

**MS. NO. PSRM-OA-2019-0126.R1 ENTITLED “TAKING DYADS
SERIOUSLY”**

Dear Editors,

Thank you for the opportunity to revise and resubmit our manuscript once more. We believe the manuscript has benefited from the Reviewers’ helpful and thoughtful comments. We have revised the manuscript, taking seriously the points raised by you and the Reviewers. Our comments and responses are shown in *BLUE* below each of the points from your letter.

We hope you agree that the manuscript has improved through this process and we are looking forward to your response.

EDITOR

- (1) ‘effect’ is not precisely defined, so you might substitute “how to estimate regression coefficients in a GLM context”
 - *We thank you for this suggestion and have changed the manuscript accordingly.*
- (2) Also, my understanding is that modeling the “underlying structural features in the data” means including these “structural features” in the conditional mean equation of the GLM rather than including them exclusively in the variance equation (or ignoring them altogether). Or, more precisely, including SRRM and LFM parameters in the conditional mean equation. The advances you make come from this specific modeling choice. Augmenting the GLM in this particular way leads to 1) unbiased (or nearly unbiased) estimates of regression coefficients under homopholic dependence, when these coefficients exist and are known (the Monte Carlo simulations) and 2) superior out-of-sample forecast predictions for real data, in which case the regression model is just an approximation with unknown parameters (the replication exercises). Including LFM parameters in the conditional mean equation also makes it more likely that IR scholars will treat dyadic dependence as substantively important.
 - *Thank you for these suggestions. We have made significant modifications to the introduction in line with the recommendations that you have posted and those from reviewer 1. We have removed references that reviewer 1 perceived as swipes at previous work. In addition, we have adopted your helpful language about how the AME approach essentially calls for including structural features in the conditional mean equation of the GLM.*

REVIEWER 1

- (1) In my previous review I pointed out some problems with vague language, mostly in the opening sections. Honestly, I still find the introduction to be problematic. Maybe it could be cut and we just open with the part of section 2 where the paper begins to discuss processes that lead to dyadic and higher order dependencies? The reason is that the intro tries to strike a very general tone. But in reality, the paper proposes a specific estimation strategy for parameters in a specific parametric model of networked interactions. I think the paper would be more useful to readers if it started with the specific applications, and then on the basis of that, motivated the statistical specification on substantive grounds. There are some swipes at other literature are neither necessary nor clear enough to be very compelling.

- *We thank the reviewer for taking the time to provide their comments and helping us to make the manuscript more focused. We have made a number of changes to the framing of the article per the reviewer's suggestions. Additionally, we have significantly modified the introduction and removed what could be considered as, but certainly not intended, swipes at other literatures. The editor also suggested a number of changes to be made in the introduction and we have adopted those as well. We disagree with the choice of removing the introduction completely and starting with section 2 because we feel some relatively non-technical context will be useful to more general readers.*

- (2) “Most of these studies use a generalized linear model (GLM). However, this approach to studying dyadic data increases the chance of faulty inferences by assuming data are conditionally independent and identically distributed (iid).” It should really read, “Most of these studies use a generalized linear model (GLM). However, this approach to studying dyadic data increases the chance of faulty inferences if the estimation and inference assume data are conditionally independent and identically distributed (iid).” But then if this is all just about mistakenly assuming iid, just say it quickly and move on.

- *We thank R1 for this suggestion, per R1's first comment, we have actually removed this from the introduction and clarified the goal of the paper along the lines of what R1 and the editor suggested.*

- (3) Why would the mean estimate be biased downward? And, biased with respect to what target of inference? A problem with the opening discussion is that it does not establish a target of inference. Without doing so, “bias” is undefined. It seems like they mean the coefficient for the linear relationship between some regressor and the latent outcome variable. But what if one is using a nonparametric model or a semiparametric model that operates on a different latent space—is bias even defined? It seems that for bias to be defined we have to be restricting ourselves to a class of models that propose the same latent space, but that makes this much less general that the discussion leads on. Moreover, that paper is not committing to talking about bias relative to causal effects under effect heterogeneity, but isn't that usually what people are after? Or, maybe the paper is really concerned with bias with respect to a forecast prediction? If so, then we need to set up a forecast prediction error analysis.

- *These are fair questions especially in terms of the complications that arise when using nonparametric or semiparametric models. Instead of trying to resolve all*

this in the introduction, we have removed the portion of the introduction that led the reviewer to ask this question.

