

Public Gender Egalitarianism: A Dataset of Dynamic Comparative Public Opinion Toward Egalitarian Gender Roles in the Public Sphere

Memo to Reviewers

March 25, 2022

Thank you for your helpful comments, as well as your general enthusiasm for this piece and the dataset it presents. We set out a list of the specific points raised in the reviews and our responses to them below, roughly in the order they appear in the article:

1. **Introduction** R2 asked if we were “only using cross-national surveys or whether you also incorporate single-country surveys.” We clarified that we are drawing on both cross-national and single-country surveys at the top of page 2. To try to avoid any confusion, we also revised the sentence on peak country coverage (at the bottom of p3) that prompted R2’s question.
2. **Examining the Source Data on Public Gender Egalitarianism** R1 noted three spots in the first few pages where our language was respectively unclear, mistaken, and repeated. We corrected each of these infelicities with gratitude.
3. **Estimating Public Gender Egalitarianism** R2 suggested that our description of the model could be made more accessible given the broad audience of the *BJPS* and the work’s potential for use in the classroom. We agree entirely. We revised these paragraphs (now at pages 4-5) carefully for readability, shifting much of the technical discussion to Appendix C and instead elaborating on how the model’s terms address the twin limitations of the source survey data that we raise in the paper’s previous section, their incomparability and their sparsity.
4. **Validating Public Gender Egalitarianism** R2 also advocated including more examples of the questions we used in the text rather than relegating all of them to the appendix. We are glad for this suggestion. Our new discussion of how the model addresses the incomparability of different survey items found on page 4 provided a good opportunity to bring in some specific illustrative examples.
5. **Estimating Public Gender Egalitarianism** R1 noted that the source data covers only “roughly 45 percent of the possible sample of country-years” and that it is decidedly unbalanced, and raised the question of whether this means that PGE estimates for country-years without source data were unavailable: if so, PGE would be limited in cross-national time-series applications due to listwise deletion and the complete series depicted in Figure 3 would be misleading. This is an important question. The DCPO model introduced in Solt (2020) that we employ here—as well as the models presented in Claassen (2019) that R1 referenced in making this point—was specifically designed to address the situation when survey data on the topic of interest is not available for all years in all countries. This, and the related issue

R1 raised about estimates based on only one or two is the problem referred to as sparsity. By using the DCPO model, then, we actually do have PGE estimates for country-years for which there is no source survey data available—and, more generally, the k -fold cross-validations presented in Solt (2020, 10–12) show that estimates generated by the DCPO model when there is no source data are very good: estimates for excluded source data on average differ from their actual values by only 5.5 percentage points, with more than 75% of the actual values falling within the estimated 80% confidence intervals. We have added a paragraph specifically addressing how this model deals with sparsity to ensure this point is clear to readers at pages 4-5.

6. **Estimating Public Gender Egalitarianism** R1 also raised the related issue of estimates based on only one or two survey items. This is another, though less dire, example of sparsity. As discussed above, the DCPO model does a very good job of dealing with sparse data. Again, we discuss sparsity at pages 4-5. For interested readers interested, we also plot the number of observed items for each country-year in Figure A1 in Appendix A.
7. **Estimating Public Gender Egalitarianism** R1 recommended including a discussion of difficulty and dispersion in the main text, and also suggested representing the estimates of these parameters in the text as in Gandhi and Sumner (2020, 1552). We now discuss difficulty and dispersion at page 4 to explain how the DCPO model we adopt addresses the incomparability of the available survey data. Regrettably, we were not able to carve out space in the text to represent and discuss the estimates of these parameters (understandable, we hope, given that we have some four times the number of indicators that Gandhi and Sumner (2020) employed). We present these parameters in Appendix A.
8. **Estimating Public Gender Egalitarianism** R1 requested “more description about how the estimates were obtained,” specifically details about iterations, chains, warmup, thinning, and convergence. We added a short paragraph with this information at p5.
9. **Validating Public Gender Egalitarianism** R2 asked us for more clarity on which questions we were using to generate the PGE estimates and which were only used for external validation checks. We worked to reinforce throughout the piece that the ‘source-data’ questions—those asking about gender roles in the public sphere of politics and the workplace—are the ones we used to generate estimates. We were particularly careful to underscore in this section that our internal test of convergent validation depicted in Figure 5 uses source-data items and that the initial test of construct validation shown in Figure 6 uses items *not* used to generate the PGE estimates.
10. **Appendix A: Survey Items Used to Estimate Public Gender Egalitarianism and Their Distribution** R1 and R2 both raised the issue of overlap and bridging among the survey items in the source data. Bridging and overlap have not been a concern in the literature on generating latent-variable estimates of public opinion across countries and over time from sparse data (see Claassen 2019; Caughey, O’Grady, and Warshaw 2019; Solt 2020). This is largely, we think, because the dynamic aspect of the models, that is, the random-walk prior that links the estimate for each country-year to the one before and the one following, are considered to do the work to tie together items that do not overlap (although this surely leads to increased uncertainty in these items’ parameters for dispersion and difficulty). Still, the most commonly observed item in our source data, `polileader4` (“On the whole, men make better political leaders than women do”), effectively serves as the bridge between the other items: it overlaps with 39 of the other 50 items, which together cover 91% of the observed

country-years, and only a single item (covering 22 country-years) had no observations within a year of a `polileader4` observation. We have included a short discussion of overlap and bridging in Appendix A.

We also took advantage of the opportunity to add more source data, increasing the number of observed country-year-items from 2,710 to 2,999. As we wrote at page 11, we will continue to update the source data and the PGE estimates as more data become available.

Thank you once more for the opportunity to make revisions. We think the paper is much stronger as a result of your comments, and we hope you agree.

1 References

- 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.
- Gandhi, Jennifer, and Jane Lawrence Sumner. 2020. “Measuring the Consolidation of Power in Nondemocracies.” *Journal of Politics* 82 (4): 1545–58.
- Solt, Frederick. 2020. “Modeling Dynamic Comparative Public Opinion.” SocArXiv. <https://osf.io/preprints/socarxiv/d5n9p>.