Panel Econometrics

Paulo M. M. Rodrigues

February 2022

Today's outline

Dynamic models II: More GMM

1. AB: Exogenous regressors correlated with the individual effect
2. AB - Some regressors are uncorrelated with the individual effect
3. The Blundell and Bond estimator

Outline

Introduction

Now let us add additional regressors:

$$y_{it} = \alpha y_{i,t-1} + \mathbf{x}_{it}\boldsymbol{\beta} + v_{it}, \qquad v_{it} = c_i + u_{it}.$$

In this part we assume that the regressors are (potentially) correlated with the individual effect. We distinguish two cases:

- The regressors are *strictly exogenous*.
- The regressors are *predetermined*.

In the next section we assume that some elements of $\mathbf{x}_{it}$ are uncorrelated with the individual effect which gives rise to Hausman-Taylor type approaches. Again, we will distinguish between strictly exogenous and predetermined regressors.

Strictly exogenous regressors: Moment conditions

Strict exogeneity implies

$$\mathrm{E}(\mathbf{x}_{it}' u_{is}) = \mathbf{0} \quad \Rightarrow \quad \mathrm{E}(\mathbf{x}_{it}' \Delta u_{is}) = \mathbf{0} \qquad \forall t = 1, \ldots, T.$$

But the $\mathbf{x}_{it}$ are correlated with c_i such that

$$\mathrm{E}(\mathbf{x}_{it}' v_{is}) = \mathrm{E}(\mathbf{x}_{it}' c_i) \neq \mathbf{0}.$$

Hence, the regressors are valid instruments for the first differenced equation

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \Delta \mathbf{x}_{it} \boldsymbol{\beta} + \Delta u_{it}.$$

but not for the level equation

$$y_{it} = \alpha y_{i,t-1} + \mathbf{x}_{it} \boldsymbol{\beta} + v_{it}.$$

Estimate the parameters from the first differenced equation

We use all moment conditions available for periods 3 to T:

$$\mathrm{E}(\mathbf{x}_{i1}'\Delta u_{it}) = \ldots = \mathrm{E}(\mathbf{x}_{iT}'\Delta u_{it}) = \mathbf{0} \quad \text{for all } t = 3, \ldots, T.$$

Defining the $1 \times \tau K$ vector $\mathbf{x}_{i,1:\tau} = [\mathbf{x}_{i1}, \ldots, \mathbf{x}_{i\tau}]$, they can be written as

$$\mathrm{E}(\mathbf{x}_{i,1:T}'\Delta u_{it}) = \mathbf{0} \quad \text{for all } t = 3, \ldots, T.$$

We add the previous moment conditions

$$\mathrm{E}[y_{i1}\Delta u_{it}] = \cdots = \mathrm{E}[y_{i,t-2}\Delta u_{it}] = 0, \qquad t = 3, \ldots, T.$$

Defining the $1 \times \tau$ vector $\mathbf{y}_{i,1:\tau} = [y_{i1}, \ldots, y_{i\tau}]$, they can be written as

$$\mathrm{E}[\mathbf{y}_{i,1:t-2}'\Delta u_{it}] = \mathbf{0}, \qquad t = 3, \ldots, T.$$

Strictly exogenous regressors: Instrument matrix

Putting the moment conditions together yields the instrument matrix

$$\mathbf{W}_i = \begin{bmatrix} [\mathbf{y}_{i,1:1}, \mathbf{x}_{i,1:T}] & & & 0 \\ & [\mathbf{y}_{i,1:2}, \mathbf{x}_{i,1:T}] & & \\ & & \ddots & \\ 0 & & & [\mathbf{y}_{i,1:T-2}, \mathbf{x}_{i,1:T}] \end{bmatrix}$$

The moment conditions are expressed as

$$\mathrm{E}(\mathbf{W}_i' \Delta \mathbf{u}_i) = 0.$$

Strictly exogenous regressors: Number of instruments

We have $(T-2)(T-1)/2 + T(T-2)K$ moment conditions:

- $(T-2)(T-1)/2$ from the lags of the endogenous variable,
- $T(T-2)K$ from the strictly exogenous regressors.

