

# Economic Inequality and Belief in Meritocracy in the United States

## Memo to Reviewers

July 12, 2016

Thank you for the opportunity to revise and resubmit our manuscript, “Economic Inequality and Belief in Meritocracy in the United States.” Based on your helpful comments and constructive suggestions for improvement, we have made substantial revisions throughout the piece. As [Andrew Gelman wrote just last week](#), “there’s an unfortunate ‘incumbency advantage’ by which published papers with ‘p less than .05’ are taken as true unless a large effort is amassed to take them down.” With your assistance, we are confident we have now reached that heightened standard, and we hope you will agree that the work is suitable for publication in Research and Politics.

Specifically, the manuscript now:

1. **Provides a Close Replication.** A major concern raised by both reviewers was with regard to the multiple changes to the NJL research design made in the replication we presented: (1) the use of a single, consistent measure for the dependent variable in a much larger survey that was designed to be representative at small geographic scales, (2) the use of commuting zones rather than counties, and (3) the omission of causally-downstream attitudinal control variables. In hindsight, we can see how the simultaneous introduction of so many moving parts naturally raised the question of which is responsible for the difference from NJL’s null results, and how the scale of the local context raised particular concern. Although the CZ may well be a better contextual unit, as Reviewer 1 noted sympathetically, it does require us to use inequality data from about a decade earlier, raising the possibility that county-specific changes in the interim could be important. Reviewer 2 further pointed to the literature on racial context as an example of a body of work where different geographic scales have been implicated in divergent findings, and was generally skeptical of the superiority of commuting zones as a unit of local context. Both advised employing counties.

We therefore redid our replication using counties as the contextual unit, which also allowed us to use the same ACS income inequality data as employed in NJL. As Reviewer 2 also questioned our omission of the attitudinal controls, and because this

decision made little difference to the results either way, we now include those variables as well. This usefully allows us to focus on the one difference that does matter: the replication employs a single, consistent measure for the dependent variable and more representative individual-level data than those used in NJL.

This close replication forms the core of our revised manuscript. Reviewer 2 had suggested that, in light of the multiple differences between our original replication and the NJL analysis, we instead frame our work as a separate inquiry on the same topic rather than a reproduction and replication. This was an option we considered at length, but in the end we concluded that providing a close replication was the better of these two mutually exclusive paths.

We relegated the analysis using commuting zones as the unit of context to Appendix B, where they join the further analyses of the potentially confounding effects of relative mobility and income segregation that we presented in the original draft.

2. **Tests the Three NJL Measures.** Reviewer 2 expressed concern that we overstated our claim that the three different measures of the dependent variable employed by NJL should not have been pooled together. Despite showing substantial differences in the levels and trends of the three measures, she or he pointed out, we had not demonstrated that the three measures actually lead to different results. This is a good point. We therefore used the same surveys employed in NJL to perform separate analyses of each of the three measures of the dependent variable. These analyses confirm that the results obtained differ across the three measures in ways that strongly suggest that the results presented in NJL for several important predictors are an artifact of pooling the three different measures together. They also confirm that our primary critique is well grounded. There is no support for the conflict theory in the data used in NJL when the interaction term is correctly interpreted, regardless of how rejection of meritocracy is measured. We point interested readers to these analyses, presented in full as Appendix A, in footnote 5 on page 8.
3. **Avoids a Snarky Tone.** Reviewer 1 suggested that the “tone of the article can be somewhat less snarky.” We took this to heart: we agree that the article is best served by simply carefully describing each of these peculiarities in the empirics of NJL without editorializing language like ‘strangely’ or ‘oddly.’ We revised throughout to avoid any such judgmental language. We were also sure to revise the paragraph at page 4, flagged in particular by this reviewer, to make more clear that we are reporting Huber and Stephens’ (2012) characterization of the conflict theory as based on an “implausible premise”—the passage we paraphrase is “With only a little reflection, the basic premise of Meltzer-Richard [(1981), the [tremendously influential](#) early [formalization of the conflict theory](#)] is implausible. It asserts a major feature of social structure, the very system of stratification of society, is self-negating” (this passage can be read in context [on Google Books here](#))—and added a sentence to indicate that regardless of their plausibility, the important thing is to find out which theory is supported by the evidence.

