

Reviewer Comments on MS 2092-1:

“A Replication of Willingness-to-pay Estimates in ‘An Adding Up Test on Contingent Valuations of River and Lake Quality’ (Land Economics, 2015)”

Economics E-Journal

December 15, 2017

This is a brief, but thorough and insightful paper. As a profession we should be appreciative of replication efforts like this, especially for methods as important and controversial as contingent valuation. Most of my comments are minor, and are discussed below. But I do have two more substantive suggestions/questions.

First, it is unclear how the statistical comparisons are performed when testing whether the sum of the WTP for each incremental part of the improvement equals the WTP for the total improvement as a whole. In some cases, failure to reject the null hypothesis that the difference of the WTP estimates is zero is discussed, and in other cases it sounds like the author is simply looking at whether the confidence intervals overlap. I suggest providing further details on the statistical test used, and include the test statistics and p-values. This is particularly important for three reasons: (i) the point estimates from the parametric models demonstrate economically significant differences, (ii) the confidence intervals are fairly wide, and (iii) the conclusions based on the upper and lower end of the confidence intervals are fairly close. Poe et al. (2005) discuss alternative methods to statistically compare differences between WTP estimates, including bootstrapping (as done in the original 2015 study by Desvousges, Matthews, and Train (DMT)) and the methods of convolutions, both of which involve simulating a distribution of the differences between the WTP estimates. Poe et al. find that drawing statistical inference based solely on whether the confidence intervals overlap is inappropriate. All this said, since the estimates for each incremental part are estimated within the same logit model, perhaps a more conventional test (e.g., Wald test) was used?

This brings me to my second question. Why were the parameters in Table 4 estimated using the same regression when each scenario could have been estimated as a separate logit model? The constant term and slope coefficient are identified based on mutually exclusive subsets of the data, correct? Is there any advantage or disadvantage to pooling the data versus estimating separate models?

More minor comments:

1. A bit more context could be provided in the first paragraph or two of the paper. In particular, a diagram of the four incremental improvements and the whole would be very useful in getting the reader up to speed. Perhaps you can even just borrow Figure 1 from the 2015 DMT paper?
2. Although this may be well-understood by most readers, I think a brief sentence in the third paragraph about why the ABERS and Turnbull estimators require smoothing would be appropriate. As I understand it, this is simply because a valid CDF must be monotonically non-decreasing. This explanation could also fit in the paragraph directly preceding table 2.

3. Table 1 would be more legible if there were vertical lines, additional space, and/or alternating background color shades denoting that each set of three columns correspond to a different increment.
4. In the conclusion the author refers to DMT's "relatively inexpensive, small non-probability sample". Given the back and forth discussions posted on 9/29/2017 regarding the earlier working paper, it seems that there is some additional information available regarding the internet panel and selection process for respondents. An additional few sentences on these details would be enlightening, particularly if it's believed that the respondents in this particular panel are not fully considering the tradeoffs in the valuation questions. The distinction between methodological shortcomings versus implementation and data quality is crucial.

Works Cited

Poe, Gregory, Kelly L. Giraud, and John Loomis. 2005. Computational Methods for Measuring the Difference of Empirical Distributions. *American Journal of Agricultural Economics* 87(2): 353-365.