Data Wrangling Before Harmonization: Best Practices for Getting Past the 'Janitor Work'*

Yue Hu[1] Yuehong Cassandra Tai[2] Frederick Solt[3]

This article focuses on a preliminary step in any ex-post data harmonization project—wrangling the pre-harmonized data—and suggests best practices for helping scholars avoid errors in this often-tedious work. To provide illustrations of these best practices, the article uses the examples of pre-harmonizing procedures used to produce the Standardized World Income Inequality Database (SWIID), a widely used database that uses Gini indices from multiple sources to create comparable estimates, and the Dynamic Comparative Public Opinion (DCPO) project, which creates a workflow for harmonizing aggregate public opinion data.

[1] Department of Political Science, Tsinghua University, Beijing, China

[2] Center for Social Data Analytics, Pennsylvania State University, University Park, USA

[3] Department of Political Science, University of Iowa, Iowa City, USA

*Corresponding author: yuehong-tai@uiowa.edu. Current version: March 28, 2025. Replication materials and complete revision history may be found at https://github.com/fsolt/wrangling_data. The authors contributed equally to this work. Yue Hu appreciates the funding support from the National Natural Science Foundation of China (72374116) and Tsinghua University Initiative Scientific Research Program (2024THZWJC01).

1 The Problem with ‘Janitor Work’

Most data harmonization projects—and a growing volume of other political science research—can be characterized as data science: they employ large quantities of data, often drawn from a large number of different sources. For such projects, data wrangling, the task of getting these data into the format required to perform harmonization or analysis, is notoriously the bulk of the work (see, e.g., Lohr 2014). Such ‘janitor work’ is often viewed as tiresome, as something to be delegated to research assistants, to someone—indeed anyone—else (see Torres 2017). Data wrangling is, however, critically important to scientific inquiry, and errors that arise during this process can undermine our data-harmonization goals.

One kind of data-wrangling error presents a particularly insidious problem: errors that occur during manual data entry. Faced with the task of getting data into the correct format before harmonizing, even some very sophisticated researchers will conclude that the most straightforward means to that end is to simply copy the needed data into a spreadsheet by hand. This technique may be straightforward, but it is very much prone to error. Barchard and Pace (2011) found that ‘research assistants’ assigned in an experiment to carefully enter data manually, even those instructed to double-check their entries against the original, had error rates approaching 1% in just a single roughly half-hour session. Rates likely go up as the tedious task goes on. Although the pernicious consequences of data-entry errors are easily grasped in everyday contexts—Haegemans, Snoeck, and Lemahieu (2019, 1) collects examples of misrouted financial transactions and airline flights—they have thus far gained little attention in political science, even among those working to harmonize large quantities of data.

We suggest three best practices for reducing the rate of data-entry errors. First, *automate data entry* to the greatest extent possible. Second, *use the double-entry method*: when manual data entry cannot be avoided, each entry should be made twice, either by separate researchers or sequentially. Third, *embrace teamwork* for any project involving entering data by hand, splitting the task up among team members will reduce the risk of errors going undetected. We demonstrate the application of these practices within two ongoing harmonization efforts, the Standardized World Income Inequality Database (SWIID) and the Dynamic Comparative Public Opinion (DCPO) project.

2 Wrangling Income Inequality Data for Harmonization

The Standardized World Income Inequality Database (SWIID) is a long-running project that seeks to provide harmonized income inequality statistics for the broadest possible coverage of countries and years (Solt 2009, 2015, 2016, 2020a). As of its most recent update at the time of this writing, its source data consists of some 27,000 observations of the Gini coefficient of income distribution in nearly 200 countries over as many as 65 years, collected from over 400 separate sources including international organizations, national statistics bureaus, and academic studies.[1]

