulates the combination of the mixed coding rule and the treatment of survey non-responses as indicating a lack of support for democracy matches, the same combination employed in Claassen (2020a). The initial drop and slow recovery in the mean of these simulations of public democratic support was the evidence presented for the "thermostatic response of public opinion" and the claim that there is "little evidence that democracy generates its own demand" (Claassen 2020a, 48). But the other panes, with a few exceptions, show smaller dips, quicker recoveries, and continued increases; these findings lead to very different conclusions. Indeed, most of the analyses—eleven of sixteen—show statistically significant increases in democratic support within three decades, the sort of generational change predicted by the classic theory since Lipset (1959). None of them show statistically significant declines that would lend credence to the argument that democratic support responds thermostatically.

Discussion

Claassen (2020a) claims that a thermostatic response to democratic change is required to explain "democracy's fading allure" among publics. It is far from clear that support for

---

[3]Complete details on these simulations are provided in the online Supplementary Materials.

Simulating the Effect of a Change in Democracy on Public Democratic Support
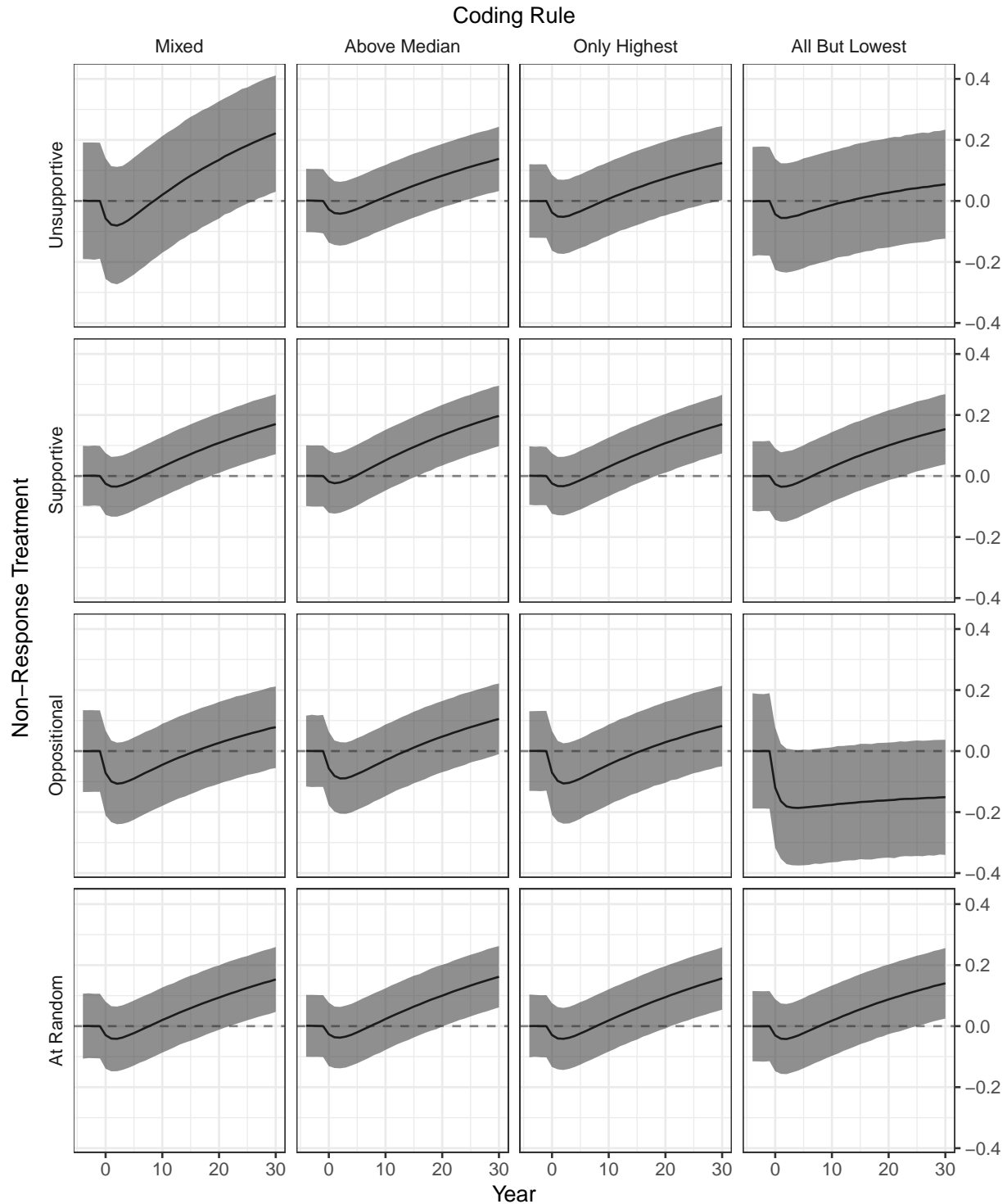
Notes: Simulated effects are estimated using coefficients from the models presented in Figure 1. The solid lines indicate the mean simulated effect; the shaded regions indicate the 95% confidence intervals of these effects.

Figure 2: Simulated Effects of Democracy on Changes in Public Democratic Support

democracy is indeed declining (for recent evidence against, see, e.g., Wuttke, Gavras, and Schoen 2022). But if it is, the fragility of the results presented in Claassen (2020a) to alternate reasonable coding decisions and non-response treatments suggests that other explanations are worth revisiting. One possibility not considered in the specification adopted in Claassen (2020a) is that rising economic inequality undermines support for democratic institutions (Magalhães 2014, 84; see also, e.g., Solt 2012). This and other theories of democratic support warrant further investigation.

Regardless, the fragility of the findings in Claassen (2020a) that is documented above has profound implications. The thermostatic theory of democratic support presented in Claassen (2020a) suggests that would-be autocrats working to erode democracy constitute little cause for concern: "should elected leaders start dismantling democratic institutions and rights, public mood is likely to swing rapidly toward democracy again, providing something of an obstacle to democratic backsliding" (Claassen 2020a, 51). To the contrary, that theory implies, it is those who attempt to deepen democracy who pose the threat, as "extending democratic rights and legal protections to minorities" will trigger a public backlash against democracy (Claassen 2020a, 51). If the theory is true, the appropriate positions for those who value democracy are, in short, complacency and satisfaction with the status quo.

That the evidence for such a perverse outcome depends entirely on researcher degrees of freedom indicates that defending democracy requires taking a very different course. Alternate combinations of plausible coding rules and non-response treatments yield no support for the thermostatic theory. In fact, most show that increases in democracy work, over time, to increase the public's support, just as suggested by the classic theory that has been advanced in scholarship over more than a half century. These findings imply that those who want to advance democracy should do just that: they should push ahead.

Acknowledgments

References

Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14 (1): 63–82.

Breznau, Nate, Eike Mark Rinke, Alexander Wuttke, Hung H. V. Nguyen, Muna Adem, Jule Adriaans, Amalia Alvarez-Benjumea, et al. 2022. "Observing Many Researchers Using the Same Data and Hypothesis Reveals a Hidden Universe of Uncertainty." *Proceedings of the National Academy of Sciences* 119 (44): e2203150119.

Claassen, Christopher. 2020a. "In the Mood for Democracy? Democratic Support as Thermostatic Opinion." *American Political Science Review* 114 (1): 36–53.

———. 2020b. "Replication Data for: In the Mood for Democracy? Democratic Support as Thermostatic Opinion." https://doi.org/10.7910/DVN/FECIO3, American Political Science Review Dataverse.

———. 2023. "In the Mood for Democracy? Democratic Support as Thermostatic Opinion-CORRIGENDUM." *American Political Science Review* 117 (4): 1531–38.

