Resurgence of Instrument Variable Estimation and Fallacy of Endogeneity

The author considers the history of instrumentals variables (IV) estimation in three contexts: simultaneous equations models (SEMs), censored regression models and the estimation of treatment effects in programme evaluation.

The author is sceptical about the value of IV estimation although she nowhere goes so far as to claim that it is invalid. The paper is strong on rhetoric but the historical perspective tends to conflate the different problems which can arise in IV modelling. There appear to be a number of different claims:

1. The author's strong preference is for conditional representations. This puts her is the tradition of time series statistics and, in econometrics, of Herman Wold. Wold regarded the world as recursive and argued that simultaneity is a consequence of temporal aggregation. Nevertheless, we are generally obliged to use temporally aggregated data. The opposing tradition in economics comes out of general equilibrium analysis and in particular the Walrasian general equilibrium model which is the context in which the Cowles SEM approach emerged. She implicitly argues (pages 17-18) for the VAR approach. This is fine in simple modelling contexts but is often not useful in policy analysis in which governments or agencies need to be able to trace the impacts of developments or changes on variables of interest. This will generally require specification of a structural model. Note that such exercises are not primarily concerned with generating optimal forecasts and so the fact that they do not do this is not conclusive evidence against them.
2. The author claims that instrument validity can only be inferred in the context of a fully specified model. This statement is correct if the word "only" is omitted. For example, a set of deterministic functions of time will often serve as valid instruments in a time series context. However, if one follows the author's line then there would not appear to be any problem with IV estimation, using as instruments those model variables excluded from the equation in question, in particular since IV and FIML estimations are asymptotically equivalent (explaining the general preference for FIML over 3SLS).

3. Heckman's limited dependent variable (LDV) model emerged as a generalization of the Tobit estimator. In Tobit, identification is the consequence of the (strong) equality restriction on the parameters of the measurement and observation equations. In the Heckman model, identification comes from exclusion restrictions on the measurement equation. The author is correct in implying that these restrictions often have little basis. It is nevertheless possible to think of circumstances in which the restrictions are valid, for example, when only subset of data is available of a variable of interest for purely administrative reasons.
4. The same points arise in relation to the use of IVs in the estimation of programme treatment effects. However, this does not imply that identification is always impossible – the Angrist (1990) paper provides a clear example.

Structure and overall argument

Given that IV estimation was born as a means of dealing with measurement error, it would be worth devoting a little more space to that application which the author appears to regard as uncontroversial. At the other end of the paper, it would be useful to add a discussion of propensity score matching which provides a different solution to the simultaneity and non-randomness issues in programme evaluation.

An overriding issue with the paper is that the author mixes historical analysis with methodological prescription in a manner which will leave many readers uncomfortable. While I am aware that this is her style, I believe that she would attract a wider readership if she could focus on the two issues separately. I wonder whether she might consider two papers- one on the history of instrumental variables and a second, perhaps drawing on the historical paper, containing the methodological considerations.

Section 2.1

The author emphasizes "incredible" identification restrictions as responsible for the move away from the SEM to VAR models. This was one factor but in my view the availability of quarterly, and later monthly, macroeconomic data series was more important. With higher frequency data, dynamic interactions became more important and the potential for simultaneous equations bias less serious.

Additional references which might be useful are Fox (1956)on the poor performance of LIML relative to OLS and Nelson (1972) on the outperformance of standard macroeconomic models by simple time series representations.

Other points: (a) More detailed reference to textbook discussions would be useful (page 7). (b) Give some detail on the LSE approach (page 8).

Section 3.1

Am I misunderstanding something here? Equation (5) is unidentified unless $a_{12,1} = 0$. (However, that would seem an incredible restriction).

References

Fox, K.A. (1956). "Econometric models of the United States". *Journal of Political Economy*, **64**, 128-142.

Nelson, C.R. (1972). "The prediction performance of the FRB-MIT-Penn model of the US economy". *American Economic Review*, **62**,902-917.