

Poverty Analysis based on Kernel Density Estimates from Grouped Data¹

Camelia Minoiu²

Abstract submission to the Second Meeting of the Society for the Study of Economic Inequality (ECINEQ)

Kernel density estimation (KDE) has been recently used to measure poverty from grouped data (representing mean incomes of a small number of population quantiles). Global poverty estimates based on this technique have been widely debated. In this paper I analyze the performance of this method. I undertake Monte Carlo simulations to determine the size of the biases associated with various poverty indicators and poverty lines using data drawn from plausible theoretical income distributions. I also use unit data from three nationally representative household surveys (for Nicaragua, Tanzania and Vietnam) with a view to comparing KDE-based poverty estimates from grouped data with their survey counterparts.

I show that the technique gives rise to biases in poverty whose sign and magnitude vary with the smoothing parameter, the kernel, the number of data-points analyzed, and the poverty indicators used. For a range of plausible unimodal distributions, I find that the average income of the poorer individuals is overstated, while the average income of the less poor is understated by the technique. The biases associated with the poverty headcount ratio can be as high as 6-7 percentage points (for unimodal distributions) and 10-11 percentage points (for a multimodal distribution).

I also demonstrate that KDE-based *global* poverty rates and headcounts are highly sensitive to the choice of smoothing parameter. Depending on that parameter, the estimated proportion of '\$1/day poor' in 2000 varies by a factor of 1.8, while the estimated number of '\$2/day poor' in 2000 varies by 287 million people.

These findings give rise to serious concern about the validity and robustness of kernel density estimation on grouped data when used for purposes of poverty assessment. The magnitude and direction of biases identified in this study can, however, serve as a framework of interpretation to researchers who wish to employ nonparametric smoothing techniques on grouped data.

¹ The full paper is available on: www.columbia.edu/~cm2036/kdepoverty.pdf

² Dept. of Economics and Institute for Social and Economic Research and Policy, Columbia University.
Email: cm2036@columbia.edu. Tel. +1 212 854 6385 (office). Fax +1 212 854 7998 (office).
Website: www.columbia.edu/~cm2036