Gelman, Andrew, and Eric Loken. 2014. "The Statistical Crisis in Science." *American Scientist* 102 (6): 460–65.

Hu, Yue, Yuehong Cassandra Tai, and Frederick Solt. 2022. "On Data 'Janitor Work' in Political Science: The Case of Thermostatic Support for Democracy." SocArXiv. https://doi.org/10.31235/osf.io/kd7mu.

Lipset, Seymour Martin. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Political Science Review* 53: 69–105.

Magalhães, Pedro C. 2014. "Government Effectiveness and Support for Democracy." *European Journal of Political Research* 53 (1): 77–97.

Solt, Frederick. 2012. "The Social Origins of Authoritarianism." *Political Research Quarterly* 65 (4): 703–13.

Solt, Frederick, and Yue Hu. 2015. "dotwhisker: Dot-and-Whisker Plots of Regression Results." Available at the Comprehensive R Archive Network (CRAN). http://CRAN.R-project.org/package=dotwhisker.

Welzel, Christian. 2013. *Freedom Rising: Human Empowerment and the Quest for Eman-*

cipation. Cambridge: Cambridge University Press.

Williams, Laron K., and Guy D. Whitten. 2012. "But Wait, There's More! Maximizing Substantive Inferences from TSCS Models." *The Journal of Politics* 74 (3): 685–93. https://doi.org/10.1017/S0022381612000473.

Wuttke, Alexander. 2019. "Why Too Many Political Science Findings Cannot Be Trusted and What We Can Do About It: A Review of Meta-Scientific Research and a Call for Academic Reform." *Politische Vierteljahresschrift* 60 (1): 1–19.

Wuttke, Alexander, Konstantin Gavras, and Harald Schoen. 2022. "Have Europeans Grown Tired of Democracy? New Evidence from Eighteen Consolidated Democracies, 1981–2018." *British Journal of Political Science* 52 (1): 416–28.

Online Supplementary Materials

Contents

A   Tabular Results

Table SI.1: Predicting the Change in Public Democratic Support

| | Non-Response Treatment | | | |
	Unsupportive	Supportive	Oppositional	At Random
Mixed Coding Rule				
Democratic Mood (t-1)	0.473***	0.512***	0.437***	0.499***
	(0.018)	(0.017)	(0.018)	(0.017)
Democratic Mood (t-2)	-0.487***	-0.522***	-0.455***	-0.510***
	(0.018)	(0.017)	(0.018)	(0.017)
Liberal Democracy (Difference)	-0.058**	-0.026**	-0.072***	-0.030**
	(0.019)	(0.010)	(0.013)	(0.011)
Liberal Democracy (t-1)	0.007*	0.004**	0.004+	0.004*
	(0.003)	(0.002)	(0.002)	(0.002)
Log GDP Per Capita (Difference)	0.063	-0.013	0.027	-0.004
	(0.046)	(0.024)	(0.032)	(0.025)
Log GDP (t-1)	0.003	0.001	0.002	0.001
	(0.002)	(0.001)	(0.002)	(0.001)
R^2	0.256	0.296	0.247	0.279
RMSE	0.10	0.05	0.07	0.05
Above Median Coding Rule				
Democratic Mood (t-1)	0.521***	0.516***	0.501***	0.538***
	(0.017)	(0.017)	(0.017)	(0.017)
Democratic Mood (t-2)	-0.531***	-0.528***	-0.514***	-0.548***
	(0.017)	(0.017)	(0.017)	(0.017)
Liberal Democracy (Difference)	-0.028**	-0.019+	-0.057***	-0.026**
	(0.010)	(0.010)	(0.012)	(0.010)
Liberal Democracy (t-1)	0.004*	0.005**	0.004*	0.004**
	(0.002)	(0.002)	(0.002)	(0.002)
Log GDP Per Capita (Difference)	0.016	-0.018	0.007	0.004
	(0.025)	(0.024)	(0.028)	(0.024)
Log GDP (t-1)	0.000	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
R^2	0.304	0.305	0.300	0.324
RMSE	0.05	0.05	0.06	0.05

