

On The Robustness of Coefficient Estimates to the Inclusion of Proxy Variables

Christopher R. Bollinger* Jenny Minier

October 2006

Preliminary and Incomplete: Please Do Not Circulate or Quote
without Permission

Abstract

This paper considers the most effective use of multiple proxy measures for the same unobserved variable. It extends the results of Lubotsky and Whittenberg (2006) to examine the impact of proxy variables on correctly measured variables. We find that including all proxy variables in the regression minimizes the bias on all other coefficients in the regression. Unlike previous results, estimates of coefficients on other regressors do not require a scaling assumption. We derive a set of bounds based on results from Klepper and Leamer (1984) and Bollinger (2003) for parameters in the model. These results are compared to Extreme Bounds Analysis. We find through Monte Carlo results that our bounds perform better than extreme bounds in most circumstances. We also find that our results may overturn many of the results found through extreme bounds analysis. We conclude with an empirical example from the cross-country growth literature in which human capital is measured through three proxy variables: literacy rates, and enrollment in primary and secondary school. We find that the coefficient estimate on initial income is "robust," as previous extreme bound analyses have concluded. However, in contrast to previous results, we find that the coefficient estimate on investment cannot be distinguished from zero, while that on population growth is robustly statistically different from zero.

*Both authors are Department of Economics, University of Kentucky, Lexington, KY 40506. Email: crboll@uky.edu and jminier@uky.edu respectively. We thank Josh Ederington and Jim Ziliak for useful discussion.