Dear Professor Vuolo,

I would like to thank Sociological Methodology for conditionally accepting my paper for publication. I have incorporated Paul von Hippel’s recommended changes into this most recent draft. In response to the request that I share my data and code, I have created a Github repository where the public data and all the code needed to recreate the tables and plots based on this data can be found. This repository provides instructions for reproducing these results. The links to the Lorenz interpolation R package and the repository for reproducing these results are in a footnote at the end of the abstract.

I have also revised the parts of the paper that cover the theory of why Lorenz interpolation provides more accurate estimates than MCIB or CDF interpolation. In the previous draft, I suggested that Lorenz interpolation produces less positively biased estimates of the closed bin means, resulting in a better estimate of the top bin mean. Paul von Hippel disputed this point in his response letter. I have removed this claim from the new draft. Instead, I simply state that Lorenz interpolation produces better estimates of the income distribution upper tail. My evidence for this is the lower relative bias and RMSE of Lorenz interpolation’s top bin mean estimates (Table 2 of the paper). I have also added an endnote in which I say that the theoretical advantage of Lorenz interpolation is not entirely clear. I credit Paul von Hippel for suggesting that Lorenz interpolation’s advantage may come from the Lorenz curve being a smoother function than the CDF and better approximated by a cubic spline. I also suggest that the improvement of Lorenz interpolation may come from the rules according to which the cubic spline produced by Lorenz interpolation is defined.

Another significant change in the most recent draft is that I have revised the section in which I derive the PDF of the income distribution produced by Lorenz interpolation. Equation 10 in the previous draft had an incorrect subscript. I fixed this. I have also specified which variables the derivatives in this section are taken with respect to. Finally, I have added a comment on why the PDF at the end of the derivation has a few spikes in the upper tail. These spikes result from the PDF implied by Lorenz interpolation, which consists of a set of square root functions. When these functions underestimate the bin means, the resulting PDF has spikes at the bin lower bounds. Lorenz interpolation still produces accurate estimates of bin means, income shares, and quantiles despite the presence of these spikes.

I also have made significant revisions to the section in which I explain how Lorenz interpolation works (the section entitled “Building a Lorenz Curve from Grouped Income Data”). Paul von Hippel found this section to be unnecessarily complicated. I have reduced the length of the section and removed some of the notation. I also moved the paragraph in which I show how a system of linear equations can be used to compute the coefficients of the estimated Lorenz curve into the endnotes.

Paul von Hippel disagreed with my description at the beginning of the background section of how the midpoint-Pareto estimator produced biased inequality estimates. He suggested that I either cut the part of the paper that talks about midpoint estimators or replace this section with a paragraph that he wrote in his response letter. I have kept the discussion of midpoint estimators in the paper and have replaced the passage in question with a couple of sentences that convey Paul von Hippel’s main point that ignoring the variation within the brackets may lead to positively or negatively biased inequality estimators depending on the shape of the underlying income distribution.

Those are all the big differences between this draft and the last one. Aside from these, I have addressed the other issues Paul von Hippel mentioned in his response letter. I have noted that the greater accuracy of Lorenz interpolation compared to CDF interpolation is at estimating income quantiles given the latter method’s reliance on the inverse of the quantile function. I have also changed the Table 4 metrics to relative bias and RMSE. Beyond these changes, I have tried to improve the quality of the writing throughout the paper.

Hopefully these changes are enough to get this paper across the finish line. Thanks again to the reviewers for their helpful comments and the editors for considering this paper for publication. Please let me know if you have any additional questions or revision requests.

Best,

Andrew