This is a lot of instruments for $K+1$ parameters to be estimated:

- $T=3$: $L = 1+3K$ instruments
- $T=4$: $L = 3+8K$ instruments
- $T=10$: $L = 36+80K$ instruments.

Consider a not-so-unusual example:

- We want to estimate a structural model with 5 strictly exogenous regressors and one lagged endogenous regressor (6 parameters).
- We have a panel with 10 time periods (not that much nowadays).
- Then we may use $L = 436$ instruments (72.7 per parameter!).
- Can we be sure that they are all relevant?

Strictly exogenous regressors: GMM estimation

We want to estimate the equation

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \Delta \mathbf{x}_{it} \boldsymbol{\beta} + \Delta u_{it} = \mathbf{z}_{it} \boldsymbol{\delta} + \Delta u_{it}, \quad t = 1, \ldots, T,$$

where $\mathbf{z}_{it} = [\Delta y_{i,t-1}, \Delta \mathbf{x}_{it}]$ and $\boldsymbol{\delta} = [\alpha, \boldsymbol{\beta}']'$.

Let us define the full regressor matrix

- for individual i: $\mathbf{Z}_i = [\Delta \mathbf{y}_{i,-1}, \Delta \mathbf{X}_i]$
- for all N individuals: $\mathbf{Z} = [\Delta \mathbf{y}_{-1}, \Delta \mathbf{X}]$.

One-step and two-step GMM estimation now works as in the AR(1) case.

Strictly exogenous regressors: One-step GMM estimator

Under equivalent homoskedasticity assumptions as in the AR(1) case, the optimal (one-step) weighting matrix is

$$\boldsymbol{\Xi}_1 = \hat{\boldsymbol{\Lambda}}_1^{-1} = \left[\mathrm{E}(\mathbf{W}_i'\Delta\mathbf{u}_i\Delta\mathbf{u}_i'\mathbf{W}_i)/\sigma_u^2\right]^{-1} = \left[\mathrm{E}(\mathbf{W}_i'\mathbf{G}\mathbf{W}_i)\right]^{-1}$$

which can be consistently estimated as

$$\hat{\boldsymbol{\Xi}}_1 = \left[N^{-1}\sum_{i=1}^{N}\mathbf{W}_i'\mathbf{G}\mathbf{W}_i\right]^{-1} = \left[N^{-1}\mathbf{W}'(\mathbf{I}_N \otimes \mathbf{G})\mathbf{W}\right]^{-1}.$$

The one-step GMM estimator is

$$\hat{\boldsymbol{\delta}}_{AB,1} = \begin{pmatrix}\hat{\alpha}_{AB,1}\\ \hat{\boldsymbol{\beta}}_{AB,1}\end{pmatrix} = \left[\mathbf{Z}'\mathbf{W}\hat{\boldsymbol{\Xi}}_1\mathbf{W}'\mathbf{Z}\right]^{-1}\mathbf{Z}'\mathbf{W}\hat{\boldsymbol{\Xi}}_1\mathbf{W}'\Delta\mathbf{y}.$$

Strictly exogenous regressors: Two-step GMM estimator

Based on first-step residuals, estimate an unconstrained weighting matrix:

$$\hat{\boldsymbol{\Xi}}_2 = \hat{\boldsymbol{\Lambda}}_2^{-1} = \left[N^{-1} \sum_{i=1}^{N} \mathbf{W}_i' \Delta \hat{\mathbf{u}}_i \Delta \hat{\mathbf{u}}_i' \mathbf{W}_i \right]^{-1}.$$

Two-step GMM estimator:

$$\hat{\boldsymbol{\delta}}_{AB,2} = \begin{pmatrix} \hat{\alpha}_{AB,2} \\ \hat{\boldsymbol{\beta}}_{AB,2} \end{pmatrix} = \left[\mathbf{Z}' \mathbf{W} \hat{\boldsymbol{\Xi}}_2 \mathbf{W}' \mathbf{Z} \right]^{-1} \mathbf{Z}' \mathbf{W} \hat{\boldsymbol{\Xi}}_2 \mathbf{W}' \Delta \mathbf{y}.$$

Predetermined regressors: Moment conditions

Again, suppose that all the $\mathbf{x}_{it}$ are correlated with c_i.