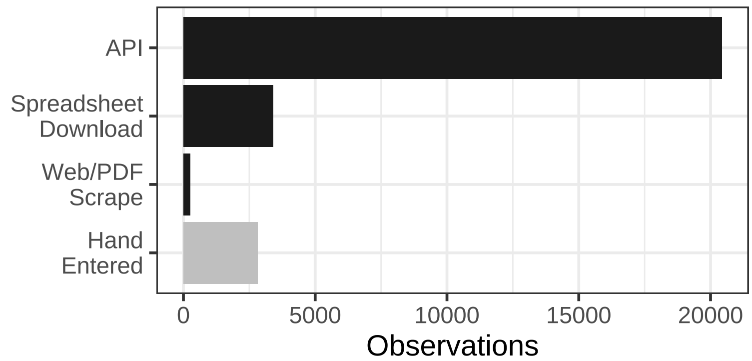


Figure 1: Income Inequality Observations by Method of Collection

In early versions of the SWIID, all of these source data were entered by hand, and checks of newly entered observations revealed that error rates were high. Moreover, many of the sources consulted frequently update or revise their figures. To avoid data-entry errors and ensure that updates and revisions are automatically incorporated as they become available, since 2015 the process of collecting the source data has been automated to the greatest extent practicable (Solt 2020a, 1184–85).[2]

Most international organizations and a few national statistical bureaus use application programming interfaces (APIs) that facilitate automating the inclusion of their data; the R community has often built packages using these APIs to make the task even easier (see Blondel 2018; Lahti et al. 2017; Lugo 2017; Magnusson, Lahti, and Hansson 2014; Wickham, Hester, and Ooms 2018). The SWIID takes as much advantage of these resources as possible, as shown in Figure 1. Although the sources with APIs are relatively few, they contain by far the most data: 76% of the observations are collected in this way. When no API is available, the automation script downloads and reads any available spreadsheets (see Wickham 2016). In the absence of a spreadsheet, the process of scraping the data either directly from the web or, preferably, from a pdf file (see Sepulveda 2024) is automated. Together the collection of 90% of the source data is scripted. This means not only that the possibility of errors introduced by hand entry for a vast majority of observations is eliminated but also that the updates and revisions that are frequent in these data are automatically incorporated as they become available.

However, it also means that some 10% of the observations are entered by hand.[3] Many sources contain just a handful or fewer observations, making the payoff to the often laborious process of data cleaning too small to justify the effort. Some sources—including most academic articles—are behind paywalls, making automation particularly challenging. When these sources contain more than a handful of observations, these are still collected using Sepulveda's (2024) `tabulapdf` R package to avoid data-entry errors. Other sources, such as many books, cannot be read directly into R. And finally, one source contains crucial information encoded in the typeface of its tables (see Mitra and Yemtsiv 2006, 6); this information would be lost if the tables were read directly into R. All such new observations are entered twice into separate spreadsheets. Most often this has been done by two different investigators, but sometimes sequentially by a single researcher. Either way, using this double-entry method allows for automated cross-checks of the newly entered data that increase the chances that errors are identified and corrected (see Barchard and Pace 2011).

To summarize, the process of collecting the source data for the SWIID is 90% automated, and the dual-entry method is employed for the remaining 10%. The upshot of this process is that the SWIID's harmonized estimates of income inequality are reliable, frequently updated, and employed around the world by international organizations, central banks, and other researchers in academia and beyond.

3 Wrangling Public Opinion Data for Harmonization

Scholarship on comparative public opinion only rarely benefits from relevant items asked annually by the same survey in many countries (see, e.g., Hagemann, Hobolt, and Wratil 2017). To address the lack of cross-national and longitudinal data on many topics, a number of works have presented latent variable models that harmonize available but incomparable survey items (see Caughey, O'Grady, and Warshaw 2019; Claassen 2019;

Kołczyńska et al. 2024; McGann, Dellepiane-Avellaneda, and Bartle 2019; Solt 2020b). This approach has been used to generate cross-national time-series measures of public opinion on a range of topics, from economic, social, and immigration conservatism (Caughey, O'Grady, and Warshaw 2019) to trust in government (Kołczyńska et al. 2024). The Dynamic Comparative Public Opinion (DCPO) model presented in Solt (2020b) in particular has been employed to measure gender egalitarianism (Woo, Allemang, and Solt 2023), political interest (Hu and Solt 2024), and support for gay rights (Woo et al. 2025), among other aspects of public opinion (see https://dcpo.org/).

