

Macrodiscontent Across Countries

Response to Reviewers

We are grateful to the reviewers for their thoughtful reviews, their enthusiasm for this project, and their suggestions on how to improve the paper. We believe that the questions and the resulting changes to the manuscript have greatly improved the accessibility of the manuscript to future readers. In the following, we respond to each of their individual queries.

1.1. Reviewer 1 worried that the DCPO model resembles a “magic trick” that stitches together fundamentally incomparable surveys, smooths them, and draws overly strong conclusions. “There is a reason no one has done this before,” R1 suggested, “and it may be that the assumptions required to make it work are quite heroic.” R1’s concern is far from trivial, and it is one that many readers are likely to share.

We have done our best to put this concern to rest. First, we note at page 8 that the DCPO model and similar efforts to estimate latent variables of public opinion using diverse survey questions and aggregate IRT models actually *have* been done before and in fact constitute a growing body of research. And second, in more direct response to R1’s suggestion of more transparency, we provide a fuller discussion of the model both in the text (at pages 9-10) and in a specific section in the supplementary material (Appendix A2). This discussion now explains each of the model’s three parameter sets functionally—i.e., what specific incomparability problem each one resolves. In short, difficulty places all survey items on a common scale by estimating how much of the latent trait each response option expresses, even when questions differ in wording and response formats. Dispersion downweights noisier questions so that poorly performing items do not distort the estimates—directly countering the concern that the model simply averages over incompatible signals. Finally, country-specific item bias absorbs translation effects, cultural response-style differences, and other country-level idiosyncrasies that make the same question non-equivalent across contexts.

In the appendix, we also address specifically that surveys contributing to a given country-year often differ in sampling design, population definitions, and weighting. It explains (at A2.3) that survey weights are incorporated where available, notes the absence of poststratification, and characterizes the resulting limitation honestly—pointing readers to the validation tests in the text for evidence that residual design effects do not undermine the estimates. Further, the revised text explains that item-level missingness reduces the effective sample size entering the beta-binomial likelihood. When nonresponse is extensive, posterior uncertainty widens, automatically limiting that observation’s influence on the country-year estimate. We also point out two channels through which the model handles sensitivity-induced bias: (a) when sensitive questions generate more variable or inconsistent response patterns (e.g., due to social-desirability pressures), the item dispersion parameter increases, implying weaker measurement information and effectively reducing that item’s influence on the latent estimate; (b) when sensitivity produces systematic shifts that differ across national contexts for particular items (e.g., corruption questions being more taboo in some countries), the country-specific item bias terms capture these context-dependent deviations rather than forcing the latent trait to explain them. Of course, we also acknowledge that these mechanisms mitigate, rather than fully eliminate, sensitivity-related error.

1.2. On the topic of the country-specific item bias parameters, δ_{kq} , R1 raises the concern that “in practice, such a parameter can end up absorbing almost everything.” The reviewer is correct that these terms absorb any distinctive characteristic of a particular question in a particular country, not only translation issues or cultural idiosyncrasies but also features of the way a survey was implemented there such as “differences in survey mode, [or] sample representativeness.” We worked this point into our discussion of these parameters in Appendix A2.1. These terms do not, however, absorb “genuine cross-national variation in discontent,” the problematic possibility that prompted R1’s preoccupation. Each of the terms is estimated for a particular

pair of country and item, and reflects how responses to that item are different *relative to other items* in that country as opposed to other countries. That is, when item q repeatedly tends to contradict what other items tell us about discontent in country k but performs well elsewhere, δ_{kq} will be larger.¹ (This explains why δ_{kq} is set to zero when item q is asked only once in country k , as we need to see how it performs relative to other items repeatedly, and when item q is asked *only* in country k , because we need to see how it performs in country k relative to other countries.) We take care to elucidate this point in Appendix 2.1.

1.3. Finally, R1 also raised the issue of question choice, noting that we omit measures of concepts like satisfaction with democracy that are often used in studies of legitimacy and discontent. In response to this issue, first we make sure to elaborate the theoretical grounds for our selection of survey items as including measures of *diffuse* rather than *specific* support, building on such classic contributions as Easton (1965) and more recent empirical works such as Jennings et al. (2017).

And second, rather than compiling a comprehensive collection of such questions and generating an alternative PPD measure that incorporates them, we address this concern through external convergent validation. Specifically, items that were purposively excluded from the construction of the PPD measure for theoretical reasons—such as trust in the incumbent government, abstract support for democracy, and evaluations of democratic performance, drawn from the ISSP, WVS/EVS, and CSES—are instead used as external validation indicators. As shown in Figure 3, these measures exhibit strong and substantively meaningful correlations with the constructed PPD series, in line with theoretical expectations. These validation analyses were initially presented in the Appendix and, with substantial revisions, have now been incorporated into the main text in the measurement and validation section (see below). We believe this approach is sufficient to demonstrate that our measures are well founded and to give readers greater confidence that the results are not artifacts of particular measurement choices.