But now assume the $\mathbf{x}_{it}$ are *predetermined* rather than strictly exogenous:

$$\mathrm{E}(\mathbf{x}'_{i,t-s} u_{it}) \begin{cases} = \mathbf{0} & \text{for } s = 0, 1, 2, \ldots \\ \neq \mathbf{0} & \text{for } s = -1, -2, \ldots. \end{cases}$$

Hence,

$$\mathrm{E}(\mathbf{x}'_{i,t-s} \Delta u_{it}) \begin{cases} = \mathbf{0} & \text{for } s = 1, 2, \ldots \\ \neq \mathbf{0} & \text{for } s = 0, -1, -2, \ldots. \end{cases}$$

Then $\mathbf{x}_{i,1:t-1} = [\mathbf{x}_{i1}, \ldots, \mathbf{x}_{i,t-1}]$ are valid instruments for the differenced equation in period t:

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \Delta \mathbf{x}_{it} \boldsymbol{\beta} + \Delta u_{it}.$$

Predetermined regressors: Illustration

- Consider period $t = 3$, the first period we observe the relation:

$$y_{i3} - y_{i2} = \alpha \left(y_{i2} - y_{i1} \right) + \left(\mathbf{x}_{i3} - \mathbf{x}_{i2} \right) \boldsymbol{\beta} + \left(u_{i3} - u_{i2} \right).$$

Here $\mathbf{x}_{i1}$ and $\mathbf{x}_{i2}$ are valid instruments, since both are uncorrelated with $u_{i3} - u_{i2}$.

- Consider period $t = 4$, the second period we observe the relation

$$y_{i4} - y_{i3} = \alpha \left(y_{i3} - y_{i2} \right) + \left(\mathbf{x}_{i4} - \mathbf{x}_{i3} \right) \boldsymbol{\beta} + \left(u_{i4} - u_{i3} \right).$$

Now $\mathbf{x}_{i1}$, $\mathbf{x}_{i2}$ and $\mathbf{x}_{i3}$ are uncorrelated with $u_{i4} - u_{i3}$ and thus valid instruments.

- Consider period $t = 5$, the third period we observe the relation

$$y_{i5} - y_{i4} = \alpha \left(y_{i4} - y_{i3} \right) + \left(\mathbf{x}_{i5} - \mathbf{x}_{i4} \right) \boldsymbol{\beta} + \left(u_{i5} - u_{i4} \right).$$

Now $\mathbf{x}_{i1}, \ldots, \mathbf{x}_{i4}$ are uncorrelated with $u_{i5} - u_{i4}$ and thus valid.

Predetermined regressors: Instrument matrix

Based on the moment conditions based on the predetermined regressors

$$\mathrm{E}(\mathbf{x}'_{i,1:t-1}\Delta u_{it}) = \mathbf{0} \qquad \text{for all } t = 3, \ldots, T.$$

and on those based on the lagged endogenous variable

$$\mathrm{E}[\mathbf{y}'_{i,1:t-2}\Delta u_{it}] = \mathbf{0}, \qquad t = 3, \ldots, T.$$

we obtain the instrument matrix

$$\mathbf{W}_i = \begin{bmatrix} [\mathbf{y}_{i,1:1}, \mathbf{x}_{i,1:2}] & & & 0 \\ & [\mathbf{y}_{i,1:2}, \mathbf{x}_{i,1:3}] & & \\ & & \ddots & \\ 0 & & & [\mathbf{y}_{i,1:T-2}, \mathbf{x}_{i,1:T-1}] \end{bmatrix}$$

Now the moment conditions can once again be expressed as

$$\mathrm{E}(\mathbf{W}'_i\Delta\mathbf{u}_i) = 0.$$

Predetermined regressors: Number of instruments

We now have $(T-2)(T-1)/2 + (T+1)(T-2)K/2$ moment conditions:

- $(T-2)(T-1)/2$ from the lags of the endogenous variable,
- $(T+1)(T-2)K/2$ from the predetermined regressors.

