[dcpo_demsupport]*

Yuehong 'Cassandra' Tai *University of Iowa*
Yue Hu *Tsinghua University*
Frederick Solt *University of Iowa*

Incorporating Uncertainty

This means not simply employing the point estimates, that is, the means of the posterior distributions (*contra* **?**Claassen 2020).

Instead, researchers should follow the recommendations of work using other latent variables (e.g., Schnakenberg and Fariss 2014; Crabtree and Fariss 2015) and other measures incorporating uncertainty, such as the Standardized World Income Inequality Database (Solt 2020a): generate many duplicate versions of the analysis dataset, assign to each a different random draw from the posterior distributions of the variables measured with uncertainty, perform the analyses repeatedly on each of these multiple versions of the dataset, and combine the results following the rules set out in Rubin (1987). The functional programming tools in the `purrr` package (Henry and Wickham 2019) make this a matter of just a few additional lines of code.

A Better Measure of Democratic Support

The DCPO model is estimated using the `DCPO` package for R (Solt 2020b), which is written in the Stan probabilistic programming language (Stan Development Team 2019a,b).

Relative Fit

To assess the ability of the DCPO model to fit public opinion data relative to the two alternative approaches, I use the set of survey questions on support for democracy employed in Claassen (2019, 7-8).

The first three columns of Table **??** present the results of an internal validation test, that is, a test that uses the same data that was used to fit the model (see, e.g., Claassen 2019, 9). In

*Corresponding author: yuehong-tai@uiowa.edu. Current version: May 31, 2020.

column 1, the mean absolute error (MAE) measures the average difference between the observed proportion of survey respondents in country k in year t with replies to question q with a response at least as positive as response r (or with an affirmative response, in the case of Claassen's (2019) Model 5) and the model's predicted proportion across all countries, years, questions, and response categories. Given, however, that Claassen's (2019) Model 5 is fit to dichotomized data, while the Caughey et al. (2019) and DCPO models are fit to the original, possibly ordinal, survey data, which have higher variance, comparing the MAE across all three models can be somewhat misleading. Therefore, the MAE of the country means—the average proportions answering affirmatively (for Claassen's (2019) Model 5) or with a response at least as positive as response r (for Caughey et al. (2019) and DCPO) in each country across all years and all questions—serve as a baseline (see Claassen 2019, 11-12); these appear in column 2. The percentage reduction in MAE achieved by each model, listed in column 3, represents the improvement in fit compared to this baseline.

The last three columns of Table **??** present the results of an external validation test, a test of the models' ability to predict out-of-sample survey responses. This test employs k-fold validation with 10 folds, that is, it randomly divides the available country-year-questions into tenths and then sequentially treats each tenth as a test set to be predicted while fitting the model on a training set consisting of the other nine tenths of the data. Column 4 shows the mean of the MAEs of the ten resulting sets of out-of-sample predictions, while column 5 presents the mean improvement over of these ten MAEs over their respective ten country-mean MAEs. Column 6 reports the discrepancy between the percentage of all response proportions that fall within their corresponding predictions' 80% confidence intervals and the expected 80%; it therefore provides a gauge of the accuracy of the models' estimates of uncertainty. A negative value in this column indicate that less than 80% of the observed out-of-sample survey proportions are included with the model's predictions' nominal 80% credible intervals, and so its uncertainty estimates are overconfident, while a positive value indicates the opposite, that more than 80% of the observed proportions fall within the credible intervals and so the model's uncertainty estimates are overly conservative.

Comparing first the results for Claassen's (2019) Model 5 with those for the Caughey et al. (2019) model reveals that, at least for this set of survey questions, while the former has a smaller MAE, the latter accounts for a larger percentage of the variation in its source data left unexplained by its respective country-means model. This is true both in the internal validation test and on average across the ten folds of the external validation test. Again, this discrepancy is possible due to the greater variance in the survey data when its ordinal nature is preserved, as in the Caughey et al. (2019) model, rather than dichotomized, as in Claassen's (2019) approach. An examination of column 6 reveals that the Caughey et al. (2019) model yielded predictions with credible intervals that are much too narrow, encompassing only about 13% of the actual sample observations, while those for Claassen's (2019) Model 5 were slightly too conservative.[1]

However, it is the DCPO model that provides what is unambiguously the best fit. It features the smallest mean absolute error: on average, in the internal validation test its predictions are just over 3 percentage points away from the actual sample observations, and in the external validation test they are only 5.5 percentage points away from the actual out-of-sample observations. In both cases, these represent the largest percentage improvement over the country-means MAE. Further, in the external validation test, the DCPO credible intervals come closest to matching the nominal 80% level of any of the three models.

Results

DCPO results here

Conclusion

[1]The relatively poor uncertainty estimates of the Caughey et al. (2019) model are similar to those found for the dichotomous version presented in Caughey and Warshaw (2015) by Claassen (2019, 12).

References

Devin Caughey and Christopher Warshaw. Dynamic estimation of latent opinion using a hierarchical group-level irt model. *Political Analysis*, 23(2):197–211, 2015.

Devin Caughey, Tom O'Grady, and Christopher Warshaw. Policy ideology in european mass publics, 1981–2016. *American Political Science Review*, 113(3):674–693, 8 2019.

Christopher Claassen. Estimating smooth country–year panels of public opinion. *Political Analysis*, 27(1):1–20, 2019. doi: 10.1017/pan.2018.32.

Christopher Claassen. In the mood for democracy? democratic support as thermostatic opinion. *American Political Science Review*, 114(1):36–53, 2020.

Charles D. Crabtree and Christopher J. Fariss. Uncovering patterns among latent variables: Human rights and de facto judicial independence. *Research & Politics*, 2(3):1–9, 2015.

Lionel Henry and Hadley Wickham. purrr: Functional programming tools. Available at the Comprehensive R Archive Network (CRAN). `https://CRAN.R-project.org/package=purrr`, 2019.

Donald B Rubin. *Multiple Imputation for Nonresponse in Surveys.* J. Wiley & Sons, New York, 1987.

Keith E. Schnakenberg and Christopher J. Fariss. Dynamic patterns of human rights practices. *Political Science Research and Methods*, 2(1):1–31, 2014.

Frederick Solt. Measuring income inequality across countries and over time: The standardized world income inequality database. *Social Science Quarterly*, 2020a.

Frederick Solt. Dcpo: Dynamic comparative public opinion. Available at the Comprehensive R Archive Network (CRAN). `https://CRAN.R-project.org/package=DCPO`, 2020b.

Stan Development Team. RStan: the R interface to Stan. R package version 2.19.2. `https://CRAN.R-project.org/package=rstan`, 2019a.

Stan Development Team. Stan user's guide, version 2.21. `https://mc-stan.org/docs/2_21/stan-users-guide-2_21.pdf`, 2019b.