2.1. Reviewer 2 questioned whether the piece fits the EPSR’s ‘original article’ format, seeing it instead as a data paper and encouraging us to refocus our contribution in this direction. We thank the reviewer for this thoughtful and constructive comment. We agree that the manuscript, in its original form, may have appeared to sit uncomfortably between a theory-driven article and a data-focused contribution. In response, we have revised the manuscript throughout to more clearly position it as a data-focused paper whose primary contribution is the introduction of a novel time-series cross-national dataset on public political discontent, covering a substantially broader set of countries and years than existing measures. This new dataset allows, for example, a far more comprehensive examination of the extent and determinants of political discontent than has been possible in prior research. In particular, focusing on OECD countries—where previous studies have produced conflicting findings due to reliance on different datasets or survey items—our new data show that political discontent has increased over time and that this rise is primarily driven by worsening economic conditions, while institutional factors play a more limited role. This analysis clearly demonstrates how our novel data can contribute to existing literature on political discontent.

At the same time, our contribution is not limited to the creation of a novel dataset. As discussed in the manuscript, prior research on political discontent has often relied on divergent conceptualizations, resulting in heterogeneous measures and conflicting empirical conclusions. To address this limitation, we advance a clearer conceptualization of political discontent as a lack of diffuse support for the political system, drawing on Easton’s distinction between diffuse and specific support. This conceptualization integrates multiple dimensions of system-level evaluations emphasized in prior studies and provides a coherent foundation for comparative measurement. Taken together, the main contribution of our paper lies in advancing a clearer concept of political discontent and providing a novel dataset grounded in that concept.

To clarify these contributions, we have substantially revised the introduction, conceptualization, and conclusion sections to more explicitly discuss how our conceptual framework and the resulting dataset contribute to the literature. In addition, in response to the reviewer’s suggestion to more fully embrace the data-driven nature of the paper, we have expanded the discussion of our measurement strategy, including how the estimates are generated, how uncertainty should be interpreted, and how researchers can incorporate this uncertainty in subsequent analyses. Finally, we strengthen the conclusion by outlining how the PPD dataset can be used in

¹If item q often contradicts what other items tell us about discontent across *all* countries, this noisiness will drive its dispersion parameter α_q will be larger. We discuss this in Appendix 2.1 also.

future research on democratic backsliding, regime stability, and related topics. We sincerely appreciate the reviewer's suggestions in this regard; they have helped us clarify the core purpose of the paper and strengthen its identity.

2.2. R2 also raised the issue of the role of measurement validation in the paper. We are grateful for these thoughtful and careful comments on the validation part, which have helped us improve both the clarity and the prominence of the validation strategy in the revised manuscript. We fully agree that validation is central to the contribution of the paper, and we have always intended to treat it as such. In the original version of the manuscript, the full set of validation analyses was presented in detail in the appendix, with only a brief summary provided in the main text. In response to the reviewer's comment, we have revised the paper to move the validation analyses into the main text and to present them more prominently and transparently.

At the outset, we would like to clarify the overall structure of the paper. All validation of the PPD measure is conducted exclusively in the section titled "Validating Public Political Discontent" and consists of three distinct and standard validation exercises—internal convergent validation, external convergent validation, and construct validation—presented in Figures 2, 3, and 4, respectively. These three figures together constitute the full set of validation tests for the measure. In the revised manuscript, we now state this structure explicitly in the main text and more clearly signal the purpose of each validation exercise.

With respect to the external convergent validation test, we emphasize that this test compares PPD scores to evaluations of democratic performance drawn from the ISSP, WVS/EVS, and CSES, which were deliberately excluded from the latent-variable estimation. Because these variables are entirely external to the measurement model, they provide strong benchmarks for assessing whether the PPD measure behaves as theoretically expected. Consistent with the reviewer's point about validation, these relationships are presented as bivariate associations without covariates or adjustments, rather than as attempts at causal inference. We have revised the discussion of Figure 3 to make this role unambiguous.

We would also like to clarify that all validation tests are analytically separate from the analyses presented in the section "Explaining Political Dissatisfaction." Any explanatory variables introduced in that later section—shown in Figure 3 in the originally submitted manuscript and now presented in Figure 6 in the revised version—are not used for validation purposes. The "Explaining Political Dissatisfaction" section instead serves as an illustrative application of the newly constructed dataset, following a common practice in data-oriented contributions to demonstrate substantive utility after validation has been established. Among other things, this section illustrates how the uncertainty in the PPD estimates can be incorporated into analyses, which as we noted above is a point R2 helpfully reminded us to address. Variables such as election years are therefore introduced solely in this application context and are not intended to function as validation benchmarks. In response to the reviewer's comment, we have revised the language and structure of the manuscript to more clearly distinguish the validation section from the subsequent substantive analysis.

Finally, in line with the reviewer's suggestion—and in full agreement that validation should be prominent—we have moved key validation analyses from the appendix into the main text. The appendix continues to provide detailed supporting information—including model specifications, survey-item documentation, and full summary and numeric results—to ensure transparency and replicability. We hope that these revisions address the reviewer's concern by foregrounding the validation strategy and maintaining a clear separation between validation and the subsequent illustrative analysis.

References

- Easton, David. 1965. *A Systems Analysis of Political Life*. John Wiley & Sons.
Jennings, Will, Nick Clarke, Jonathan Moss, and Gerry Stoker. 2017. "The Decline in Diffuse Support for National Politics: The Long View on Political Discontent in Britain." *Public Opinion Quarterly* 81 (3): 748–58.