Again, this is a lot for $K+1$ parameters:

- $T=3$: $L = 1 + 2K$ instruments
- $T=4$: $L = 3 + 5K$ instruments
- $T=10$: $L = 36 + 44K$ instruments.

Consider the previous example:

- We want to estimate a structural model with 5 strictly exogenous regressors and one lagged endogenous regressor (6 parameters).
- We have a panel with 10 time periods.
- Then we may use $L = 256$ instruments (42.7 per parameter!).

GMM estimation

One-step and two-step estimators are obtained as described above for the case of strictly exogenous regressors.

Only the instrument matrices differ but the estimation approach is the same.

In empirical studies, a combination of *predetermined* and *strictly exogenous* variables may be appropriate. The researcher has to decide for each variable which assumption fits best.

Potential numerical problems

- The weighting matrix is of size $L \times L$ and may thus be huge.
- In our previous example ($\mathbf{x}_{it}$ strictly exogenous, $K = 5$, $T = 10$) we have to estimate a 436×436 matrix with $L(L+1)/2 = 95266$ distinct elements (due to symmetry).
- Estimating so many elements may induce large variance and thus imprecise estimates.
- If N is small, it may even become infeasible to invert $\hat{\mathbf{\Lambda}}$.
- This is the reason why some researchers "collapse" the moment conditions to stop their "proliferation".
  We will discuss this option later.

Outline

Predetermined regressors: Additional moment conditions

The assumption that *all* the $\mathbf{x}_{it}$ are correlated with c_i might be too cautious.

- As for Hausman-Taylor, we may rather assume that some regressors are uncorrelated with the individual effects.
- This yields additional moment restrictions.

To see this, separate

$$\mathbf{x}_{it} = [\mathbf{x}_{1,it}, \mathbf{x}_{2,it}]$$

where the K_1 regressors $\mathbf{x}_{1,it}$ are uncorrelated with c_i, while the K_2 regressors $\mathbf{x}_{2,it}$ are correlated with c_i.

If the $\mathbf{x}_{it}$ are predetermined, the additional TK_1 moment restrictions are

$$\mathrm{E}[\mathbf{x}'_{1,i1} u_{i2}] = \mathbf{0} \qquad \text{and} \qquad \mathrm{E}[\mathbf{x}'_{1,it} u_{it}] = \mathbf{0} \quad \text{for } t = 2, \ldots, T.$$

All additional linear restrictions from the level equations are redundant given those already exploited from the first-differenced equations.

Predetermined regressors: All moment conditions

- First, we have the $(T-2)(T-1)/2$ moment conditions obtained from the lagged endogenous variable and applied to the differenced equation:

$$\mathrm{E}[y_{is}\Delta u_{it}] = 0 \qquad \text{for } t = 3, \ldots, T \text{ and } s \le t-2.$$

- In addition, we have the $K(T-2)(T+1)/2$ moment conditions obtained from $\mathbf{x}_{it}$ and applied to the differenced equation:

$$\mathrm{E}[\mathbf{x}_{is}'\Delta u_{it}] = \mathbf{0} \qquad \text{for } s \le t-1.$$

- Finally, we have the $K_1 T$ moment conditions obtained from $\mathbf{x}_{1,it}$ and applied to the level equation:

$$\mathrm{E}[\mathbf{x}_{1,i1}' u_{i2}] = \mathbf{0} \qquad \text{and} \qquad \mathrm{E}[\mathbf{x}_{1,it}' u_{it}] = \mathbf{0} \quad \text{for } t = 2, \ldots, T.$$

  Since $\mathbf{x}_{1,it}$ is uncorrelated with c_i, this implies

$$\mathrm{E}[\mathbf{x}_{1,i1}' v_{i2}] = \mathbf{0} \qquad \text{and} \qquad \mathrm{E}[\mathbf{x}_{1,it}' v_{it}] = \mathbf{0} \quad \text{for } t = 2, \ldots, T.$$