This sort of work is by nature extremely data intensive, drawing on information extracted from dozens if not hundreds of survey datasets. In DCPO projects, the collection of the source data proceeds in two steps: first, identifying the relevant survey items, and second, accumulating the survey responses.

The first step, identifying the relevant survey items, involves finding surveys that contain questions on the topic of interest and then recording in a spreadsheet the survey, the variable representing the question of interest in the dataset, the question text, the response values ordered from least to most of the concept being investigated, and the original textual response categories. At present, this step is done entirely by hand. To minimize data-entry errors, the dual-entry method is used here too. Multiple collaborators go through the process separately, and the resulting spreadsheets are compared to catch the omnipresent data-entry errors.

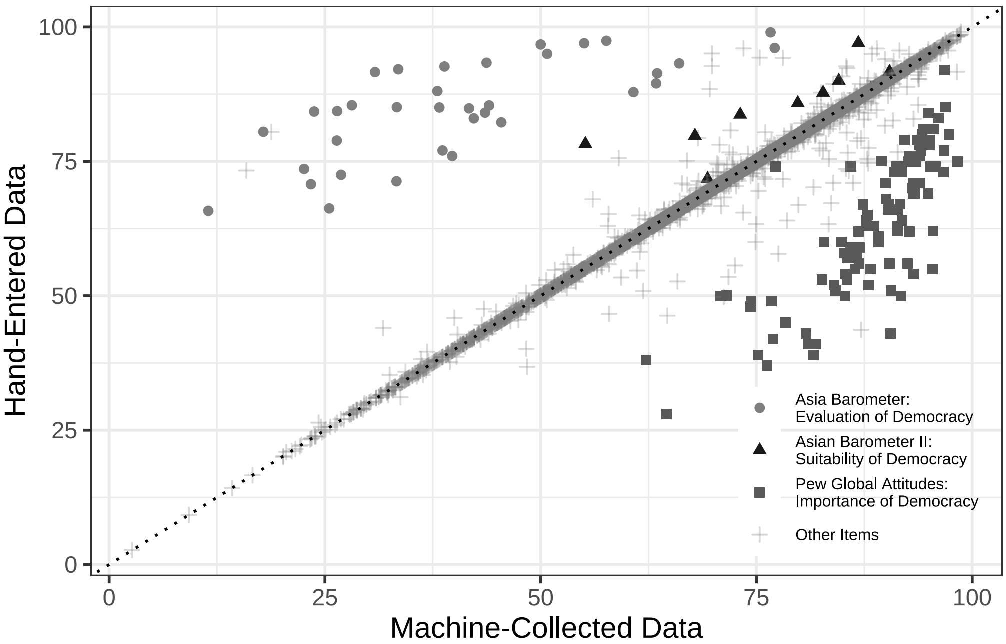
The second step, accumulating the survey responses, is fully automated through the use of the `DCPOtools` R package (Solt, Hu, and Tai 2018). When passed the spreadsheet created in the first step, this package reads in each recorded survey dataset, extracts the variable of interest, reorders the response values for this variable from least to most of the concept investigated, applies survey weights, and then aggregates the weighted number of respondents in each of the reordered response categories in each country for each year in which the survey was fielded. `DCPOtools` also automatically ensures that country names are standardized using the excellent `countrycode` package for R (Arel-Bundock, Enevoldsen, and Yetman 2018) and that the years accurately reflect actual fieldwork dates using internal crosswalk tables. The aggregated number of respondents for each observed response-item-country-year then serve as the source data for the latent variable model.

To illustrate the advantages of this more robust process relative to manual entry, we reexamine the source data for latent variable estimates of public support for democracy that were used first in a pair of very prominent publications (see Claassen 2020a, 2020b) and subsequently in additional studies (see Claassen and Magalhães 2022; Jacob 2025). These source data consist of thousands of nationally aggregated responses to dozens of different questions asked in cross-national survey projects that were apparently copied by hand into a spreadsheet from the crosstabulations included in the surveys' codebooks (see Claassen 2020c). Two pieces of data were collected for each distinct survey item in each country and year it was asked: the number of respondents to give a democracy-supporting response—defined, for ordinal questions, as those above the median value of the response scale (Claassen 2020b, Appendix 1.3)—and the total number of respondents to whom the question was posed.

