Estimators and tests for unbalanced multi-way error component models with correlated effects

Giovanni S.F. Bruno
Department of Economics, Bocconi University

"5° Incontro degli utenti di Stata"
Milano, 20-21 ottobre 2008

1 Structure of the presentation

- Motivations
- Related literature
- The multiway Error Component Model (ECM)
- Results
- Conclusions

2 Motivations

- New a) tests of correlated effects and b) estimators for the (possibly) unbalanced multiway ECM.
- New algebraic results, useful for computations.

3 Related literature

- Tests for correlated effects: Hausman (1978), Mundlak (1978), Hausman and Taylor (1982), Kang (1985), Arellano (1993), Ahn and Low (1996), Wooldridge (2002), Krishnakumar (2006).
- Estimators: Kaptein and Wansbeek (1989), Davis (2002).
- Algebra for the multiway ECM: Davis (2002).

4 The multiway ECM

4.1 Notation for column-wise partitioned matrices

Given a column-wise partitioned matrix $A = \begin{pmatrix} A_1 & A_2 & \cdots & A_m \end{pmatrix}$, define $\mathfrak{D}(A)$ as the set of all column-wise partitioned matrices formed by any number $1 \leq k \leq m$ of distinct blocks of A, taken in the same order as in A. For example, if $A = \begin{pmatrix} A_1 & A_2 & A_3 & A_4 \end{pmatrix}$, then $\begin{pmatrix} A_1 & A_3 & A_4 \end{pmatrix} \in \mathfrak{D}(A)$. $A \in \mathfrak{D}(A)$ and the size of $\mathfrak{D}(A)$ is $\sum_{g=1}^{m} \binom{m}{g}$.

4.2 Projection matrices

Given an arbitrary matrix A, A^- denotes a generalized inverse of A. $P_{[A]} = A(A'A)^- A'$ indicates the projection matrix onto the space spanned by the columns of A. $Q_{[A]} = I - P_{[A]}$

4.3 The Model

I focus on the general multi-way ECM with generic number of levels $m+1$

$$y = W\delta + \Gamma u \tag{1}$$

where

$$\begin{aligned} W &= \begin{pmatrix} X & \Delta Z \end{pmatrix} \\ \Gamma &= \begin{pmatrix} I_n & \Delta \end{pmatrix} \text{ and } \Delta = \begin{pmatrix} \Delta_1 & \Delta_2 & \cdots & \Delta_m \end{pmatrix} \\ \delta &= \begin{pmatrix} \beta' & \lambda' \end{pmatrix}' \\ u &= \begin{pmatrix} u_0' & u_1' & \cdots & u_m' \end{pmatrix}' \end{aligned}$$

and

- Δ_i denotes the $(n \times g_i)$ matrix of dummy variables indicating the groups at the level $i = 1, ...m$
- u_i denotes the error component vector of dimension $(g_i \times 1)$;
- u_0 stands for the idiosyncratic error component vector of dimension $(n \times 1)$

The following identification assumptions holds throughout.

A.1 Both X and ΔZ are of full-column rank (f.c.r.).

The following assumption characterises the columns of X as the regressors with idiosyncratic (observation specific) variation.

A.2 No linear combination of the columns of X lies in the subspace spanned by the columns of Δ.

A.1 and **A.2** together imply that the regressor matrix W is of f.c.r.

A.3 ECM variance-covariance matrix of the composite error Γu (Kaptein and Wansbeek, 1987; Davis, 2002)

$$\Sigma = \sigma_0^2 I_n + \sigma_1^2 \Delta_1 \Delta_1' + ... + \sigma_m^2 \Delta_m \Delta_m' \tag{2}$$

Convenient nonsingular transformations of Δ and Γ are defined below.

Definition 1 *Let* $\widetilde{\Delta}_i = \frac{\sigma_i}{\sigma_0}\Delta_i$ *for all* $i = 1, ..., m$. *Then, let* $\widetilde{\Delta} = \begin{pmatrix} \widetilde{\Delta}_1 & \cdots & \widetilde{\Delta}_m \end{pmatrix}$ *and* $\widetilde{\Gamma} = \begin{pmatrix} I_n & \widetilde{\Delta} \end{pmatrix}$.

