An Instrumental Variable Consistent Estimation Procedure to Overcome the Problem of Endogenous Variables in Multilevel Models

Neil H Spencer
University of Hertfordshire

Antony Fielding
University of Birmingham
& Multilevel Models Project, Institute of Education, University of London

The Problem of Endogeneity

It is far from unusual for a multilevel model to contain a regressor that can be regarded as an endogenous variable. The term endogeneity as opposed to exogeneity is a familiar term in econometrics. Often it manifests itself by explanatory variable being subject to the same influences as the response variable. It is thus not exogenous in the model being fitted. More particularly it may mean that variables which are regarded as endogenous are not independent of the random effects in the model. In such circumstances a basic assumption of modelling is not met and obtaining consistent estimators of the parameters is not straightforward. Many standard multilevel procedures (e.g. Iterative Generalised Least Squares) rely on the independence of regressors and model disturbances for their consistency properties. Obviously in the case of the presence of endogenous explanatory variables this does not generally hold. We present here a modelling strategy based on instrumental variables and introduce a MlwiN macro that provides desired consistent estimators.

We will consider an example in which we wish to use a multilevel model and where we suppose an explanatory variable may be endogenous. Fielding (1998) introduces a dataset drawn from children in primary schools of the City of Birmingham Local Education Authority. Data is available on a range of school and pupil characteristics but of prime importance are factors influencing the results of National Curriculum Key Stage 1 tests. In a simple model we may wish to relate one of these test results to gender and age of the child in months taking into account as controls the results of baseline tests carried out when the child entered reception classes in school. A model such as this may be used to examine the progress children are making in different schools. For pupil i in school j and where we have p baseline test results we may write a model which we label (1) as

$$\mathbf{KS1TEST}_{ij} = \beta_0 + \beta_1 \mathbf{GENDER}_{ij} + \beta_2 \mathbf{AGE}_{ij} + \sum_{k=1}^{p} \mathbf{BASELINE - TEST(k)}_{ij} + \mathbf{u}_j + \varepsilon_{ij}$$

The term u_{ij} in model (1) represents a random school effect and ε_{ij} is a within school random pupil effect. The endogeneity in this model arises because the baseline tests may be supposed to be related to the random pupil effect through the existence of important unmeasured and unmeasurable influences acting at this lowest level of the hierarchy (e.g. home circumstances). These influences are incorporated in the disturbance ε_{ij} but may also influence baseline test performances. It is also possible that there are some influences that make the baseline tests related to the school effect

u_j. The common influences may be such things as the locality in which the school is situated and from which the pupils generally come.

Overcoming the Problem of Endogeneity

Solutions to the problem of inconsistency caused by endogenous regressors, particularly when they are thought to relate to higher level effects such as u_j, have been proposed by Kiviet (1995) and Rice *et. al.* (1998). Kiviet uses a bias corrected version of the least squares dummy variable estimator (LSDV). Rice use conditioned iterative generalised least squares (CIGLS). The first of these approaches suffers from a problem that the bias correction applied may actually increase the bias in some circumstances. Neither approach can directly easily cope with the case where the level 1 (pupil) random effect is correlated with regressors. It is this latter situation on which we mainly focus.

In econometrics and other literature a frequently used method of overcoming such endogeneity problems in fixed effects models is to use instrumental variable techniques. We adapt these techniques to cover multilevel random effects models. The latter possibility has been mentioned briefly and independently by Rice *et. al.* (1998). Spencer (1997) also suggested such an approach for repeat testing in educational situations where explanatory variables are lagged versions of the response.. Here we operationalise an instrumental variable approach by constructing instruments for each of the endogenous variables (baseline tests) in model (1) above. Separate multilevel models for each endogenous explanatory variable are constructed using regressors assumed exogenous and independent of the random part of model (1). Prediction estimates of the endogenous variables are then obtained from the fixed parts of the supplementary models. These predicted values, being independent of the random part of the model of interest (1), are used as instruments.

Armed with data on the original set of regressors of model (1) and the set of instruments (being the original regressor set with endogenous variables replaced by their instruments), we can conduct an instrumental variable estimation of the fixed effect parameters in model 1 (see e.g. Bowden & Turkington (1984)). This provides us with consistent estimates of the fixed parameters but at this stage adequate estimates of their standard errors are not available.

The next stage, then, is to obtain estimates of the random part of Model (1). This is done by using MlwiN procedures to create constraints on the fixed parameters. They are forced to be equal to those calculated from the instrumental variable procedure. The resulting estimates of the random part of the model can then be used to obtain standard errors of the instrumental variable fixed effects estimates.

Mln and MlwiN macros called IV have been written to implement this procedure for obtaining consistent parameter estimates via instrumental variable estimation. They are available from the author's on request (A.Fielding@bham.ac.uk , **N.H.Spencer@herts.ac.uk**) or can be downloaded from the Birmingham web page www.bham.ac.uk/economics/staff/tony.htm. Preparatory work on the part of the user involves identifying which variables are to be endogenous and creating appropriate instruments for them. Model (1) or whatever is set up in the normal manner in MlwiN. The macros then need to have indicated to them which worksheet columns

the set of instrumental variables occupy. Following the set of simple instructions in the 'READ ME' file accompanying the macros does this.