We re-collected all of these source data from the original surveys. Following the two-step process described above, for each of the 52 survey items in the source data, we first identified the variables, response categories, and so on, and then we used `DCPOtools` to collect the needed data from the survey datasets by machine. In Figure 2, we compare the percentage of respondents to give a democracy-supporting response in the original hand-entered spreadsheet with the percentage in the machine-collected data. When points fall along the plot's dotted line, it indicates that the hand-entered and machine-collected source data report the same percentages. Points above this diagonal represent observations for which the hand-entered data are higher than the percentage calculated directly from the survey datasets, while points below this line are observations where the hand-entered data are lower than this machine-collected percentage.

For 85% of the country-year-item observations, the difference between these percentages was negligible—less than half a percent—yielding points approximately along the plot's dotted line. But for the remaining observations, the difference was often substantial due to data-entry errors in the hand-entered data. For three survey items contributing observations to a total of 135 country-years, the differences were found to result from deviations from the original study's coding rules: for two items the median category was mistakenly also counted as a democracy-supporting response, and for one an above-median category was not (for details, see online Appendix A).

While these issues involve mistakes in recording the numerator of the percentage, the



Notes: Each point represents the percentage of respondents in a country-year to give a democracy-supporting response to a particular survey item. Hand-entered data is as reported in Claassen (2020c); the machine-collected data was collected directly from the original surveys. The Asia Barometer's item on the evaluation of democracy accounts for most overreports, and the Pew Global Attitudes item on the importance of democracy accounts for most substantial underreports. In both cases, as well as the overreports of the suitability of democracy item in the second wave of the Asian Barometer, the issues can be easily explained by errors in transcribing the data in accordance with the reported coding rule. Deviations in other items result from inconsistent treatment of missing data and/or survey weights, reflecting in part differences in codebook reporting practices across surveys.

Figure 2: Comparing Democracy-Supporting Responses in Hand-Entered and Machine-Collected Data

number of respondents who provided a democracy-supporting answer, entering the denominator, the total number of respondents asked a question, was also problematic on occasion. For example, when the Americas Barometer surveyed Canada in 2010, it asked half its sample, whether when “democracy doesn’t work,” Canadians “need a strong leader who doesn’t have to be elected through voting.” Those who were not asked the question were included in the total number of respondents. According to the study’s coding rules, refusing to answer is equivalent to answering in a fashion not supporting democracy (see Claassen 2020b, Appendix 1.3). This rule may or may not be a reasonable coding choice (see Hu, Tai, and Solt 2025), but including in this category those who were never asked the question at all is clearly a data-entry error.

Another source of data-entry errors here involves survey weights. Weighting raw survey results to maximize the extent to which they are representative of the target population is important. Relying on toplines reported in codebooks rather than the survey data itself evidently caused some mistakes in correctly entering the needed information here, as codebooks do not always take survey weights into account. These errors shifted the percentage of democracy-supporting responses in both directions, typically by relatively small amounts.

Finally, although not depicted on this plot, data-entry errors were also evident in the variable recording the year in which a survey was conducted: these typically reflected differences between the nominal year of a survey wave and when the survey was actually in the field in a particular country. This was an issue for some 9% of the country-year-item observations.

Any one of these mistakes should give pause to any researchers tempted to meet the challenge of ‘janitor work’ in harmonization research by hand entering the needed data. Correcting all of them is sufficient to undermine the results published in Claassen (2020b) on the sources of democratic support (see online Appendix B).

4 Discussion

Harmonization projects like those discussed above are data-intensive efforts, and data-wrangling ‘janitor work’ is often a substantial part of the research process. This means data-entry errors are particularly dangerous to these undertakings. The threat is a subtle one that is not easily detected. To discern it requires close scrutiny of every manual entry; merely examining the data and their distribution will uncover few errors (Barchard and Pace 2011, 1837–38).

