

A Reply to “Reducing Political Bias in Political Science Estimates”

Daniel Maliniak, *College of William & Mary*

Ryan Powers, *University of Wisconsin–Madison*

Barbara Walter, *University of California, San Diego*

ABSTRACT

Zigerell (this issue) cites the findings of his recent reanalysis (Zigerell 2015) of the data in our 2013 study of the gender citation gap in the international relations literature to support his claim that our study showed a “preference for statistically-significant results.” We thank Zigerell for so closely engaging with our work. However, we note that he is focused on how his changes to our sample affect a single model in our original paper, highlight the fact that we reported statistically insignificant results when they arose in our original analyses, and review the findings of other recent re-analyses of our data. Ultimately, while we disagree with Zigerell’s conclusions about our work, we join Zigerell in calling for greater diversity in the discipline.

Zigerell (this issue) is concerned that “reporting flexibility” allows researchers to make post hoc changes to research designs in order to find “support for the policy preferences of a particular political party or ideology.”¹ To justify his concerns, Zigerell cites the findings of his recent reanalysis (Zigerell 2015) of the data used in our 2013 article on the role of gender in citation patterns in the international relations literature (Maliniak et al. 2013) and then asserts that our study showed “a preference for statistically-significant results.” However, Zigerell focuses on how his reanalysis affects a single model from our original article and ignores the fact that we reported statistically insignificant results when they arose in our original article. We report below how Zigerell’s (2015) changes to our sample have little effect on the substantive implications of most of the models that we reported in our article.

Our study demonstrated that articles written by women, and untended women in particular, are cited fewer times and are less central in the IR literature citation network compared to similar articles written by men alone or coauthored with men. Our sample included roughly half of the IR articles published in 12 top peer-reviewed journals over the period 1980–2006. Zigerell (2015) notes that our dataset, but not our published analysis, included data on articles published in 2007 and reports on how the inclusion of those observations affects a single model from our article. We thank Zigerell for so closely reviewing our work and highlighting our omission of the 2007 observations. This omission

resulted from a miscommunication between coauthors and we take full responsibility for it. After adding the observations from 2007 to our “Kitchen Sink” citation count model, the p-value on the variable of interest, *All female*, inflated from .093 to .223. Despite this, the substantive implications of nearly every other specification of the citation count models in our original article are similar after including observations from 2007 (table 1). The coefficient on *All Female* remains negative across all models, is of similar magnitude to those in our original analyses, and is statistically significant across all but the most saturated models (the last two).² Given his present concern with reporting flexibility, we find it odd that Zigerell does not make this clear to readers in either the present essay or in his 2015 sensitivity analysis.

Additionally, we were not shy about reporting statistically insignificant results when they arose in our original analysis. We reported a statistically insignificant effect of author gender on citation counts in our decade sub-sample analysis of the 1990s and 2000s.³ Similarly, we reported that the gender effect appears to be smaller for more recently published articles and is not statistically detectable for articles less than a decade old at the time of our data collection. Finally, the p-value of .093 on the *All Female* coefficient in the “Kitchen Sink” model that we reported in the original article is not, by some standards, statistically significant. We think it is clear from these examples that we were not selectively reporting only statistically significant results.

We are encouraged to see that Zigerell’s (2015) effort is joined by others who continue to study gender bias in citations. For example, Fogerty (2015) argues that the distribution of our citation count variable is more consistent with distributional assumptions of a Poisson inverse Gaussian (PIG) regression model than those of a negative binomial (NB) regression model. Fogerty re-estimates the “Kitchen Sink” citation count model and includes the

Daniel Maliniak is assistant professor of government at the College of William & Mary. He can be reached at dxmali@wm.edu.

Ryan Powers is a PhD candidate in the Department of Political Science at the University of Wisconsin–Madison. He can be reached at rpowers@wisc.edu.

Barbara Walter is professor of political science in the School of Global Policy & Strategy at the University of California, San Diego.

2007 observations. The coefficient is negative and statistically significant, but estimated effect of gender is substantially smaller in magnitude. While we view Fogerty's reanalysis of the models that we reported in our original paper as broadly consistent with our findings, his method does produce less dramatic and more mixed results than ours. Separately, Roberts, Stewart, and Nielsen (2016) reanalyze our data (including articles from 2007) using their newly-developed matching technique for high-dimensional data, like the text of journal articles, to balance the sample and find results

In our view, the best response to such a pattern of results is to develop more robust theories of how gender bias might manifest in academia and then test those theories using of variety of theoretically-motivated research designs or newly collected data.

very consistent with those in our original study: articles written by women are cited less than articles written by men or coauthored by men and women. In our view, the best response to such a pattern of results is to develop more robust theories of how gender bias might manifest in academia and then test those theories using of variety of theoretically-motivated research designs or newly collected data.