It follows that

$$\Sigma = \sigma_0^2 \left(I_n + \widetilde{\Delta}_1 \widetilde{\Delta}_1' + ... + \widetilde{\Delta}_m \widetilde{\Delta}_m' \right)$$

5 Algebraic results

Definition 2 *Given a real matrix A, define the operator $V_{[A]}$ as $V_{[A]} = (AA')^{-1}$.*

The importance of $V_{[\cdot]}$ hinges upon the following

$$V_{[\widetilde{\Gamma}]} = \sigma_0^2 \Sigma^{-1}. \tag{3}$$

$V_{[\cdot]}$ is well defined for any column-wise partitioned matrix A of the form $A = \begin{pmatrix} I & B \end{pmatrix}$ as $AA' = I + BB'$ is positive definite.

The following Lemma (Davis, 2002) is useful to compute $V_{[\widetilde{\Gamma}]}$

Lemma 3 *Let $C = \begin{pmatrix} I & D_1 & D_2 \end{pmatrix}$. Then,*

$$V_{[C]} = V_{[I\ D_2]} - V_{[I\ D_2]} D_1 \left[I + D_1' V_{[I\ D_2]} D_1 \right]^{-1} D_1' V_{[I\ D_2]}$$

and

$$V_{[I\ D_2]} = I - D_2 \left[I + D_2' D_2 \right]^{-1} D_2'.$$

The following extension to Davis (2002) (and to Wansbeek and Kapteyn (1989)) expands the set of possible representations for $V_{[\widetilde{\Gamma}]}$.

Lemma 4 *Given the column-wise partitioned real matrix B, let $B_1 \in \mathfrak{D}(B)$, $A = \begin{pmatrix} I & B \end{pmatrix}$ and $r \equiv rank(B_1)$. Then, there exists a mapping $m : \mathfrak{L}(B_1) \rightarrow \mathfrak{M}_r$ defined as*

$$m(B_1^*) = \begin{cases} (B_1^{*\prime} B_1^*)^{-1} B_1^{*\prime} B_1 B_1' B_1^* (B_1^{*\prime} B_1^*)^{-1} & \text{if } B_1^* \text{ has f.c.r.} \\ I_r & \text{else} \end{cases}$$

and such that

$$V_{[A]} = V_{[A \backslash B_1]} - V_{[A \backslash B_1]} B_1^* \left[m^{-1}(B_1^*) + B_1^{*\prime} V_{[A \backslash B_1]} B_1^* \right]^{-1} B_1^{*\prime} V_{[A \backslash B_1]} \tag{4}$$

for all $B_1^* \in \mathfrak{L}(B_1)$; where $\mathfrak{L}(B_1)$ is the set containing B_1 and all the submatrices of B_1 having f.c.r. and $\mathfrak{M}_r$ is the collection of all $r \times r$ symmetric positive definite matrices.

Lemma 3 emerges as a corollary of Lemma 4.

A convenient operator is defined.

Given a positive definite symmetric matrix Ω and any matrix A define $P_{[\Omega,A]}$ as

$$P_{[\Omega,A]} = A\left(A'\Omega A\right)^{-} A'\Omega \tag{5}$$

and $Q_{[\Omega,A]}$ as

$$Q_{[\Omega,A]} = I - P_{[\Omega,A]}$$

Specific properties of $P_{[\Omega,A]}$ may emerge depending on A and Ω. The following results establishes two important properties for $P_{\left[V_{[\tilde{\Delta}]},\Delta_{(k)}\right]}$.

Theorem 5 $P_{\left[V_{[\tilde{\Gamma}]},\Delta_{(k)}\right]} = P_{\left[V_{\left[\tilde{\Gamma}\setminus\tilde{\Delta}_{(k)}\right]},\Delta_{(k)}\right]}$ *for any* $\Delta_{(k)} \in \mathfrak{D}\left(\Delta\right)$.

Theorem 6 $V_{[\tilde{\Gamma}]}Q_{\left[V_{[\tilde{\Gamma}]},\Delta_{(k)}\right]} = V_{\left[\tilde{\Gamma}\setminus\tilde{\Delta}_{(k)}\right]}Q_{\left[V_{\left[\tilde{\Gamma}\setminus\tilde{\Delta}_{(k)}\right]},\Delta_{(k)}\right]}$ *for any* $\Delta_{(k)} \in \mathfrak{D}\left(\Delta\right)$.