System of level and differenced equations

We now have moment conditions with respect to the level equation

$$y_{it} = \alpha y_{i,t-1} + \mathbf{x}_{it}\boldsymbol{\beta} + v_{it}$$

and with respect to the differenced equation

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \Delta \mathbf{x}_{it}\boldsymbol{\beta} + \Delta u_{it}.$$

To jointly use all moment conditions, stack the two equations

$$\begin{pmatrix} \Delta y_{i3} \\ \vdots \\ \Delta y_{iT} \\ y_{i2} \\ \vdots \\ y_{iT} \end{pmatrix} = \begin{pmatrix} \Delta y_{i2} \\ \vdots \\ \Delta y_{i,T-1} \\ y_{i1} \\ \vdots \\ y_{i,T-1} \end{pmatrix} \alpha + \begin{pmatrix} \Delta \mathbf{x}_{i3} \\ \vdots \\ \Delta \mathbf{x}_{iT} \\ \mathbf{x}_{i2} \\ \vdots \\ \mathbf{x}_{iT} \end{pmatrix} \boldsymbol{\beta} + \begin{pmatrix} \Delta u_{i3} \\ \vdots \\ \Delta u_{iT} \\ v_{i2} \\ \vdots \\ v_{iT} \end{pmatrix}$$

or compactly

$$\mathbf{y}_i^+ = \mathbf{y}_{i,-1}^+ \alpha + \mathbf{X}_i^+ \boldsymbol{\beta} + \mathbf{v}_i^+.$$

Predetermined regressors: Instrument matrices

Collecting all instruments for the differenced equation yields the known instrument matrix

$$\mathbf{W}_i = \begin{bmatrix} [\mathbf{y}_{i,1:1}, \mathbf{x}_{i,1:2}] & & & 0 \\ & [\mathbf{y}_{i,1:2}, \mathbf{x}_{i,1:3}] & & \\ & & \ddots & \\ 0 & & & [\mathbf{y}_{i,1:T-2}, \mathbf{x}_{i,1:T-1}] \end{bmatrix}$$

Collecting the instruments for the level equation yields the instrument matrix

$$\overline{\mathbf{W}}_i = \begin{bmatrix} [\mathbf{x}_{1,i1}, \mathbf{x}_{1,i2}] & & & 0 \\ & \mathbf{x}_{1,i3} & & \\ & & \ddots & \\ 0 & & & \mathbf{x}_{i,1T} \end{bmatrix}$$

Putting the instruments together yields

$$\mathbf{W}_i^{+} = \begin{bmatrix} \mathbf{W}_i & 0 \\ 0 & \overline{\mathbf{W}}_i \end{bmatrix}$$

Predetermined regressors: GMM estimator

The moment conditions

$$\mathrm{E}(\mathbf{W}_i'\Delta\mathbf{u}_i) = \mathbf{0} \qquad \text{and} \qquad \mathrm{E}(\overline{\mathbf{W}}_i'\mathbf{v}_i) = \mathbf{0}$$

can be written as

$$\mathrm{E}(\mathbf{W}_i^{+\prime}\mathbf{v}_i^{+}) = \mathbf{0}.$$

- The two-step GMM estimator is of the same form as before. Just replace $\Delta\mathbf{y}$, $\Delta\mathbf{y}_{-1}$, $\Delta\mathbf{X}$ and $\mathbf{W}$ by $\mathbf{y}^{+}$, $\mathbf{y}_{-1}^{+}$, $\mathbf{X}^{+}$ and $\mathbf{W}^{+}$.
- To make the GMM estimator feasible, we need a consistent one-step estimator to generate residuals from which a the covariance matrix of $\mathbf{v}_i^{+}$ can be estimated.
- A possible first step is to estimate the model neglecting the moment restrictions applied to the level equation. This is the estimator discussed previously.

Strict exogeneity: Additional moment conditions

Again assume that some regressors are uncorrelated with the individual effects and separate

$$\mathbf{x}_{it} = [\mathbf{x}_{1,it}, \mathbf{x}_{2,it}],$$

where $\mathbf{x}_{1,it}$ is uncorrelated with c_i, while $\mathbf{x}_{2,it}$ is correlated with c_i.