$+ \ p < 0.1, \ ^* \ p < 0.05, \ ^{**} \ p < 0.01, \ ^{***} \ p < 0.001$

	Non-Response Treatment			
	Unsupportive	Supportive	Oppositional	At Random
Only Highest Coding Rule				
Democratic Mood (t-1)	0.456***	0.523***	0.448***	0.515***
	(0.018)	(0.017)	(0.018)	(0.017)
Democratic Mood (t-2)	-0.470***	-0.533***	-0.465***	-0.525***
	(0.018)	(0.017)	(0.018)	(0.017)
Liberal Democracy (Difference)	-0.038**	-0.025**	-0.071***	-0.029**
	(0.012)	(0.010)	(0.013)	(0.010)
Liberal Democracy (t-1)	0.004*	0.004**	0.004+	0.004**
	(0.002)	(0.001)	(0.002)	(0.002)
Log GDP Per Capita (Difference)	0.048+	-0.009	0.028	0.002
	(0.029)	(0.023)	(0.032)	(0.024)
Log GDP (t-1)	0.002	0.001	0.002	0.001
	(0.002)	(0.001)	(0.002)	(0.001)
R^2	0.241	0.306	0.256	0.295
RMSE	0.06	0.05	0.07	0.05
All But Lowest Coding Rule				
Democratic Mood (t-1)	0.343***	0.478***	0.349***	0.480***
	(0.019)	(0.017)	(0.018)	(0.017)
Democratic Mood (t-2)	-0.368***	-0.492***	-0.374***	-0.493***
	(0.019)	(0.017)	(0.018)	(0.017)
Liberal Democracy (Difference)	-0.043*	-0.027*	-0.120***	-0.031**
	(0.018)	(0.011)	(0.019)	(0.011)
Liberal Democracy (t-1)	0.003	0.004*	-0.003	0.004*
	(0.003)	(0.002)	(0.003)	(0.002)
Log GDP Per Capita (Difference)	0.054	-0.027	0.024	-0.019
	(0.043)	(0.027)	(0.045)	(0.028)
Log GDP (t-1)	0.007**	0.003*	0.007**	0.003*
	(0.002)	(0.001)	(0.002)	(0.001)
R^2	0.157	0.275	0.190	0.272
RMSE	0.09	0.06	0.10	0.06

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

B Calculating Effects Via Simulation

To simulate the effects of changes in democracy on public support for democracy in the error-correction models (see Williams and Whitten 2012), we follow the same procedure employed in Claassen (2020a, 48–50). That is, for each of our sixteen sets of results, we first set all of the independent variables to the same moderate values used in Claassen (2020a) and ran the model for 200 years, which is a sufficient time for the system of equations to stabilize. Next, the value of democracy was increased one standard deviation, from its previous value of a half standard deviation below its mean to a half standard deviation above its mean. Finally, the system of equations was allowed to run for thirty additional years. These thirty years of simulated results are shown in the sixteen panes of Figure 2 in the text. Again following Claassen (2020a, Supplementary Information 3) and Claassen (2020b), the uncertainty in the model was captured by taking 10,000 draws from a multivariate normal distribution with the expectation that the vector of model coefficients and variance constitute the robust covariance matrix, $\tilde{\Theta} \sim MVN(\Theta, \Sigma)$, and adding the noise estimated in the regression standard error, $\tilde{Y}_i \sim N(X_k\tilde{\Theta}_{ki}, \sigma)$. To get first differences, the mean value of $\tilde{Y}_i$ in the year before the increase in democracy ($t = -1$) was subtracted from each $\tilde{Y}_i$, and the 0.025 and 0.975 quantiles of the first difference were used as its lower and upper confidence bounds.