6 Estimators and tests

6.1 Efficient GLS estimators

Under assumptions **A.1-A.3**, if all effects are not correlated to the regressors, that is if

$$E(u|W) = 0,$$

then the Gauss-Marcov estimator for β and λ is the *Multi-way GLS*

$$d^{GLS} = \begin{pmatrix} b^{GLS} \\ l^{GLS} \end{pmatrix} = \left(W'V_{[\widetilde{\Gamma}]}W\right)^{-1} W'V_{[\widetilde{\Gamma}]}y. \tag{6}$$

The formula for b^{GLS} is the following

$$b^{GLS} = \left(X'V_{[\widetilde{\Gamma}]}Q_{\left[V_{[\widetilde{\Gamma}]},\Delta Z\right]}X\right)^{-1} X'V_{[\widetilde{\Gamma}]}Q_{\left[V_{[\widetilde{\Gamma}]},\Delta Z\right]}y \tag{7}$$

The *Multi-way Within* estimator for β is the following

$$b^{within} = \left(X'Q_{[\Delta]}X\right)^{-1} X'Q_{[\Delta]}y. \tag{8}$$

It is a robust estimator in that it leaves the correlation between regressors and all error components unrestricted. A more general class of efficient estimators encompassing d^{GLS} and b^{within} as particular cases is derived

Theorem 7 *Assume* ***A.1-A.3*** *and let* $\Delta_{(k)} \in \mathfrak{D}(\Delta)$. *Then, the efficient multi-way GLS estimator for* β *and* λ *in the presence of (possibly) correlated effects at the levels* $\Delta_{(k)}$, $d^{GLS|\Delta_{(k)}}$, *is*

$$\begin{aligned} d^{GLS|\Delta_{(k)}} &= \begin{pmatrix} b^{GLS|\Delta_{(k)}} \\ l^{GLS|\Delta_{(k)}} \end{pmatrix} \\ &= \left(W'HQ_{\left[H,\Delta_{(k)}\right]}W\right)^{-} W'HQ_{\left[H,\Delta_{(k)}\right]}y. \end{aligned} \tag{9}$$

with

$$b^{GLS|\Delta_{(k)}} = (X'MX)^{-1} X'My, \tag{10}$$

where $H \equiv V_{\left[\widetilde{\Gamma}\backslash\widetilde{\Delta}_{(k)}\right]}$ *and* $M = H\left(Q_{\left[H,\Delta_{(k)}\right]} - P_{\left[H,Q_{\left[H,\Delta_{(k)}\right]}\Delta Z\right]}\right)$.

6.2 Between estimators

The *Multi-way Between estimator*, considering the variation between all groups in Δ, is defined as

$$\widetilde{d}^{B} = \left(W' V_{\left[\widetilde{\Gamma}\right]} P_{[\Delta]} W\right)^{-1} W' V_{\left[\widetilde{\Gamma}\right]} P_{[\Delta]} y. \tag{11}$$

The following general formula for the between estimator of β is suggested, which is useful in the context of specification tests

$$\widetilde{b}^{B\left(\Delta_{(k)}\right)} = \left(X' V_{\left[\widetilde{\Gamma}\right]} P_{\left[V_{\left[\widetilde{\Gamma}\right]}, Q_{\left[V_{\left[\widetilde{\Gamma}\right]}, \Delta Z\right]} \Delta_{(k)}\right]} X \right)^{-1} X' V_{\left[\widetilde{\Gamma}\right]} P_{\left[V_{\left[\widetilde{\Gamma}\right]}, Q_{\left[V_{\left[\widetilde{\Gamma}\right]}, \Delta Z\right]} \Delta_{(k)}\right]} y. \tag{12}$$

It generalizes the extended between estimator derived in Krishnakumar (2006) to an unbalanced multilevel setting with generic non-idiosyncratic variables that do not lie necessarily onto the space spanned by the correlated effects. One can think of $\widetilde{b}^{B\left(\Delta_{(k)}\right)}$ as an estimator that exploits only the residual variation between the groups in $\Delta_{(k)}$ once the variation in ΔZ has been partialled out (in the metric $V_{\left[\widetilde{\Delta}\right]}$).