If the $\mathbf{x}_{it}$ are strictly exogenous, there are additional $K_1 T$ moment restrictions available with respect to the level equation (all other restrictions are already exploited for the differenced equation).

There are many ways to write down these additional $K_1 T$ moment restrictions. A possible way is

$$\mathrm{E}\left[\mathbf{x}'_{1,it} \sum_{s=1}^{T} u_{is}/T\right] = \mathrm{E}\left[\mathbf{x}'_{1,it}\bar{u}_i\right] = \mathbf{0} \qquad \text{for } t = 1, \ldots, T.$$

Strictly exogenous regressors: All moment conditions

- First, we have the $(T-2)(T-1)/2$ moment conditions obtained from the lagged endogenous variable and applied to the differenced equation:

$$\mathrm{E}[y_{is}\Delta u_{it}] = 0 \qquad \text{for } t = 3, \ldots, T \text{ and } s \leq t-2.$$

- In addition, we have the $K(T-2)(T+1)/2$ moment conditions obtained from $\mathbf{x}_{it}$ and applied to the differenced equation:

$$\mathrm{E}[\mathbf{x}'_{is}\Delta u_{it}] = \mathbf{0} \qquad \text{for } s \leq t-1.$$

- Finally, we have the $K_1 T$ moment conditions obtained from $\mathbf{x}_{1,it}$ and applied to the level equation:

$$\mathrm{E}[\mathbf{x}'_{1,it}\bar{u}_i] = \mathbf{0} \quad \text{for } t = 1, \ldots, T.$$

  Since $\mathbf{x}_{1,it}$ is uncorrelated with c_i, this implies

$$\mathrm{E}[\mathbf{x}'_{1,it}\bar{v}_i] = \mathbf{0} \quad \text{for } t = 1, \ldots, T.$$

System of level and differenced equations

Hence, we again have moment conditions with respect to the average level equation

$$\bar{y}_i = \alpha \bar{y}_{i,-1} + \bar{\mathbf{x}}_i \boldsymbol{\beta} + \bar{v}_i,$$

and with respect to the differenced equation

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \Delta \mathbf{x}_{it} \boldsymbol{\beta} + \Delta u_{it}, \quad t = 3, \ldots, T.$$

To jointly use all moment conditions, stack the two equations one on another

$$\begin{pmatrix} \Delta y_{i3} \\ \vdots \\ \Delta y_{iT} \\ \bar{y}_i \end{pmatrix} = \begin{pmatrix} \Delta y_{i2} \\ \vdots \\ \Delta y_{i,T-1} \\ \bar{y}_{i,-1} \end{pmatrix} \alpha + \begin{pmatrix} \Delta \mathbf{x}_{i3} \\ \vdots \\ \Delta \mathbf{x}_{iT} \\ \bar{x}_i \end{pmatrix} \boldsymbol{\beta} + \begin{pmatrix} \Delta u_{i3} \\ \vdots \\ \Delta u_{iT} \\ \bar{v}_i \end{pmatrix}$$

or compactly

$$\mathbf{y}_i^+ = \mathbf{y}_{i,-1}^+ \alpha + \mathbf{X}_i^+ \boldsymbol{\beta} + \mathbf{v}_i^+.$$

Strictly exogenous regressors: Instrument matrices

Collecting all instruments for the differenced equation yields the instrument matrix

$$\mathbf{W}_i = \begin{bmatrix} [\mathbf{y}_{i,1:1}, \mathbf{x}_{i,1:T}] & & & 0 \\ & [\mathbf{y}_{i,1:2}, \mathbf{x}_{i,1:T}] & & \\ & & \ddots & \\ 0 & & & [\mathbf{y}_{i,1:T-2}, \mathbf{x}_{i,1:T}] \end{bmatrix}$$

Collecting all valid instruments for the level equation yields the instrument matrix

$$\overline{\mathbf{W}}_i = [\mathbf{x}_{1,i1}, \ldots, \mathbf{x}_{1,iT}] \, .$$

Putting the instruments together yields

$$\mathbf{W}_i^+ = \begin{bmatrix} \mathbf{W}_i & 0 \\ 0 & \overline{\mathbf{W}}_i \end{bmatrix}$$

