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

Amemiya, T. (1978), "The Estimation of a Simultaneous Equation Generalized Probit Model," **Econometrica,** 46, 1193-1205.

Basman, R. (1960), "On Finite Sample Distributions of Generalized Classical Linear Identifiability Test Statistics," **Journal of the American Statistical Association,** 5, 650-659.

Berry, S., J. Levinsohn and A. Pakes (1995), "Automobile Prices in Market Equilibrium," **Econometrica**, 63(4), 841-90.

Bresnahan, T.F. (1997), "Comment", in T.F. Bresnahan and R. Gordon eds., **The Economics of New Goods**. (University of Chicago Press, Chicago).

C. Angelo Guevara, C.A., Tirachini, A., Hurtubia, R. and Dekker, T. (2018) Correcting for Endogeneity due to Omitted Crowding in Public Transport Choice Using the Multiple Indicator Solution (MIS) Method. Working paper, Universidad de Chile.

Card, D. (1995), "Using Geographic Variation in the College Proximity to Estimate the Return to Schooling," in Christophides, Garnt and Swidinsky eds. **Aspects of Labour Market Behaviour: Essays in Honor of John Vanderkamp**, University of Toronto Press, Toronto, Canada, 201-222.

Chesher, A. (2010), "Instrumental Variables Models for Discrete Outcomes", **Econometrica**, 78, 575-601.

Cox, D. and E. Snell (1968), "A General Definition of Residuals," **Journal of the Royal Statistical Society B**, 30, 248-275.

De Blander, R. (2008), "Which Null Hypothesis Do Overidentification Restrictions Actually Test?" **Economics Bulletin**, 3(76), 1-9.

Engle, R. (1984), "Wald, Likelihood Ratio and Lagrange Multiplier Tests in Econometrics," in Griliches and Intrilligator eds. **Handbook of Econometrics**, II, Amsterdam, North Holland.

Fernández-Antolín, A., Guevara, C. A., de Lapparent, M., & Bierlaire, M. (2016). Correcting for endogeneity due to omitted attitudes: Empirical assessment of a modified MIS method using RP mode choice data. **Journal of Choice Modelling**, 20, 1-15.

Greene, W. (2003), **Econometric Analysis**, 5th Edition, Prentice Hall, New York.

Guevara, C. (2010), "Endogeneity and Sampling of Alternatives in Spatial Choice Models," Ph.D. Thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Guevara, C. A., Tang, Y., & Gao, S. (2017). The initial condition problem with complete history dependency in learning models for travel choices. **Transportation Research Part B: Methodological.** Forthcomming

Guevara, C. and M. Ben-Akiva (2006), "Endogeneity in Residential Location Choice Models," **Transportation Research Record,** 1977, 60-66.

Guevara, C. and M. Ben-Akiva (2008), "A Lagrange Multiplier Test for the Exogeneity of instruments in MNL Models: An Application to Residential Choice," presented at the European Transport Conference, Leeuwenhorst, The Netherlands.

Hahn, J. and J. Hausman (2002), "Weak Instruments: Diagnosis and Cures in Empirical Econometrics," **American Economic Review,** 93, 118-128.

Hahn, J. and J. Hausman (2002), "Weak Instruments: Diagnosis and Cures in Empirical Econometrics," **American Economic Review,** 93, 118-128.