6.3 Efficient GLS estimators as weighted averages

Theorem 8 *For all* $\Delta_{(k)} \in \mathfrak{D}\left(\Delta\right)$

$$b^{GLS} = F b^{GLS|\Delta_{(K)}} + \left(I - F\right) \widetilde{b}^{B\left(\Delta_{(k)}\right)}$$

Theorem 9 *Let* $\Delta_{(\cdot)} \in \mathfrak{D}\left(\Delta\right)$ *and* $\Delta_{(k)} \in \mathfrak{D}\left(\Delta|\Delta_{(\cdot)}\right)$ *then*

$$b^{GLS|\Delta_{(\cdot)}} = F b^{GLS|\Delta_{(K)}} + G \widetilde{b}^{B\left(\Delta_{(k)}\right)} - H \widetilde{b}^{B\left(\Delta_{(\cdot)}\right)}$$

with $F + G + H = I$

6.4 Tests for correlated effects

Borrowing the same terminology as Kang's (1985), the following definitions hold.

Definition 10 *For some level $i = 1, ..., m$, the unobserved effect u_i is said uncorrelated if $E(u_i|W) = 0$.*

Definition 11 *For some level $i = 1, ..., m$, the unobserved effect u_i is said (possibly) correlated if $E(u_i|W)$ is left unrestricted.*

In a multi-level framework the number of possible specifications for the unobserved effects, h_m, increases rapidly with the number of error components m. For example, Kang (1985) focussing on the two-level model considers $h_2 = 1 + \binom{2}{1}2 = 5$ possible specifications for the error components and consequently 5 specification tests. These are reported in Table 1.

Table 1: Specification tests in the two-level model

Test	H_o	Given:
1	u_2 uncorrelated	u_1 correlated
2	u_2 uncorrelated	u_1 uncorrelated
3	u_1 uncorrelated	u_2 correlated
4	u_1 uncorrelated	u_2 uncorrelated
5	u_1 and u_2 uncorrelated	

If only m increases to 3, the number of specification tests increases to $h_3 = 19$ $(1 + \binom{3}{2}2 + 3\left[2 + \binom{2}{1}\right] = 19)$. The specification tests are spelled out in Table 2

Table 2: Specification tests in the three-way model

Test	H_o	Given:
1	u_3 uncorrelated	u_1 and u_2 correlated
2	u_2 uncorrelated	u_1 and u_3 correlated
3	u_1 uncorrelated	u_2 and u_3 correlated
4	u_3 and u_2 uncorrelated	u_1 correlated
5	u_3 and u_1 uncorrelated	u_2 correlated
6	u_1 and u_2 uncorrelated	u_3 correlated
7	u_3 uncorrelated	u_1 uncorrelated and u_2 correlated
8	u_3 uncorrelated	u_2 uncorrelated and u_1 correlated
9	u_2 uncorrelated	u_1 uncorrelated and u_3 correlated
10	u_2 uncorrelated	u_3 uncorrelated and u_1 correlated
11	u_1 uncorrelated	u_2 uncorrelated and u_3 correlated
12	u_1 uncorrelated	u_3 uncorrelated and u_2 correlated
13	u_3 uncorrelated	u_1 and u_2 uncorrelated
14	u_2 uncorrelated	u_1 and u_3 uncorrelated
15	u_1 uncorrelated	u_2 and u_3 uncorrelated
16	u_3 and u_2 uncorrelated	u_1 uncorrelated
17	u_3 and u_1 uncorrelated	u_2 uncorrelated
18	u_1 and u_2 uncorrelated	u_3 uncorrelated
19	u_1, u_2 and u_3 uncorrelated	

In general, with m error components the number h_m of tests is

$$\begin{aligned}
h_m &= 1+\binom{m}{m-1}2+...+\binom{m}{2}\left[2+\binom{m-2}{m-3}+\right. \\
&\quad ...+\left.\binom{m-2}{2}+\binom{m-2}{1}\right]+m\left[2+\binom{m-1}{m-2}+\right. \\
&\quad ...+\left.\binom{m-1}{2}+\binom{m-1}{1}\right] \\
&= 1+\binom{m}{m-1}2+\sum_{g=1}^{m-2}\binom{m}{m-1-g}\left(2+\sum_{h=1}^{g}\binom{g+1}{h}\right)
\end{aligned}$$

Fortunately, the notation used in this paper is general enough to deal with any number of error components. Indeed, as large as h_m may be, the specification tests can always be classified according to the following four-type partition.