Strictly exogenous regressors: GMM estimator

The moment conditions

$$\mathrm{E}(\mathbf{W}_i'\Delta\mathbf{u}_i) = \mathbf{0} \qquad \text{and} \qquad \mathrm{E}(\overline{\mathbf{W}}_i'\bar{v}_i) = \mathbf{0}$$

can be written as

$$\mathrm{E}(\mathbf{W}_i^{+\prime}\mathbf{v}_i^{+}) = \mathbf{0}.$$

- The two-step GMM estimator is of the same form as before. Just replace $\Delta\mathbf{y}$, $\Delta\mathbf{y}_{-1}$, $\Delta\mathbf{X}$ and $\mathbf{W}$ by $\mathbf{y}^{+}$, $\mathbf{y}_{-1}^{+}$, $\mathbf{X}^{+}$ and $\mathbf{W}^{+}$.
- To make the GMM estimator feasible, we need a consistent one-step estimator to generate residuals from which a the covariance matrix of $\mathbf{v}_i^{+}$ can be estimated.
- A possible first step is to estimate the model neglecting the moment restrictions applied to the level equation. This is the estimator discussed previously.

Outline

Introduction

Blundell and Bond (1998) examine the relevance of the instruments suggested by Arellano and Bond (1991).

For the simple AR(1) model, they show that

- the Arellano-Bond (AB) instruments may be weak, but
- under a specific initial condition there are additional moment conditions that yield strong instruments.

In the following we replicate their results and discuss their “system” GMM estimator.

Potential weakness of the AB instruments

Consider the simple AR(1) model with $T = 3$ observations,

$$y_{it} = \alpha y_{i,t-1} + c_i + u_{it}, \qquad t = 2, 3.$$

Note that so far, we have not made an assumption concerning the initial observation, y_{i1}.

Let us, however, assume that $\mathrm{E}(c_i) = 0$, $\mathrm{E}(u_{it})$, and $\mathrm{E}(c_i u_{it}) = 0$ for $t = 2, 3$. Let us also assume u_{it} is white noise.

The AB estimator applied GMM to the differenced equation,

$$\Delta y_{i3} = \alpha \Delta y_{i,2} + \Delta u_{i3}$$

using the single instrument $y_{i,1}$ (so α is just-identified).

Due to differencing, there is only one observation per individual which simplifies our subsequent arguments (but they extend to $T > 3$).

Potential weakness: Details

Recall that the quality of the first-stage regression

$$\Delta y_{i,2} = \pi y_{i,1} + r_i$$

hinges on instrument relevance; if π is “small”, $y_{i,1}$ is a weak instrument.

Subtracting y_{i1} from both sides of the AR(1) model for $t = 2$ yields

$$y_{i2} - y_{i1} = (\alpha - 1) y_{i1} + c_i + u_{i2}.$$

Hence, $\pi = \alpha - 1$ and $r_i = c_i + u_{i2}$.

Assuming **stationarity** (which includes $|\alpha| < 1$), we show on the next slides that, as $N \to \infty$, the POLS estimator of π is biased towards zero

$$\text{plim}\,(\hat{\alpha} - 1) = (\alpha - 1)\,\frac{k}{(\sigma_c^2/\sigma_u^2) + k}, \qquad k = \frac{1-\alpha}{1+\alpha} > 0,$$

because

$$0 < \frac{k}{(\sigma_c^2/\sigma_u^2) + k} < 1.$$

Discussion

- The result that the first-stage coefficient $\pi = \alpha - 1$ implies that for α near 1 (unit root), the lagged level y_{i1} has not much to say about $\Delta y_{i,2}$.
- The result that the OLS estimator of π is generally biased towards zero (and thus towards instrument irrelevance) strengthens this view, see Fig.
- In a simulation study, Blundell and Bond confirm that for $\alpha = 0.8$ and above, the AB estimator is extremely poor.

