Generalized Instrumental Variable Models

Andrew Chesher

CeMMAP and University College London

Abstract

A procedure for determining the identifying power of a very general class of instrumental variable (IV) models is presented.

The models considered have a scalar outcome Y determined by a function of a vector of explanatory variables, X, and a vector of unobservable variables, U, with distribution function F_U, thus.

$$Y = h(X, U)$$

Instrumental variables Z are excluded from the function h. Unobservable U and the instrumental variables Z are independently distributed. X may be endogenous in the sense that U and X are not restricted to be independently distributed. The identification analysis focuses on the identifiability of the function h.

Until now IV models have required (i) Y to be continuous, and (ii) U to be scalar and (iii) h to be monotone in U. The proposed procedure delivers results when Y is discrete, when U has higher dimension than Y and when h is not restricted to be monotone in unobservable U.

An important feature of these models is that they are typically set rather than point identifying. The extent of identified sets depends on the nature of the restrictions imposed on the components of structures: $\{h, F_U\}$.

The application of the procedure is illustrated using IV models for univariate binary and ordered outcomes and an IV model of choices amongst unordered outcomes. The latter is an IV extension of Daniel McFadden's classic conditional logit model in which explanatory variables are allowed to be endogenous.