1. Test that the effects at the levels $\Delta_{(\cdot)} \in \mathfrak{D}(\Delta)$ are uncorrelated given that the effects at all other levels $\Delta^c_{(\cdot)}$ are uncorrelated. There are $\sum_{g=1}^{m-1}\binom{m}{m-g}$ Hausman tests based on the differences $q_1\left(\Delta_{(\cdot)}\right) = b^{GLS} - b^{GLS|\Delta_{(\cdot)}}$ over all $\Delta_{(\cdot)} \in \mathfrak{D}(\Delta)$. If $m=2$, these are Test 2 and Test 4 of Table 1. If $m=3$ these are Test 13 to 18 in Table 2.

2. Test that the effects at the levels $\Delta_{(\cdot)} \in \mathfrak{D}(\Delta)$ are uncorrelated, leaving the effects at all other levels, $\Delta\backslash\Delta_{(\cdot)}$, possibly correlated. There are $\sum_{g=1}^{m-1}\binom{m}{m-g}$ Hausman tests based on the differences $q_2\left(\Delta_{(\cdot)}\right) = b^{GLS|\Delta\backslash\Delta_{(\cdot)}} - b^{within}$ over all $\Delta_{(\cdot)} \in \mathfrak{D}(\Delta)$. If $m=2$ these are Test 1 and Test 3 of Table 1. If $m=3$, these are Test 1 to 6 of Table 2.

3. Test that the effects at the levels $\Delta_{(k)} \in \mathfrak{D}(\Delta)$ are uncorrelated, maintaining a mixed specification for the effects at all other levels,$\Delta\backslash\Delta_{(k)}$; that is assume that the effects at the levels $\Delta_{(\cdot)} \in \mathfrak{D}\left(\Delta\backslash\Delta_{(k)}\right)$ are uncorrelated and leave the effects at the remaining levels $\Delta\backslash\Delta_{(k)}\backslash\Delta_{(\cdot)}$ possibly correlated, $k=1,...,m-2$. There are

$$\sum_{g=1}^{m-2}\binom{m}{m-1-g}\sum_{h=1}^{g}\binom{g+1}{h}$$

Hausman tests based on the differences $q_3\left(\Delta_{(\cdot)}, \Delta_{(k)}\right) = b^{GLS|\Delta\backslash\Delta_{(k)}\backslash\Delta_{(\cdot)}} - b^{GLS|\Delta\backslash\Delta_{(\cdot)}}$ over all $\Delta_{(\cdot)} \in \mathfrak{D}\left(\Delta\backslash\Delta_{(k)}\right)$. If $m = 2$, there are no such tests. If $m = 3$ these are Test 7 to 12 of Table 2.

4. Test that the effects at all levels are uncorrelated. Regardless the number of levels in the data, there is 1 Hausman test based on the difference $q_4 = b^{GLS} - b^{within}$. This is Test 5 in Table 1 and Tests 19 in Table 2.

Remark 12 *Particular tests of type 4 have been examined in the ECM literature, notably Hausman and Taylor (1982), Arellano (1993) and Ahn and Low (1996) for $m = 1$ and Kang (1987) for $m = 2$. Particular tests of type 1 and 2 have been examined by Kang (1987) for $m = 2$. Conversely, tests of type 3 have never been considered, since they emerge only for $m \geq 3$. Given that efficient GLS can be obtained as weighted averages of other estimators, identical tests can be derived using differences that involve the between estimators.*

7 Conclusion

What's left to do?

- Mata implementation
- Regression based tests a la Mundlak

(MAIN) REFERENCES

DAVIS, P. (2002) "Estimating multi-way error component models with unbalanced data structures" *Journal of Econometrics*, 106, 67-95.

KANG, S. (1985) "A note on the equivalence of specification tests in the two factor multivariate variance components model" *Journal of Econometrics*, 106, 67-95.

HAUSMAN, J.A.and W.E. TAYLOR, (1981) "Panel Data and Unobservable Individual Effects" *Econometrica*, 49, 1377-1398.

HUSSAIN, A. (1969) "A mixed model for regressions" *Biometrika*, 56, 327-336

MUNDLAK, Y. (1978) "On the Pooling of Time Series and Cross Section Data" *Econometrica*, 46, 69-85.

WANSBEEK, T. and A. KAPTEYN (1989) "Estimation of the error component model with incomplete panels", *Journal of Econometrics*, 41, 341-361.