Our three recommendations—to maximize automation, use the double-entry method, and embrace teamwork—like similar open-science prescriptions, undoubtedly take effort (see Engzell and Rohrer 2021). But as researchers automate more of their data entry, the chances that they can reuse their code in subsequent projects improve; moreover, as the examples of the API packages used by the SWIID and `DCPOtools` demonstrate, many common janitor-work chores already have been packaged as open-source software to make researchers' task even easier and more straightforward. And while the double-entry method is labor intensive, experiments have shown that it reduces error rates by thirty-fold even when done immediately after the initial collection and by the same person (Barchard and Pace 2011, 1837); this payoff justifies the extra effort. Teamwork cuts the other way. For any project involving entering data by hand, splitting the task up among team members lessens the probability of errors due to fatigue arising in the first place, and coupled with the double-entry method, allows discrepancies to be noted, discussed, and resolved correctly. Having two sets of eyes on complex materials such as survey codebooks also increases the chances that nuances of the presentation like survey weights will be uncovered.

Data-entry errors are inevitable, and even following these recommendations is unlikely to eliminate them entirely. Nonetheless, with careful attention, not only can the threat of data-entry errors to our 'janitor work', our efforts at data harmonization, and our understanding of the world be minimized, but the transparency, openness, and credibility of our research can continue to grow.

Notes

[1] Those who are interested can access and explore these data on the web at https://fsolt.org/swiid/swiid_source.html.

[2] The R code for this automated data collection can be viewed here: https://github.com/fsolt/swiid/blob/master/R/data_setup.R.

[3] The resulting spreadsheet can be found at https://github.com/fsolt/swiid/blob/master/data-raw/fs_added_data.csv.

References

Arel-Bundock, Vincent, Nils Enevoldsen, and C. J. Yetman. 2018. "countrycode: Convert Country Names and Country Codes." *Journal of Open Source Software* 3(28): 848–49.

Barchard, Kimberly A., and Larry A. Pace. 2011. "Preventing Human Error: The Impact of Data Entry Methods on Data Accuracy and Statistical Results." *Computers in Human Behavior* 27(5): 1834–39.

Blondel, Emmanuel. 2018. "rsdmx: Tools for Reading SDMX Data and Metadata."

Caughey, Devin, Tom O'Grady, and Christopher Warshaw. 2019. "Policy Ideology in European Mass Publics, 1981–2016." *American Political Science Review* 113(3): 674–93.

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

Claassen, Christopher. 2020a. "Does Public Support Help Democracy Survive?" *American Journal of Political Science* 64(1): 118–34.

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

Claassen, Christopher. 2020c. "Replication Data for: Does Public Support Help Democracy Survive?"

Claassen, Christopher, and Pedro C. Magalhães. 2022. "Effective Government and Evaluations of Democracy." *Comparative Political Studies* 55(5): 869–94.

Engzell, Per, and Julia M. Rohrer. 2021. "Improving Social Science: Lessons from the Open Science Movement." *PS: Political Science & Politics* 54(2): 297–300.

Haegemans, Tom, Monique Snoeck, and Wilfried Lemahieu. 2019. "A Theoretical Framework to Improve the Quality of Manually Acquired Data." *Information & Management* 56(1): 1–14.

Hagemann, Sara, Sara B. Hobolt, and Christopher Wratil. 2017. "Government Responsiveness in the European Union: Evidence from Council Voting." *Comparative Political Studies* 50(6): 850–76.

Hu, Yue, and Frederick Solt. 2024. "Macrointerest Across Countries." *British Journal of Political Science* Forthcoming.

Hu, Yue, Yuehong Cassandra Tai, and Frederick Solt. 2025. "Revisiting the Evidence on Thermostatic Response to Democratic Change: Degrees of Democratic Support or Researcher Degrees of Freedom?" *Political Science Research and Methods* 13(1): 237–43.