We conclude with a general comment. Given Zigerell's particular concerns about ideological bias, we expected him to propose mechanisms uniquely suited to addressing such biases. He does not do this. Instead, he calls for preregistration of research designs and pre-acceptance of journal articles. We are not opposed to either of these efforts, but neither of them addresses the underlying issue of ideological diversity. Indeed, if a lack of ideological diversity biases the questions we ask and conclusions we reach in political science, it is likely that a lack of diversity on other dimensions that affect how individuals interpret the world around them are problematic as well. One tractable solution is to strengthen departmental diversity policies to help ensure recruitment of faculty and students that are more diverse in terms of age, physical ability, ethnicity, gender, race, sexual orientation, socioeconomic status, and potentially, political ideology.⁴ As such, we view Zigerell's concerns as a subset of broader concerns about the lack of diversity in our discipline. In that light, we join the author in calling for greater diversity in political science. ■

NOTES

1. Zigerell motivates his concerns about reporting flexibility with references to LaCour and Green (2014) and reproduction failures in psychology reported by the Open Science Collaboration (OSC) (2015). We find this puzzling since neither has much to do with "post-hoc research design choices." Further, existing reporting and data sharing norms allowed for problems with the analyses to be discovered. For example, LaCour and Green was retracted because Broockman, Kalla, and Aronow (2015) found evidence of fraud in the publically-available replication data. More recently, Gilbert et al. (2016), studied OSC's work and concluded that it "contains three major statistical errors and, when corrected, provides no evidence of a replication crisis. ... [OSC's] data are consistent with the opposite
2. We estimated these models by modifying the original replication code that we posted on the *International Organization* website.
3. The same analyses of the citation network did find a statistically significant effect for author gender. Articles written by all women are less central in the IR citation network than those written by men or coauthored with men.
4. Zigerell (this issue) cites Duarte et al. 2015 in his discussion of the benefits of ideological diversity. The authors of that study explicitly call for diversity statements to include political ideology. See also Shields and Dunn (2016) who outline a number of other potential ways to increase ideological diversity in the social sciences and humanities.

REFERENCES

- Broockman, David, Joshua Kalla, and Peter Aronow. 2014. "Irregularities in LaCour (2014)." (2015): A36–A38.
- Duarte, José L., Jarret T. Crawford, Charlotta Stern, Jonathan Haidt, Lee Jussim, and Phillip E. Tetlock. 2015. "Political Diversity Will Improve Social Psychological Science." *Behavioral and Brain Sciences* 38: e130.
- Fogarty, Brian J. 2015. "Comment on Zigerell (2015): Using Poisson Inverse Gaussian Regression on Citation Data." *Research & Politics* 2 (4): 2053168015617496.
- Gilbert, Daniel T., Gary King, Stephen Pettigrew, Timothy D. Wilson. 2016. "Comment on 'Estimating the Reproducibility of Psychological Science.'" *Science* 351.6277: 1037–1037.
- LaCour, Michael J. and Donald P. Green. 2014. "When Contact Changes Minds: An Experiment on Transmission of Support for Gay Equality." *Science* 346 (6215): 1366–1369.
- Maliniak, Daniel, Ryan Powers, and Barbara F. Walter. 2013. "The Gender Citation Gap In International Relations." *International Organization* 67 (4): 889–922.
- Monogan, James E. 2015. "Research Preregistration in Political Science: The Case, Counterarguments, and a Response to Critiques." *PS: Political Science & Politics* 48 (03): 425–429.
- Open Science Collaboration. 2015. "Estimating the Reproducibility of Psychological Science." *Science* 349 (6251): aac4716.
- Roberts, Margaret E., Brandon M. Stewart, and Richard Nielsen. 2016. "Matching Methods for High-Dimensional Data with Applications to Text." Working Paper, (<http://www.margaretroberts.net/wp-content/uploads/2015/07/textmatching.pdf>).
- Shields, Jon A. and Joshua M. Dunn Sr. 2016. *Passing on the Right: Conservative Professors in the Progressive University*. Oxford University Press.
- Zigerell, L. J. 2015. "Is the Gender Citation Gap in International Relations Driven by Elite Papers?" *Research & Politics* 2 (2): 2053168015585192.

Table 1

Estimates from Maliniak et al. (2013) with 2007 observations. Model titles represent the addition of different sets of article or author attributes

	Gender	Article Age	Career	Epistemology	Material
All Female	-.199**	-.199**	-.195**	-.198**	-.197**
	(.0814)	(.0814)	(.0958)	(.0935)	(.0934)
AIC	21998	21998	21881	21826	21828
	Ideational	Paradigm	Issue	Method	Kitchen Sink
All Female	-.211**	-.19**	-.158*	-.14	-.104
	(.0937)	(.0941)	(.0914)	(.0892)	(.0853)
AIC	21810	21715	21649	21579	21159

Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.