Overcoming Endogeneity and Sparse-Data Bias in Consumer-level Random-Coefficient Discrete Choice Models Using Copulas

Fan Yang
NEOMA Business School
Xixi Hu
University of British Columbia
Yi Qian
University of British Columbia
Hui Xie
Simon Fraser University

Overcoming Endogeneity and Sparse-Data Bias in Consumer-level Random-Coefficient Discrete Choice Models Using Copulas

Abstract

The pervasive presence of endogeneity poses a considerable challenge for establishing causal inference in choice models, and the increasing availability of rich consumer-level choice data further intensifies the technical difficulties. The classical econometric method to address endogeneity in the choice model (e.g., the product-market controls approach initially developed by Berry 1994 and Berry et al. 1995) requires IVs that must satisfy the stringent condition of exclusion restriction. Moreover, it imposes rigorous conditions on the data and is prone to failure in certain scenarios. In this paper, we develop an alternative IV-free control function method named as CF-2sCOPE that applies copulas to address the endogeneity concern in the demand model. The main contribution of this approach is three-fold. First, unlike the existing methods in the literature, the proposed CF-2sCOPE method eliminates the reliance on IVs, which are often difficult to find and validate in practice. We achieve this by incorporating generated regressors derived from existing regressors in the choice model to control for endogeneity. Second, our method will not suffer from the data sparsity issue (i.e., close-to-zero market shares), which can be very common in practice and would lead to the failure of the conventional BLP method. Third, our method addresses the challenges posed by heterogeneous regressors and can even account for heterogeneous endogeneity, which often arises in individual targeting scenarios, a context in which the conventional BLP method is not applicable. We further demonstrate the performance of CF-2sCOPE and compare it with existing methods using simulation studies and illustrate its use in an empirical application.

Our paper is intended for the 'Methods, Modelling & Marketing Analytics' track.

Keywords: Random Coefficient Logit Model, Endogeneity, Instrument-free Copula method