The Application

We now use the example and data of Fielding (1998) discussed above to demonstrate the use and performance of the instrumental variable method embodied in the macros. The particular response variable used is the Mathematics Test level at Key Stage 1 standardised to have mean zero and unit variance.

The seven baseline tests available in the data (various forms of Mathematics and English tests) are inevitably highly correlated. To avoid potential multicollinearity and for heuristic purposes principal components are formed. The first principal component accounts for 60% of the variation in the data and none of the remaining components accounts for more than 8.5%. A decision was taken to use just the first principal component and this turns out to be a near equally weighted average of the seven original test scores. This variable was taken as endogenous and replaced the seven variables in model (1) above. Further, it was then modelled using a multilevel approach with intercept random effects for school and pupils with fixed effect explanatory   dummy variables for the pupils ethnicity and first language.
Excluded from use were some dummies that had been shown by Fielding (1998) to have a significant effect on pupil progress. Fielding also found that whether or not the pupil had been to a nursery school did have an impact on baseline performance but no significant effect on progress.. Thus a dummy for this variable was also used in forming the instrument.. The resultant model for the principal component of baseline scores was thus thought to provide predictions as an instrument, free of the problem of dependence on the disturbances of the original progress model of interest.

Table 1 shows estimates of the fixed parameters (and estimated standard errors)of the adapted model (1) obtained with and without the consistent instrumental variable estimation procedure (IV). It is noticeable that the influence of gender and baseline testing decreases and that of age increases (indeed almost doubles) when the consistent procedure is applied. Thus although the estimates from the two procedures are not radically different the interpretation of the results is affected by the choice of estimation method. Additionally using the instrumental variable technique gives us the added benefit of providing estimates which are consistent. Had we not used the consistent estimation procedure we would not have known the extent to which the potential problem of endogeneity was affecting the results that we were obtaining.

Table 1: Results with and without Instrumental Variable Procedures

	Without IV		With IV	
Coefficient for	**Estimate**	**Est std err**.	**Estimate**	**Est. std. err**.
Intercept	-0.0671	0.0520	0.0353	0.0611
Male gender dummy	0.102	0.0244	0.0758	0.0335
Centred age in months	0.0145	0.00379	0.0281	0.00828
Baseline Ist Principal	0.314	0.00775	0.211	0.0540

Component				

In Table 1 it is also noticeable that the estimated standard errors produced by the IV procedures are higher (though not much higher) than those produced without. A well known drawback of IV procedures is that if good instruments for the endogenous variables cannot be found , then the resulting estimates , although consistent, may be quite imprecise. In some cases standard errors can become so large so as to make results uninterpretable. This in turn has implications for good practice in data collection where modelling such as that described above intended. Sometimes information may not be collected on variables which may be important in this respect but are ignored because they are thought from theoretical and empirical considerations *a priori* to have no direct influence on a response The variable of interest. The relatively low standard errors for the IV estimators displayed in Table 1 may be noted . They were obtained because the instrument for the baseline principal component variable
was constructed from a number of useful variables that were collected alongside the baseline results. These variables had no net direct effect on the response once other controls had been included and could not be part of unmeasured variation included in the random part of the model. Had these details of nursery attendance and ethnicity/first language not been available, the estimates of the first principle component used to construct its instrument would inevitably have been much cruder and larger standard errors than those in Table 1 might well have resulted. It is also possible that ,if a much wider range of information relevant to baseline results had been available, the precision of the IV estimators might have been further improved. Experimentation with good instrument selection depends on the availability of such information. There is then a real lesson to be learned for study design. It is important when studies are planned , with ends similar to the above in view, that allowance is made for the necessity to have sufficient (and possibly relevant) background variables to use when constructing instruments.

Conclusions

The problem of inconsistency caused by the presence of endogenous variables in a multilevel model has been identified and a solution suggested by use of adaptations of instrumental variable procedures. The implementation of the consistent estimation method suggested has been made possible using the flexible macro facilities of MlwiN. These macros , available from the authors , make the implementation as a MlwiN procedure a fairly routine operation. An illustration of the method in action has been presented and the results contrasted with those produced when the problem of heterogeneity is ignored. Due to the availability of some reasonable variables with which to form instruments the IV procedure produced satisfactory efficient estimates. The importance of sound planning in decisions on data needs have been emphasised.

References

Bowden, R.J. and Turkington, D. A. (1984). *Instrumental Variables*, Cambridge, England, CUP

Fielding, A. (1998). *Why Use Arbitrary Points Scores? Ordered Categories in Models*

of Educational Progress. Discussion Paper 98-23, Department of Economics, University of Birmingham

Kiviet, J. F. (1995). On Bias, Inconsistency and Efficiency of Various Estimators in Dynamic Panel Data Models. *Journal of Econometrics*, 68, 53-78

Rice, N.. , Jones, A. and Goldstein, H. (1998). Multilevel Models where the Random Effects are Correlated with the Fixed Predictors: A Conditioned Iterative Generalised Least Squares Estimator (CIGLS). *Multilevel Modelling Newsletter*, 10, 1, 7-11

Spencer, N. H.. (1998) Consistent Parameter Estimation for Lagged Multuilevel Models, University of Hertfordshire Business School Technical Report 1 , UHBS 1998:19