Jacob, Marc S. 2025. "Citizen Support for Democracy, Anti-Pluralist Parties in Power and Democratic Backsliding." *European Journal of Political Research* 64(1): 348–73.

Kołczyńska, Marta, Paul-Christian Bürkner, Lauren Kennedy, and Aki Vehtari. 2024. "Modeling Public Opinion over Time and Space: Trust in State Institutions in Europe, 1989-2019." *Survey Research Methods* 18(1): 1–19.

Lahti, Leo, Janne Huovari, Markus Kainu, and Przemysław Biecek. 2017. "Retrieval and Analysis of Eurostat Open Data with the eurostat Package." *The R Journal* 9(1): 385–92.

Lohr, Steve. 2014. "For Data Scientists, 'Janitor Work' Is Hurdle to Insights." *New York Times*: B4.

Lugo, Marco. 2017. "CANSIM2R: Directly Extracts Complete CANSIM Data Tables."

Magnusson, Mans, Leo Lahti, and Love Hansson. 2014. "pxweb: R Tools for PX-WEB API."

McGann, Anthony, Sabastian Dellepiane-Avellaneda, and John Bartle. 2019. "Parallel Lines? Policy Mood in a Plurinational Democracy." *Electoral Studies* 58: 48–57.

Mitra, Pradeep, and Ruslan Yemtsiv. 2006. "Increasing Inequality in Transition Economies: Is There More to Come?"

Sepulveda, Mauricio Vargas. 2024. "tabulapdf: Extract Tables from PDF Documents."

Solt, Frederick. 2009. "Standardizing the World Income Inequality Database." *Social Science Quarterly* 90(2): 231–42.

Solt, Frederick. 2015. "On the Assessment and Use of Cross-National Income Inequality Datasets." *Journal of Economic Inequality* 13(4): 683–91.

Solt, Frederick. 2016. "The Standardized World Income Inequality Database." *Social Science Quarterly* 97(5): 1267–81.

Solt, Frederick. 2020a. "Measuring Income Inequality Across Countries and over Time: The Standardized World Income Inequality Database." *Social Science Quarterly* 101(3): 1183–99.

Solt, Frederick. 2020b. "Modeling Dynamic Comparative Public Opinion." doi:10.31235/osf.io/d5n9p.

Solt, Frederick, and Yue Hu. 2015. "dotwhisker: Dot-and-Whisker Plots of Regression Results." http://CRAN.R-project.org/package=dotwhisker.

Solt, Frederick, Yue Hu, and Yuehong Cassandra Tai. 2018. "DCPOtools: Tools for Dynamic Comparative Public Opinion." https://github.com/fsolt/DCPOtools.

Torres, Rachel. 2017. "Me: Shouldn't There Be Someone in a Basement That We Just Pay to Do All This Awful Data Cleaning? Advisor: That's Who You Are."

Wickham, Hadley. 2016. "rvest: Easily Harvest (Scrape) Web Pages."

Wickham, Hadley, James Hester, and Jeroen Ooms. 2018. "xml2: Parse XML."

Woo, Byung-Deuk, Lindsey Allemang, and Frederick Solt. 2023. "Public Gender Egalitarianism: A Dataset of Dynamic Comparative Public Opinion Toward Egalitarian Gender Roles in the Public Sphere." *British Journal of Political Science* 53(2): 766–75.

Woo, Byung-Deuk, Hyein Ko, Yuehong Cassandra Tai, Yue Hu, and Frederick Solt. 2025. "Public Support for Gay Rights Across Countries and Over Time." *Social Science Quarterly* 106(1): 1–7.

Online Supplementary Materials

A Details on Data-Entry Problems in the Democratic Support Source Data

Comparing the original hand-entered dataset (Claassen 2020c) with data that were machine-collected using the `DCPOtools` package for R (Solt, Hu, and Tai 2018) revealed three survey items for which the hand-entered data did not match the data's documented coding rules. These rules indicate that responses above the median value in the response scale are to be considered as supporting democracy, while those at the median value and below are not (see Claassen 2020b, Appendix 1.3).

