COMMENTS ON

"The Strengths and Weaknesses of Factor Analysis in Predicting Cuba's GDP," by Nicolás Sánchez and Miles Cahill

Roger R. Betancourt

This paper is interesting, well written, and for the most part insightful. Nonetheless, it fails to appreciate the main weakness of factor analysis for predicting Cuba's GDP, or that of any other command economy.

The essence of the technique, as the authors clearly explain, is to gather data on a number of socioeconomic indicators (37 of them in this case) for various countries (19 Latin American ones plus Cuba in this case) and to perform a factor analysis on these indicators that generates orthogonal components or factors and corresponding factor scores. These factor scores become (in this case the analysis generated seven factors) seven explanatory variables in a regression through the origin with GDP per capita as the dependent variable for the 19 Latin American countries that have data on GDP per capita. Cuba's GDP per capita can then be predicted by using the coefficients of this regression for each of the seven factors and Cuba's factor scores, based on the 37 underlying indicators for which Cuba does have data.

The results of this exercise are reported in Table 3 of the paper. If one uses GDP per capita in nominal or official exchange rates for the 19 Latin American countries, then Cuba's GDP will also be in terms of these units; if one uses GDP per capita in terms of purchasing power parity exchange rates or ICP$, then Cuba's GDP prediction would be in terms of these rates. Table 3 is in terms of official or nominal exchange rates and it predicts a GDP per capita of $2,578 for Cuba in 1990. One of the figures cited in the text for comparison, however, is from a United Nations *Human Development Report* that gives a figure for Cuba of $2,200 *in terms of purchasing power parity rates or ICP$* in 1990. The prediction for Cuba if one uses factor analysis to predict Cuba's GDP per capita in ICP $ is given by the authors in footnote 8, namely $5,223. This brings out that the prediction is way out of line, in contrast to the authors' assertion in the paper. The latter results from emphasizing the comparison with GDP's per capita in nominal exchange rates and ignoring that at least one of the estimates they use for comparison, namely the one from the *Human Development Report*, refers to ICP$.

This raises two related issues: first, which is the relevant magnitude to predict? and second, why is the difference so great when using ICP$? It is now well accepted in the economics literature that GDP per capita in ICP$ is the proper measure of welfare for international comparisons, because it incorporates both traded goods and nontraded goods and their prices. The difference is so large when using ICP$ precisely because command economies are even more distorted when it comes to nontraded goods and their prices than with respect to other sectors. The GDPs in ICP$ of the Latin American countries used as the dependent variables contained sizable nontraded sectors, for example retail and wholesale trade or banking and financial services. In 1990 Cuba these sectors barely existed (for example banking and in-

surance) or existed at a low level (for example retail and wholesale trade). This is an important if not the most important reason why it makes no sense to use factor analysis to predict Cuba's GDP per capita in ICP$, which is what one wants to measure in terms of welfare. The authors' own numbers indicate how far the prediction is from other **comparable** estimates—100% overprediction!