The Robustness of the Monetary Cost of Children to Data Problems

by

Stefan C. Norrbin
Professor of Economics
Department of Economics
Florida State University
Tallahassee, FL
snorrbin@fsu.edu

David Macpherson
E.M. Stevens Professor of Economics
Department of Economics
Trinity University
San Antonio, TX
dmacpher@trinity.edu

Thomas S. McCaleb
Associate Professor of Economics
Department of Economics
Florida State University
Tallahassee, FL
tmccaleb@fsu.edu

Topic Area: Economics
Abstract:

The cost of children is of considerable societal importance. Either explicitly or implicitly society supports the cost of raising a child. Examples of such monetary support are: tax deductions for children, welfare payments for children in low income families, international aid programs to combat child poverty, and child support payments if the parents are not jointly raising the child. Therefore much attention has been devoted to attempting to estimate the monetary cost of children.

The precise amount of monetary costs of raising children is difficult to establish, due to the joint nature of actual spending by families. Therefore, econometric estimation techniques have been developed to compute the likely cost of children. Most of the early studies used methodologies that were based on some form of Engel curve estimation. The original Engel curve estimation (Engel, 1857) used food as a cost indicator, and compensated the childless couple for the additional consumption necessary to reach the same food to total consumption ratio as a couple with children. For example, Espenshade (1984) used this approach, and it has become the basis for most of the child support guidelines in the U.S. Deaton and Muellbauer (1986) show that the Engel models generally overstate the cost of children. This has resulted in a search for other theoretical frameworks that can be applied to the existing data. No clearly superior framework has yet been adopted.

Instead of examining the differences in monetary costs based on different theoretical approaches, we focus in this paper on the data used in the computation. We show that the assumptions made about data selection affect the estimated costs substantially. Although similar techniques are used in different states in the U.S. to calculate child support guidelines, the estimated cost of children differs substantially. In 2011, the cost of a single child in a particular income bracket would range from 14%-25% of gross income across states in the U.S. (Norrbin et al., 2011). We show that these differences can be explained by data selection problems. This finding is similar to the finding in the Indian National Sample Survey that changing the recall period in the survey reduces poverty rates by half, removing 200 million people poverty (Deaton, 1983). Data issues might be more important than theoretical modeling, in the case of computing the cost of children.

Typically, empirical researchers apply univariate bounds to variables to avoid outliers. We show that such bounds are insufficient, and that outliers that remain in the data set substantially alter the estimated costs of children. To find the outliers, we use a multivariate outlier detection method by Filzmoser (2005). This method extends the Mahalanobis measure by finding a statistical break for when the data point can be classified as an outlier. We explore the findings from regressions with and without outliers, and compare these findings with robust regression methods. Our preliminary estimates indicate that the cost of children as a percent of income can more than double when outliers are removed from the data set.
References:


