

Correction Memo:

Democracy, Public Support, and Measurement Uncertainty

Yuehong Cassandra Tai, Yue Hu, and Frederick Solt

2023-06-05

Thank you for the opportunity to address the issues of (1) the inclusion of nonrepresentative surveys in the expanded dataset, (2) the treatment of non-responses, and (3) the lack of standardization of survey weights. It is perhaps worth noting that, soon after our letter was published last spring, Christopher Claassen shared with us via email that he had identified these issues. He offered no evidence or indication that he believed that any or all of them were of consequence, and we did not think they were either; as both he and our team are continuing to work in this area, we understood our exchange as instead a friendly effort to ensure these points would be adequately addressed in the future research both he and we conduct. At your prompting, we have now attended to these concerns specifically in the context of our letter, and we can confirm that they are inconsequential to its conclusions. They do suggest, however, that revisions would be warranted to the replication materials, to the online supplemental materials, and to both plots and a few numbers in the text of the letter itself.

Nonrepresentative Surveys

The Issue. In the 2002 and 2005 Pew Global Attitudes Surveys, the codebooks identify a total of nineteen national surveys with samples that are biased toward urban areas. Although Claassen (2020a, 2020b) identified and excluded seventeen of the nineteen (inadvertently, apparently, including two), we did not catch this issue at all and included all of these national surveys in our expanded sample.

Resolution. Data from these nineteen national surveys—including the two that were included in the sample used in Claassen (2020a, 2020b)—are now excluded from our data.

Non-Responses

The Issue. In Claassen (2020a, 2020b), all respondents who did not respond to a survey item are assumed to not support democracy and an undemocratic response is (usually but not always, as we noted in the OSM at page A2; see also Hu, Tai, and Solt (2022) at 4) singly imputed to them. Single imputation has long been well understood to incorporate overly strong assumptions and understate measurement error (see, e.g., Rubin 1987). In our letter, we treated these non-responses as ‘missing at random’—that is, that the distribution of attitudes among those who did not respond was similar to that among those who did—and applied listwise deletion, excluding the non-responses from the sample. This was a mistake: like single imputation, listwise deletion is known to understate measurement error (again see Rubin 1987 and many others). So neither approach is suitable, especially not in a letter on the importance of avoiding understating measurement error.

Resolution. In some cases (for example, the Americas Barometer survey of Canada in 2010; see Hu, Tai, and Solt (2022) at 5), non-responses are the result of split samples, where many respondents were not even presented with the survey item in question. In such cases, imputing undemocratic responses to these respondents is clearly inapt. Instead, they are missing at random—respondents were explicitly randomly assigned to the split samples that excluded the relevant item—and so listwise deletion is appropriate. We therefore continue to exclude from the sample those respondents who were not even asked the relevant question. These cases are rare, though; most non-responses in these surveys are the result of refusal, ‘don’t know’, and other non-responsive answers.

For these remaining cases, it is *multiple*, not single, imputation that is the preferred method for addressing such missing data and the measurement error that results (see, e.g., Rubin 1987). Typically, multiple imputation of survey non-response involves building a model of the missing data with the other variables of interest from the survey. This preserves the relationships among these survey variables and allows the measurement uncertainty that remains after modeling to propagate to the quantity of interest (see, e.g., King et al. 2001; Blackwell, Honaker, and King 2017). Here, however, there is only one variable of interest (or, sometimes, a few) in each of thousands of national surveys, and building such an MI model for every survey is not practicable.

The single-imputation approach employed in Claassen (2020a, 2020b) is inappropriate on its own, but it points the way to a feasible solution that, unlike either single imputation or listwise deletion, incorporates measurement uncertainty in the proportion of the population that supports democracy. To assume that all respondents who did not answer an item do not support democracy is not theoretically grounded: indeed, it is often argued that non-response to such politically sensitive items in nondemocracies reflects a fear of government reprisal for holding *pro-democratic* views (see, e.g., Shen and Truex 2021). Nevertheless, the Claassen assumption sets a lower bound on the proportion of the population who support democracy as evidenced by a particular survey item. Conversely, assuming that all non-responses reflect unvoiced *support* for democracy establishes an upper bound on democratic support indicated by an item. A more theoretically-driven assumption than either of these is that non-response in democratic contexts reflects social desirability bias and so a lack of support for democracy, while non-response in non-democratic contexts reflects fear of reprisal (or perhaps again social desirability bias, only yielding different results in a different political setting) and so democratic support. An intermediate possibility, of course, is that attitudes among those who did not respond to an item are actually similar to the attitudes of those who did; that is, that the data are missing at random and non-responses are appropriately listwise deleted.

We therefore now incorporate the measurement uncertainty due to non-response as follows. The number of democracy-supporting responses and the total sample size was imputed four times—that is, using in turn each of these four assumptions about the distribution of non-responses described above—for each country-year-item in the source survey data. The latent variable was then estimated using each of these imputations, and the resulting draws combined as in model-based multiple imputation. The result is a distribution of draws that reflects the measurement uncertainty in the latent variable, avoiding the strong assumptions of either single imputation or listwise deletion.

Survey Weights

The Issue. Claassen (2020a, 2020b) ignored survey weights (well, again usually but not always; see Hu, Tai, and Solt (2022) at 4-5), instead using the unweighted survey data. This is problematic, of course, as the unweighted data may not be representative of the underlying population. In our data collection, we corrected this by employing each survey’s weights, but did not appreciate that some surveys’ weights are not standardized; that is, applying them caused (occasionally dramatic) changes in the sample size. As the sample size is an important input in the latent variable model—smaller samples yield larger uncertainty in the population mean, which then propagates into the estimate of democratic support—this is undesirable.

Resolution. We now standardize survey weights to have a mean of one before applying them and so ensure that these weights preserve the sample size.

Conclusions

Applying these corrections does not yield substantive differences in the results of our analyses and so does not affect the conclusions reached. But “no substantive differences” does not mean “no differences.” As one would anticipate, all of the regression results change marginally, and the few numbers we included in our description of the expanded dataset are no longer accurate either. Updating the replication materials to reflect these corrections (and, as a bonus, make them more user friendly); revising the tables in OSM B and the figures and tables in OSM D; adding a discussion of non-responses to OSM A; and updating Figure 1, Figure 2, and the specific numbers reported under “Adding More Data” in the text itself would be warranted.

Drafts of the revised text and OSM are attached, and we have prepared the updated replication materials for upload to the Dataverse. Please let us know how you would like us to proceed.

References

- Blackwell, Matthew, James Honaker, and Gary King. 2017. “A Unified Approach to Measurement Error and Missing Data: Overview and Applications.” *Sociological Methods & Research* 46 (3): 303–41.
- Claassen, Christopher. 2020a. “Does Public Support Help Democracy Survive?” *American Journal of Political Science* 64 (1): 118–34.
- . 2020b. “In the Mood for Democracy? Democratic Support as Thermostatic Opinion.” *American Political Science Review* 114 (1): 36–53.
- 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>.
- King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. 2001. “Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation.” *American Political Science Review* 95 (1): 49–69.
- Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: J. Wiley & Sons.
- Shen, Xiaoxiao, and Rory Truex. 2021. “In Search of Self-Censorship.” *British Journal of Political Science* 51 (4): 1672–84.