First, the Asia Barometer asked respondents in 35 country-years to indicate whether they thought "a democratic political system" would be very good, fairly good, or bad for their country. According to the original study's coding rules (see Claassen 2020b, Appendix 1.3), only answers above the median of the response categories should be considered as democracy supporting, yet in this case the lukewarm intermediate category was coded as supporting democracy as well. Similarly tepid responses at and below the median response category to similar questions (e.g., in the Arab Barometer, that democracy was "somewhat appropriate" for the country) were coded as not supportive, confirming that this is indeed a data-entry error. This discrepancy resulted in hand-entered percentages of democracy-supporting responses ranging from 19 to 63 percentage points higher than the data automatically collected directly from the survey datasets.

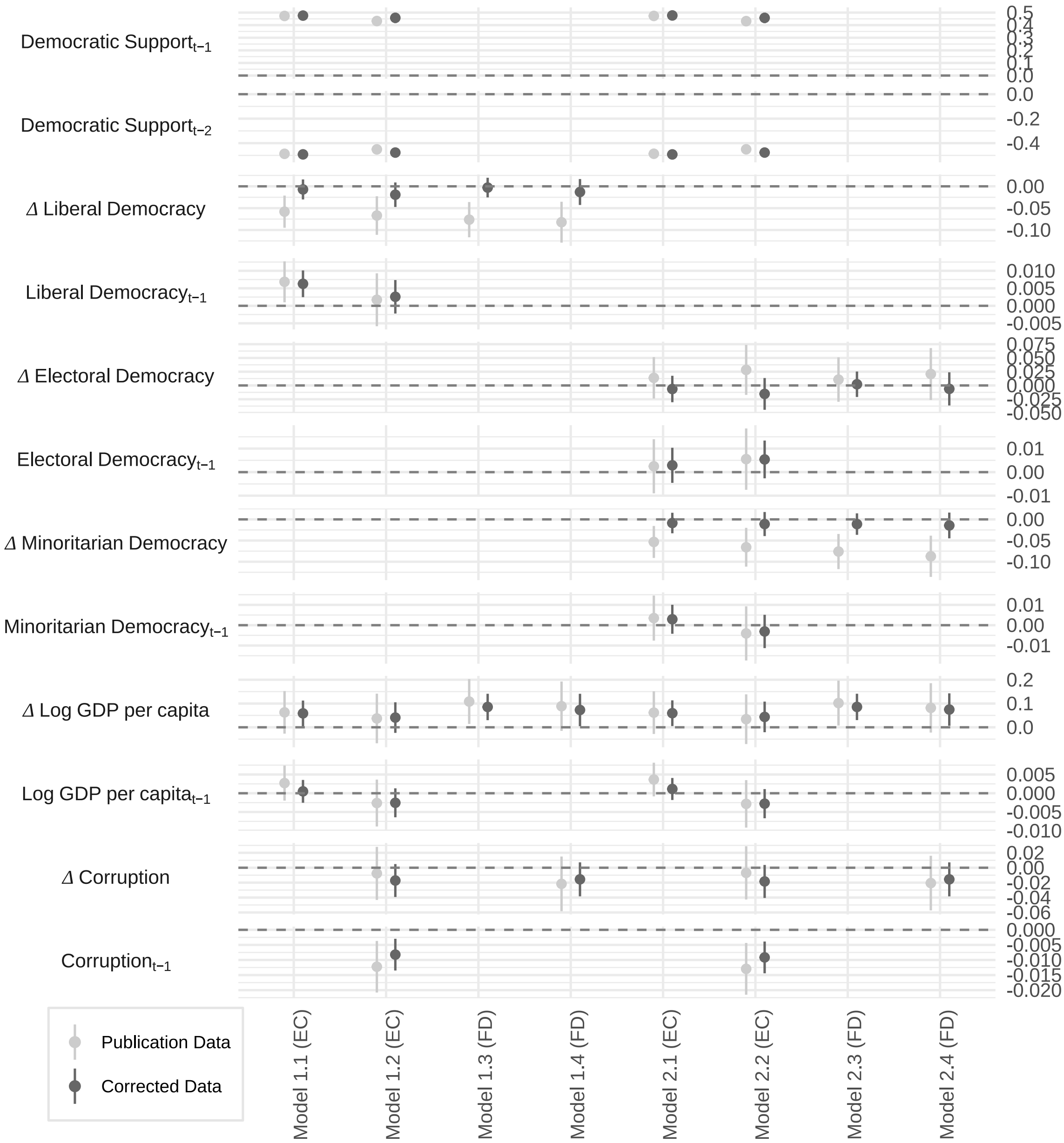
Second, the four waves of the Asian Barometer included the following item: "Here is a similar scale of 1 to 10 measuring the extent to which people think democracy is suitable for our country. If 1 means that democracy is completely unsuitable for [name of country] today and 10 means that it is completely suitable, where would you place our country today?" In accordance with the coding rules of the study, responses of 6 through 10 are considered democracy supporting, and that is how the first, third, and fourth waves of the survey are coded. For the second wave, however, 5 was erroneously also included among the democracy-supporting responses. This data-entry error resulted percentages overstated by as much as 23 percentage points in 9 country-years.