4. **Elaborates on NJL’s Error in Interpretation.** Reviewer 2 suggested that, because our reproduction shows that the estimated coefficient of income inequality for those with incomes below \$10,000 is positive and approaches statistical significance, NJL’s misinterpretation of the interaction term is perhaps not really that consequential. Although this view overlooks the fact there is no hint of support for the article’s claim that its results support the conflict theory’s prediction that *higher*-income individuals will be *less* likely to reject meritocracy in contexts of greater income inequality, it is nevertheless one that some readers of our piece may share.

We certainly agree that the difference between  $p$ -values just below and just above 0.05 should not be overstated. As Gelman and Stern (2006, 328) famously wrote, “the difference between ‘significant’ and ‘not significant’ is not itself statistically significant.” But even if the coefficient estimate for the very lowest-income Americans were statistically significant, it would provide only minimal support for the conflict theory.

We again direct attention to the substantial literature in political methodology regarding the correct use and interpretation of interaction terms. Berry, Golder, and Milton (2012) addresses this issue directly. In discussing the circumstance in which the estimated conditional effect of  $X$  is statistically significant only for a range of values of the conditioning variable  $Z$ , Berry, Golder, and Milton (2012, 660-661) points out that the extent of support for the hypothesis that  $X$  has a positive effect “depends on the percentage of observations having values of  $Z$  at which the marginal effect of  $X$  is positive and significant . . . The higher this percentage, the more inclined we would be to accept the empirical evidence as supportive.” In this case only 4% of the sample has incomes below \$10,000. The support for the conflict theory, even conceding the question of statistical significance, is minimal. We include a discussion of this point in footnote 3 on page 7.

Relatedly, Reviewer 1 inquired about NJL’s predicted probability plots—“Did the authors of this article look at the NJL code and see how they arrived at these figures?”—and Reviewer 2 also noted these diverging predicted probabilities. Unfortunately, the NJL reproducibility materials do not provide the code for producing the article’s figures; in fact, these materials consist of little more than the analysis data and a single line of code for each model in the article. Despite considerable effort, we were unable to approximate the predicted probabilities depicted in NJL using the reproduced results. Predicted probabilities derived from logistic regression models depend not only on the values of the variables of interest but also on the values assumed for the other variables in the model, and NJL does not indicate what values were adopted. Neither mean nor median values—the most easily justified and commonly assumed values—for the other predictors yield predicted probabilities similar to those shown in NJL’s Figure 2.

5. **Expands the Conclusion.** Presenting a close replication had the happy consequence of allowing us to dispense with the discussion of commuting zones, freeing space below the RaP word limit for us to incorporate Reviewer 1’s suggestion that we expand the paper’s conclusion to include a discussion of the implications of our findings. These

implications are indeed important. If the conflict theory were correct, advocates for redistribution would need only wait for inequality to activate conflict and so deliver the votes needed for the policies they prefer. That the relative power theory better describes political reality, however, means that the social structure will not simply undermine itself. Instead, the hard work of organization and mobilization will be crucial.

6. **Labels Figure 5.** Reviewer 1 wrote that she or he could not see the labels on the axes and lines on Figure 5 and so needed to refer to the text to understand what was being depicted. This is a concern: we completely agree that all such plots should be easily understood at a glance. We verified that these axis and line labels appear in our revised submission, but as the labels also appear in our proof of the original submission we received from Manuscript Central, this is not completely reassuring. We will be sure to confirm with the RaP editors that the figures appear as intended should we be given the opportunity to publish the work.

## References

- Berry, William D., Matt Golder, and Daniel Milton. 2012. "Improving Tests of Theories Positing Interaction." *Journal of Politics* 74(3):653–671.
- Gelman, Andrew, and Hal Stern. 2006. "The Difference Between 'Significant' and 'Not Significant' Is Not Itself Statistically Significant." *American Statistician* 60(4):328–331.
- Huber, Evelyne, and John D. Stephens. 2012. *Democracy and the Left: Social Policy and Inequality in Latin America*. Chicago: University of Chicago Press.
- Meltzer, Allan H., and Scott F. Richard. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy* 89(5):914–927.