Hausman, J. (1978), “Specification Tests in Econometrics,” **Econometrica,** 46, 1251-1272.
Hausman, J. (1996) ‘‘Valuation of New Goods under Perfect and Imperfect Competition,’’ in **The Economics of New Goods, Studies in Income and Wealth** 58, Bresnahan and Gordon eds., Chicago: National Bureau of Economic Research.
Hausman, J. (1997), “Valuation of New Goods under Perfect and Imperfect Competition”, in T.F. Bresnahan and R. Gordon eds., **The Economics of New Goods**. (University of Chicago Press, Chicago).
Heckman, J. (1978), “Dummy Endogenous Variables in a Simultaneous Equation System,” **Econometrica,** 46, 931-959.
Helveston, J. P. (2016). Development and Adoption of Plug-in Electric Vehicles in China: Markets, Policy, and Innovation. PhD Thesis, Carnegie Mellon Univeristy.
Karaca-Mandic, P. and K. Train (2003), “Standard Error Correction in Two-stage Estimation with Nested Samples,” **Econometrics Journal**, 6(2), 401-407.
Lee, L. (1992), “Amemiya's Generalized Least Squares and Tests of Overidentification in Simultaneous Equation Models with Qualitative or Limited Dependent Variables,” **Econometric Reviews**, 11(3), 319-328.
Lurkin, V., Garrow, L. A., Higgins, M. J., Newman, J. P., & Schyns, M. (2017). Accounting for price endogeneity in airline itinerary choice models: An application to Continental US markets. **Transportation Research Part A: Policy and Practice**, 100, 228-246.
Martinez, L. and J. Viegas (2009), “Effects of Transportation Accessibility on Residential Property Values: A Hedonic Price Model in the Lisbon Metropolitan Area,” **Transportation Research Record**, 2115, 127-137.
Martinez, L., J. Abreu and J. Viegas (2010), “Assessment of Residential Location Satisfaction in Lisbon Metropolitan Area,” presented at the 89th Transportation Research Board Annual Meeting, Washington, DC.
McFadden, D. (1987), “Regression Based Specification Tests for the Multinomial Logit Model,” **Journal of Econometrics**, 34, 63-82.
Nevo, A. (2001) “Measuring Market Power In The Ready-To-Eat Cereal Industry,” **Econometrica**, 69 (2), 307-342.
Newey, W. (1985), “Generalized Method of Moments Specification Testing,” **Journal of Econometrics,** 29, 229–256.
Newey, W. (1987), “Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables,” **Journal of Econometrics,** 36, 231-250.
Nichols, A. (2007), “Causal Inference with Observational Data,” **Stata Journal**, 7 (4), 507–541.
Parente, P. M. D. C. and J. M. C. S. Silva (2012): “A Cautionary Note on Tests of Overidentifying Restrictions,” **Economic Letters**, 115, 314–317.
Park, S. and S. Gupta (2009), “A Simulated Maximum Likelihood Estimator for the Random Coefficient Logit Model Using Aggregate Data,” **Journal of Marketing Research**, 46(4), 531-542.
Petrin, A. and K. Train (2002), “Omitted Product Attributes in Discrete Choice Models,” Working Paper, Department of Economics, University of California, Berkeley, CA.
R Development Core Team (2008), **R: A Language and Environment for Statistical Computing**, R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org.
Rao, C. and S. Mitra (1971), **Generalized Inverse of a Matrix and its Applications**, J. Wiley, New York, NY.
Sargan, J. (1958), “The Estimation of Economic Relationships Using Instrumental Variables,” **Econometrica**, 26, 393-415.
Stock, J. (2001), “Instrumental Variables in Statistics and Econometrics,” in Smelser and Baltes eds. **International Encyclopaedia of the Behavioural Sciences**, Elsevier Publishing, New York, 7577–7582.

Stock, J., J. Wright and M, Yogo (2002), "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," **American Statistical Association Journal of Business and Economic Statistics,** 20(4), 518-529.

Villas-Boas, A. and R. Winer (1999), "Endogeneity in Brand Choice Models," **Management Science**, 45, 1324–1338.

Wen, C. H., & Chen, P. H. (2017). Passenger booking timing for low-cost airlines: A continuous logit approach. **Journal of Air Transport Management**, 64, 91-99.

Zegras, P. C., Li, M., Kilic, T., Lozano-Gracia, N., Ghorpade, A., Tiberti, M., ... & Zhao, F. (2018). Assessing the representativeness of a smartphone-based household travel survey in Dar es Salaam, Tanzania. **Transportation**, 1-29.