

Estimation of the distribution of income from survey data, adjusting for compatibility with other sources.¹

Víctor Alfredo Bustos y de la Tijera²

INEGI

Abstract

In this paper we present an approach for the estimation of income distributions, which is aimed at dealing with survey data shortcomings through simultaneous consideration of other statistical sources and through adjustment for compatibility with all of them. Our proposal is based on well-established statistical criteria and methods and thus reduces the need for subjective or arbitrary choices. It has the purpose of selecting the distributional model that best fits the data from the survey, using a Constrained Pseudo Log-likelihood criterion. We show how our proposal deals both with income under-reporting and with truncation which are known to be present in the survey. We then apply our procedure to Mexican data from the National Survey on Household Income and Expenditure for the year 2012, and from Mexico's System of National Accounts, sources that produce widely differing results regarding total household current income for the country. We show that, among all fitted models, a satisfactory explanation is given by a 4-parameter Generalized Beta Type 2 distribution. The chosen distribution has little impact on the official poverty measurement. The Gini coefficient, however, reaches a value as high as 0.803.

¹ The results shown in this paper do not necessarily reflect those of Mexico's National Institute of Statistics and Geography (INEGI) and are not to be considered official in any way.

² Dr. Víctor Alfredo Bustos y de la Tijera, researcher, INEGI, Av. Héroe de Nacozari Sur No. 2301, Aguascalientes, Ags., 20276 México, Tel. (+52) 449 910-5431, e-mail: alfredo.bustos@inegi.org.mx