Testing the Exogeneity of Instrumental Variables in Discrete Choice Models

by

C. ANGELO GUEVARA
Departamento de Ingeniería Civil
Universidad de Chile
Facultad de Ingeniería y Ciencias Aplicadas
Blanco Encalada 2002, Santiago, Chile
Tel: +562 29784380
crguevar@ing.uchile.cl

EXTENDED ABSTRACT

When the error term of a discrete choice model is not independent of the observed variables, conventional estimators of the model parameters are inconsistent, making the model misleading for behavioral assessment and policy design (see e.g. Guevara and Thomas, 2007). This problem is known as endogeneity and may be caused by three main reasons: errors in variables, simultaneous determination and the omission of attributes of alternatives (Guevara, 2015). Endogeneity is common to several types of discrete choice models used in transportation analysis, including, but not limited to the following: airline itinerary choice (Lurkin et al, 2017), mode choice (Fernandez-Antolin, et al, 2016), passenger booking timing (Wen and Chen, 2017), learning models of route choice (Guevara et al, 2017), mobility data collection (Zegras et al, 2018), demand for electric vehicles (Helveston, 2016), valuation of public transport attributes (Guevara et al, 2018) and residential choice (Guevara 2006, 2010; Guevara and Ben-Akiva 2006, 2012; Guevara and Polanco, 2016).

The canonical methods to correct for endogeneity in discrete choice models rely in the availability of suitable instrumental variables. Instruments have to be relevant (correlated with the endogenous variable) and also exogenous (independent of the error term of the model). Testing whether or not the instruments are relevant (or not weak), can be achieved by analyzing the correlation between the instrument and the endogenous variable. Instead, testing instruments exogeneity is particularly difficult because the error term of the model is not observed. This article proposes two novel tests for achieving this goal in discrete choice-models, and assesses their size, power and robustness using a Monte Carlo experiment.

Tests for the exogeneity of instruments rely on over-identification, that is, on having more instruments than endogenous variables. For linear models, Sargan (1958) noted that when the problem is over-identified the residuals of the instrumental-variables regression can be used to test for the exogeneity of the instruments. For nonlinear models, including discrete choice models, such as the Logit or the Probit, Lee (1992) noted that an estimator developed by Amemiya (1978), and studied by Newey (1987), can play the role of the Sargan test in the validation of instruments in this context. This test is usually termed as
the Amemiya-Lee-Newey test, it relies in the estimation of an auxiliary GMM model build from reduced-form estimates, and is the state of the art in the subject.

Over-identification test for the exogeneity of the instruments have an important limitation. Newey (1985) showed that these tests are inconsistent, which means that they are blind to certain alternative hypotheses, for which power is equal to size, even if the sample size goes to infinity. One way to recover consistency is to consider that the over-identification tests work under the assumption that, at least, a subset of the instruments, for which the model becomes just identified, is exogenous (Stock, 2001). This additional assumption cannot be proven, discouraging the use of methods to correct for endogeneity that are based on instrumental variables, because they rely on an unverifiable assumption.

De Blander (2008) proposes an alternative way to attain consistency of over-identification tests. He notes that the alternative hypothesis for which over-identification tests are blind is very peculiar, so he recommends instead to assume that this alternate hypothesis does not hold, putting then “the burden of proof … on the critic, who has to make the case why the instruments” would fulfil this rare condition. De Blander (2008) shows that consistency would fail if the way in which the instruments appear in the structural equation and the reduced form equation, are linearly dependent. Pleus (2015) provides a more general expression for this result, building on Newey (1985), and presents a graphical representation to illustrate the nature of the problem. Parente and Silva (2012) identify one plausible case in which this may occur in practice, when both instruments are of the same nature, that is, if they come from the same source, because then their correlation with both the endogenous variable and the error term will likely be very similar. A similar warning was suggested by Nichols (2007), although justified in a different ground. Beyond Parente and Silva’s (2012) warning, which is fully addressable in practice, it seems easier to defend than to attack the plausibility of the consistency of the over-identification tests.

This article focuses in the development and the assessment of tests for the exogeneity of instruments in discrete choice models. First, the state of the art Amemiya-Lee-Newey test is reviewed and compared with two novel tests that are constructed as adaptations of the Refutability and the Hausman test into the discrete choice framework. In the Refutability test, one or up to all instruments are included as additional variables in an auxiliary model that was corrected for endogeneity using the full set of instrumental variables. A version of this test, when not all instruments are included in the auxiliary model, was originally suggested by Guevara (2010), who termed it the Direct test. The Hausman test is built from the comparison of the estimates attained using all and a subset of instruments in the correction of endogeneity, taking advantage of the differences on efficiency that would be attained for each case.

The three types of tests are assessed using a binary Logit Monte Carlo experiment in terms of power, size, and their robustness to De Blander’s alternate hypothesis to which all over-identification tests of this kind are blind. Results suggests that the version of the proposed Refutability test that uses all instruments has larger power, smaller size distortion, and is more robust, compared to the state of the art Amemiya-Lee-Newey test.
Also, the proposed Refutability test is superior to the Hausman test, not only for the same reasons, but also in that it has a null hypothesis that allows being agnostic about which instrument might be endogenous. Furthermore, the proposed Refutability test can be readily applied with canned estimation software with much lower computation costs, making it an attractive tool for practical applications and for model building.

**References**