- (4) Moving to section 2, the extended discussion of specifying a likelihood under iid is overkill. Again, I would just move on to the discussion of the application. The point that iid is inappropriate will receive due attention in the course of developing the networked interaction model.
 - *We understand the position that a detailed discussion of specifying a likelihood under iid seems laborious and potentially unnecessary. After a lengthy discussion among our co-authors, we have decided to retain this section of the paper. We believe that since our approach is directly aimed at applied scholarship in International Relations, and because we focus on applications that have appeared in top journals during the last 10 years, this section enables our study to serve a broad audience of readers. By keeping it, we ensure that a reader can directly follow our main points about GLM specifications and how the AME addresses these concerns.*
- (5) Explain how inference is done for the parameters from the AME model. How are intervals produced exactly? (Need this because you examine these in the simulation and make many statements about intervals throughout the text.) What are the operating characteristics, in terms of convergence properties, etc.?
 - *Bayesian inference for the parameters from the AME model proceed via Gibbs sampling algorithms where we iteratively simulate values of the model parameters from their full conditional distributions. We have added additional discussion about how the parameters are estimated to both the simulation section and the applications. In Appendix A we provide a brief outline of the MCMC algorithm used. We do not re-derive the full conditional distributions for the AME models in this paper as it has been done in previous work that we have cited: Hoff 2005, 2008; Minhas et al. 2019. Additionally, Hoff (2021), which we now cite in the paper, more fully details the history of the AME model and how it has evolved since its original formulation in Hoff et al. (2002).*
- (6) The DGP is simple enough that the paper needs to include the analytical calculation of the bias from the simple probit regression of Y on X. We get no insight by just having monte carlo simulation results. For the simulations, clarify the “standard international relations approach” as just Y regressed on X, when you introduce it. Then explain how confidence intervals were produced. Are these just normal-approximation Wald intervals with non-robust standard errors? Explain exactly how the confidence intervals are produced for the AME model. They were never characterized in the section that discussion the model.
 - *Given the various forms of network dependence that can arise it’s not clear to us how to actually construct an analytical calculation like what R1 is suggesting. If the purpose of this analytical calculation would be to highlight the consequences that dyadic dependence would have in the context of inference, then that is exactly why we open Section 2 in the paper with a discussion of the assumptiong underlying GLMs.*
 - *In terms of the simulation section, we thank R1 for pointing out the parts that are still unclear and have added in their suggested language with respect to clarifying the standard IR approach.*

- *We thank R1 for asking a question about how confidence intervals are produced because that was an error on our part as we meant to say credible interval. Within our framework These are generated directly from the posterior distribution generated from the three models. Specifically, for every simulation and model we calculate the 95% credible interval and then across all simulations we estimate the proportion of times that the true value fell within those intervals. We have clarified this in the simulation section of the manuscript.*

REVIEWER 2

- (1) I would have liked to see a justification for the choice of papers used in the replications. It is still unclear. Are these cases where the theory fail to characterize the direct and indirect effects, and implicitly assume that the latent dependencies are nuisances?
- *In footnote 20 we state the primary criteria for our choice of papers used: we focus on studies explicitly about International Relations, were published after the year 2000, and were published in a top ranking general political science journal (for consistency in editorial standards and reviews, we focus on one journal, the American Political Science Review). In addition to these criteria, we searched for studies we understood to have hypotheses in which interdependencies are theoretically consequential due to the dyadic setting in which their theories operate (p.14) but where-in dependencies are not directly modeled (beyond the use of standard error corrections). We chose papers that, in spirit, seem to treat interdependence as technical problem rather than a threat to inference.*
- (2) The presentation of the AME model clearly states that the Xs are exogenous; it is likely that they are not in replications. Still, the AME estimates (which in the simulations do not seem to outperform the standard method for estimating the betas when correlation between X and W is high) suggest that the models in the original papers are misspecified, and that these differences are likely driven by a better modeling of the DGP. This is a very important take-home point.
- *Thank you. We hope that this take home point is made clear throughout the paper, but especially in sections 5.4 ‘lessons learned’ and in our conclusion. We have added the following change to section 5.4 “In each of the models the AME substantially outperforms the original model out of sample. This may be because by ignoring dependencies, the original models are misspecifying the DGP, and the AME better accounts for it.”*