And third, the Pew Global Attitudes surveys' four-point item asking about the importance of living in a country with regular and fair contested elections: the question wording is “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?” In this case, 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. This error caused the hand-entered percentages to be substantially lower in 91 country-years.

B Consequences for Inference of Democratic Support Data-Entry Errors

Claassen (2020b) argued that public support for democracy moved thermostatically in response to changes in democracy, that is, that that changes in the latter prompt opposite changes in the former. Here we show that correcting the data-entry errors we document yield results that provide no support for this conclusion. After generating the latent variable of democratic support with these corrections, we replicated each of the models presented in Claassen (2020b) exactly using both the original and the new version of the latent variable. The results using the corrected data reveal that there is no evidence of a thermostatic relationship.

Dependent Variable: Change in Public Democratic Support



Notes: Replications of Claassen (2020), Table 1, 47, and Table 2, 49. Models denoted 'EC' are error-correction models; those marked 'FD' are first-difference models.

Figure **??** presents our results in a "small multiple" plot (Solt and Hu 2015) for a clear comparison of the coefficients of each variable in the article's models. In the plot, the dots represent point estimates and the whiskers show the associated 95% confidence intervals. Models 1.1 through 1.4, which replicate those presented in Table 1 of Claassen (2020b, 47), examine the effects of overall liberal democracy using error-correction models and first-difference models. As Claassen (2020b, 46) notes, the thermostatic theory predicts that the estimated coefficient of the change in liberal democracy will be negative, while the classic theory suggests that lagged levels of liberal democracy will be positive. When using the original publication data with their data-entry errors, we replicate the results of the article exactly: the coefficients estimated for the change in liberal democracy are large, negative, and statistically significant across all four models, just as the thermostatic theory predicts. The positive and statistically significant result for the lagged level of liberal democracy found in Model 1.1—supporting the classic theory—disappears when corruption is taken into account in Model 1.2, suggesting that "this effect is not particularly robust" (Claassen 2020b, 47).

When the data-entry errors are corrected, however, the results for these models suggest a very different set of conclusions. The standard errors shrink across the board—indicating that the models are better estimated in the corrected data—but so do the magnitudes of the coefficients. The positive and statistically significant result for the lagged level of liberal democracy remains in Model 1.1. The estimate is only slightly smaller than in the publication data, and as with the publication data, it disappears when corruption is added in Model 1.2: the evidence, such as it is, for the classic theory, operationalized as a short-run process, remains substantively unchanged. On the other hand, the estimates for the change in liberal democracy that provided support for the thermostatic theory are much smaller—very nearly exactly zero—and fail to reach statistical significance in any of these four models. Models 2.1 through 2.4, which break liberal democracy into its electoral democracy and minoritarian democracy components, similarly undermine claims for the thermostatic theory. The strong and statistically significant negative coefficients for the change in minoritarian democracy on public democratic support that are found using the publication dataset evaporate when the data-entry errors are corrected. There is no support for the thermostatic theory.