Demography (2020) 57:2377–2381
https://doi.org/10.1007/s13524-020-00936-x

COMMENTARY

Another Plea for Caution When Using Survey Income Data From the Far-Left Tail

James X. Sullivan [1]

Accepted: 30 October 2020 / Published online: 19 November 2020

Introduction

As an original reviewer of Brady and Parolin's article (published in this issue of *Demography*), I expressed concerns about measurement error given that the study relies on extreme outliers in the distribution of survey income. Recognizing the importance of this issue, the editorial team at *Demography* invited me to share these concerns in a comment. In my comment (published in this issue of *Demography*), I highlight evidence from a growing literature indicating that income is significantly underreported in household surveys. This underreporting is evident across many surveys in the United States and internationally, has worsened noticeably over time, and appears to be most problematic for values at the very bottom of the reported income distribution. Given this evidence, one should be very cautious about drawing strong conclusions based only on data from extremely low values of income reported in surveys. In their response to my comment, Brady and Parolin argue that the concerns I raise are not problematic for their analyses. They also present results from additional robustness exercises intended to address measurement error concerns and argue that these results confirm their conclusions.

In this response, I note that the evidence that Brady and Parolin present in their response does not address the key concerns in my comment. My goal here is to clarify the key arguments and findings from the relevant literature. Brady and Parolin offer additional analyses in their response, noting that their main conclusions do not change qualitatively after they make small changes to their approach. These robustness checks, however, do not address the primary issues raised in the literature. A major recent development in this literature is the launch of the Comprehensive Income Dataset (CID) by researchers internal and external to the U.S. Census Bureau. The CID links data from several national surveys—including the Current Population Survey (CPS), which Brady and Parolin use—to administrative data on many sources of income,

✉ James X. Sullivan
james.x.sullivan.197@nd.edu

[1] Department of Economics and the Wilson Sheehan Lab for Economic Opportunities (LEO), University of Notre Dame, Notre Dame, IN 46556, USA

including earnings, several cash and in-kind transfer programs, and retirement income (Medalia et al. 2019). A primary goal of the CID is to improve the accuracy of income data in large surveys. A key feature of the CID is that it can be used not only to confirm earlier studies finding that many income sources are significantly underreported at the bottom but also to correct survey-based estimates that rely heavily on these poorly measured sources. Making such corrections shows that Brady and Parolin's estimates are biased significantly upward. For $2/day poverty, the bias is more than 120%. In addition, if the underreporting of income is rising over time, as several studies have suggested, then Brady and Parolin's estimates of the trends in extreme poverty may be biased upward.

As I noted in my comment, Brady and Parolin's study is important because it brings greater attention to critically important questions about the extent of extreme deprivation in the United States and how this deprivation has changed over time. Furthermore, it is important to understand what we can learn about the levels and trends in extreme and deep poverty using survey data alone. It is also important, however, for such analyses to acknowledge the potential for significant bias due to measurement error and to specify the underlying assumptions about measurement error that are required to draw valid conclusions.

The Validity of Survey-Based Extreme Poverty Estimates

Many studies have documented the underreporting of income in surveys. New evidence from Meyer et al. (2019), which relies on the CID, is the most directly applicable here because it examines the robustness of estimates of extreme poverty using the CPS. In my comment, I emphasize that the evidence from Meyer et al. shows that underreported earnings are the most important source of bias for survey-based estimates of extreme poverty. In their study, Meyer et al. make two types of adjustments for underreported earnings: survey-based adjustments (for example, setting a lower bound on household earnings as reported annual hours worked times the federal minimum wage) and administrative data–based adjustments. I emphasize only the latter adjustments in my comment, because Meyer et al. showed that survey-based adjustments have little impact in the CPS given that very few households in this survey reported positive hours but extremely low earnings.

Despite Meyer et al.'s finding that survey-based adjustments to earnings have little impact in the CPS, Brady and Parolin's additional analyses related to earnings underreporting consider the impact of only survey-based adjustments. In figure 3 of their response, they present estimates that adjust earnings using reported hours, following the procedure in Meyer et al. These results show that such adjustments have virtually no impact on their estimates, which is exactly what one would expect given Meyer et al.'s point that few households in the CPS reported positive hours but extremely low earnings. This adjustment is inconsequential in the CPS.

Brady and Parolin characterize these new results in their figure 3 as confirming their conclusions. However, the much more substantive issue (and the one I point out in my comment) remains: *administrative*-based adjustments to earnings have a very large impact on estimates of extreme poverty. Meyer et al. showed that correcting survey-reported earnings using administrative data cuts estimates of extreme poverty in half in

the CPS (see their table 6), and the effect of this adjustment is nearly identical in the Survey of Income and Program Participation (SIPP).

Brady and Parolin argue that "this evidence cannot be used to directly critique our analyses" because Meyer et al.'s results are for a lower threshold and for a single year. Regarding the threshold, the results from Meyer et al. clearly show that for the SIPP, administrative-based earnings adjustments have a significant impact at higher thresholds: among the 0.84% of households in the survey below $2/day based on their reported income, 33% would have been falsely classified as deep poor, and 24% would have been falsely classified as poor (Meyer et al. 2019: figure 3). In other words, because the errors are so large, they affect higher thresholds as well.

Although Meyer et al. did not report these same estimates for the CPS, the large impact of this earnings adjustment at higher thresholds in the SIPP certainly gives rise to concerns that underreported earnings could significantly bias similar estimates in the CPS, particularly given the evidence that this adjustment has the same effect in both the CPS and the SIPP at the extreme poverty threshold.

The authors also discount the relevance of Meyer et al.'s results because the results are for only a single year; Brady and Parolin argue that significant measurement error at a point in time is not relevant for changes over time. In order to maintain that sizable measurement error at a point in time does not bias trends, one needs to impose the strong assumption that this error is constant over time. This is a strong assumption, and there are important reasons why one might be skeptical about it given other evidence showing a rise in income underreporting and, more generally, a decline in survey quality over time (for a summary, see Meyer et al. 2015).

In addition to the upward bias due to underreported earnings, estimates of extreme poverty based on survey data will be biased upward because of the well-documented underreporting of other sources of income, such as transfer and retirement income. Brady and Parolin do not address the significant underreporting of Supplemental Security Income, retirement income, and other important income sources. They do adjust for the underreporting of Supplemental Nutrition Assistance Program and Temporary Assistance for Needy Families benefits using a microsimulation model, but, as I note in my comment, several studies have shown that adjustments using microsimulation models misallocate imputed benefits to the wrong parts of the distribution (Mittag 2019; Shantz and Fox 2018; Stevens et al. 2018). In particular, these imputations overallocate benefits to those at the very bottom. Although any overallocation to the bottom would bias estimates of extreme poverty downward, comparisons of Brady and Parolin's estimates to those that correct for underreporting of earnings, transfer income, and other sources using administrative data show that on net, their estimates are overstated.

Brady and Parolin argue that their results are not biased upward by suggesting that their estimates are "nearly identical" to those of Meyer et al. (2019). They note that their estimate of $2/day poverty in 2011 is 0.29%, and those from Meyer et al. (2019) for the same year range from 0.18% to 0.29%. The range of estimates they report from Meyer et al. (2019), however, are for households, and the 0.29% estimate is for the SIPP. The comparable estimate from Meyer et al., the one for individuals in the CPS, is 0.13% (the SIPP estimate for individuals is 0.11%), indicating that Brady and Parolin's estimate of extreme poverty using the $2/day threshold is biased upward by more than 120% (see Meyer et al. 2019:32, and table 6). Moreover, this upward bias is not unique

to the $2/day threshold. Adjusting for underreporting indicates a significant upward bias in survey-based estimates at higher thresholds (see Meyer et al. 2019: tables 7 and figures 2 and 3).

Conclusions

A growing literature using survey data linked to administrative microdata presents strong evidence that income is underreported at the bottom. Results from this literature show that the estimates of extreme poverty from Brady and Parolin are significantly overstated. The additional analyses that Brady and Parolin present in their response do not address this large bias.

An important limitation of studies of extreme poverty based on linked survey and administrative data is that the evidence is available for only a single year. For this reason, these studies cannot be used to assess the validity of Brady and Parolin's conclusion that extreme poverty has risen sharply over time. However, if the significant underreporting of income from many sources is rising over time, as suggested by several studies, then Brady and Parolin's estimates of the trends in extreme poverty may be biased upward.

It is important to note that the studies of extreme poverty discussed here are based on samples from surveys that exclude the homeless, those living in prisons and nursing care facilities, and other institutionalized individuals. The exclusion of these individuals may lead to a significant understatement of extreme poverty. Both Meyer et al. (2019) and Brady and Parolin emphasize this point. Future research should explore further the economic circumstances of the homeless and other populations missed in surveys and examine the impact that including such populations would have on estimates of extreme poverty.

In addition, very low estimates of extreme poverty do not necessarily imply an absence of extreme economic deprivation. All the estimates discussed here consider poverty over a year. Individuals and families with incomes above an extreme poverty threshold for a year may very well experience bouts of extreme poverty over shorter periods.

Acknowledgments I thank Bruce Meyer and Derek Wu for helpful comments and suggestions.

References

Medalia, C., Meyer, B. D., O'Hara, A. B., & Wu, D. (2019). Linking survey and administrative data to measure income, inequality, and mobility, *International Journal of Population Data Science, 4*(1). https://doi.org/10.23889/ijpds.v4i1.939

Meyer, B. D., Mok, W. K. C., & Sullivan, J. X. (2015). Household surveys in crisis. *Journal of Economic Perspectives, 29*(4), 199–226.

Meyer, B. D., Wu, D., Mooers, V. R., & Medalia, C. (2019). *The use and misuse of income data and extreme poverty in the United States* (NBER Working Paper No. 25907). Cambridge, MA: National Bureau of Economic Research.

Mittag, N. (2019). Correcting for misreporting of government benefits. *American Economic Journal: Economic Policy, 11*(2), 142–164.

Shantz, K., & Fox, L. E. (2018). *Precision in measurement: Using state-level Supplemental Nutrition Assistance Program and Temporary Assistance for Needy Families administrative records and the Transfer Income Model (TRIM3) to evaluate poverty measurement* (SEHSD Working Paper #2018-30). Retrieved from https://www.census.gov/content/dam/Census/library/working-papers/2018/demo/SEHSD-WP2018-30.pdf

Stevens, K., Fox, L. E., & Heggeness, M. L. (2018). *Precision in measurement: Using state-level SNAP administrative records and the Transfer Income Model (TRIM3) to evaluate poverty measurement* (SEHSD Working Paper #2018-15). Retrieved from https://www.census.gov/content/dam/Census/library/working-papers/2018/demo/SEHSD-WP2018-15.pdf