# Response memo for JCR 20-0346

2020-09-25

Thanks for the chance to revise and improve our manuscript. We respond to each of the three major suggestions in JCR’s initial recommendations below. We also briefly review a few other changes we made to the manuscript.

## JCR suggestions

*1. On p.5, lines 3-4 (as numbered by JCR system) you state that standard practice is to use empirical vs. smoothed ROC curves. You need to provide some evidence to support this claim of standard practice. This turns out to be a critical point so you need to back this up to be more convincing.*

We conducted a literature survey of ROC/AUC-ROC use in regard to this question by checking all 60+ references in the B&S paper and also 15 articles from two recent JCR issues. We essentially found that only empirical ROC curves are used. There are no clear instances where parametric smoothing of ROC curves was used.

We established this by either visually checking ROC curves were empirical (i.e. they have clear discontinuities and are not smooth) and assuming that associated AUC values were similarly empirical, or, if there were only AUC values but no ROC curves shown, by checking replication code. There were two ambiguous cases where we could not clearly establish that parametric smoothed ROC curves were *not* used, but nor that parametric smoothed ROC curves were used. Both involve Blair as co-author. Details on these and also the instances where we checked replication code are documented at <https://github.com/andybega/Blair-Sambanis-replication/tree/master/journal-survey>.

*2. On p.14 in the opening paragraph of the conclusion there are a couple of points. First, the sentence, “It is not a procedure.” is unclear to me and should we re-written. Second, the last sentence in the opening paragraph referring to “unfortunate standard bearer...” does not seem to follow from previous sentence. The fact that many theoretical specifications might be derived from the literature does not imply that the B&S theoretical choice is unfortunate per se. Given the large literature, there could be many “strong” theoretical claims RE civil war onset and escalation. State more clearly what your concern is here.*

We have included a clearer definition of what we mean by “strong” versus “weak” theory and expounded on the dangers of variable selection in applications of weaker forms of theory. We try to explicate exactly your point: many different implementations of a broad literature on process in conflict dynamics are possible.

*3. On p.15, the last paragraph is an opportunity to open up your discussion beyond the very specific critiques you have raised about B&S and try to engage the JCR readership on the broader debate about approaches to forecasting in IR and conflict studies. My biggest concern with your MS is that many readers will not stay engaged with it as you walk them through a series of important but very specific and sometimes quite technical points. In the conclusion I would like to see more discussion of the larger issues on theory driven vs. machine learning approaches to forecasting, or other issues related to improving and advancing forecasting among conflict scholars. I know you have a clear position on these issues and this is the place to discuss them. It is imperative that you do this in the conclusion. I would also suggest that in the introduction you reference these larger issues as something readers can look forward to in the conclusion.*

We have expanded the concluding discussion regarding the use of theory in forecasting and how different aspects of machine learning relate to this in accordance with your comment. In particular, we try to point out that a horse race may not be the only way to proceed. And, a single horse race offers only a single case study, not a generalizable law. It turns out that when analyzed carefully and correctly, the preferred model of B&S is the weakest of the various alternatives examined. This doesn’t mean that ML models are better, but it certainly suggests that they may be useful ways to learn about conflict dynamics.

*4. The title may need to change a bit. I will consult with Sage editors on best practices. This can be handled a bit later if necessary.*

We have changed the title for now to “Reassessing the rôle of theory and machine learning in forecasting civil conflict”, but will otherwise wait for better suggestions.

## Other changes

The previous version of the paper contained low-level discussions referencing specific lines of code and variable names in two footnotes and portions of the main text. In the spirit of comment #3, we have moved these to the appendix. We have also changed our discussion of onsets versus incidents to reflect that this seems to be a design decision by B&S rather than a coding error.

Finally, we have each thoroughly copy edited the manuscript.