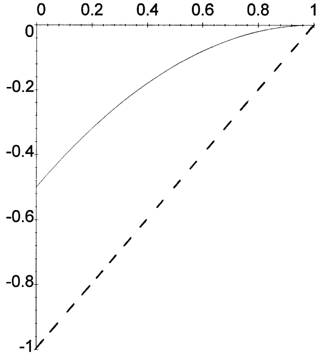

Figure $\operatorname{plim} \hat{\pi}$ against α for $T = 3$ and $\sigma_c^2 = \sigma_u^2 = 1$

Source: Blundell and Bond (1998, Fig. 1)

Details: plim of first-stage regression estimator

POLS applied to the first-stage regression

$$\Delta y_{i,2} = \pi y_{i,1} + r_i$$

yields

$$\hat{\pi} = \frac{\frac{1}{T}\sum_{i=1}^{N} y_{i,1}\Delta y_{i,2}}{\frac{1}{T}\sum_{i=1}^{N} y_{i,1}^2} = \pi + \frac{\frac{1}{T}\sum_{i=1}^{N} y_{i,1} r_i}{\frac{1}{T}\sum_{i=1}^{N} y_{i,1}^2} = \alpha - 1 + \frac{\frac{1}{T}\sum_{i=1}^{N} y_{i,1}(c_i + u_{i2})}{\frac{1}{T}\sum_{i=1}^{N} y_{i,1}^2}.$$

Hence, as $N \to \infty$,

$$\operatorname{plim} \hat{\pi} = \alpha - 1 + \frac{\mathrm{E}[y_{i,1}(c_i + u_{i2})]}{\mathrm{E}(y_{i,1}^2)} = \alpha - 1 + \frac{\mathrm{E}(y_{i,1}c_i) + \mathrm{E}(y_{i,1}u_{i2})}{\mathrm{Var}(y_{i,1}^2)}.$$

To find $\operatorname{plim} \hat{\pi}$, we have to make assumption on the relation between the initial observation and the error components c_i and u_{it}, $t = 2, \ldots, T$.

For this proof, let us assume (weak) stationarity which means that population moments are constant over time.

Details cont'd

A. $\mathrm{E}(y_{i,1}u_{i2})$:

Due to being white noise, u_{i3} is uncorrelated with y_{i2}. By stationarity, this also holds for u_{i2} and y_{i1}:

$$\mathrm{E}(y_{i,1}u_{i2}) = 0.$$

B. $\mathrm{E}(y_{i,1}c_i)$:

Multiplying the AR(1) model with c_i and taking expectations yields

$$\mathrm{E}(y_{it}c_i) = \alpha\,\mathrm{E}(y_{i,t-1}c_i) + \underbrace{\mathrm{E}(c_i^2)}_{\sigma_c^2} + \underbrace{\mathrm{E}(u_{i,t}c_i)}_{0}, \qquad t = 2, \ldots, T.$$

By stationarity, $\phi \equiv \mathrm{E}(y_{i,1}c_i) = \mathrm{E}(y_{it}c_i) = \mathrm{E}(y_{i,t-1}c_i)$. Solve for ϕ:

$$\phi = \sigma_c^2/(1-\alpha).$$

Details cont'd

C. $\mathrm{Var}(y_{i,1}^2)$:

Using the definition of the AR(1) model for $t = 2, \ldots, T$,

$$\mathrm{Var}(y_{it}) = \alpha^2 \,\mathrm{Var}(y_{i,t-1}) + \underbrace{\mathrm{Var}(c_i)}_{\sigma_c^2} + \underbrace{\mathrm{Var}(u_{it})}_{\sigma_u^2} + 2\alpha \underbrace{\mathrm{Cov}(y_{i,t-1}, c_i)}_{\phi},$$

where we used the fact that all other covariances are zero. By stationarity, $\sigma_y^2 \equiv \mathrm{Var}(y_{i1}) = \mathrm{Var}(y_{i2}) = \ldots = \mathrm{Var}(y_{iT})$. Solve for σ_y^2:

$$\sigma_y^2 = \frac{\sigma_c^2 + \sigma_u^2 + 2\alpha\phi}{1-\alpha^2} = \frac{\sigma_c^2 + \sigma_u^2 + \frac{2\alpha}{1-\alpha}\sigma_c^2}{1-\alpha^2} = \frac{\sigma_u^2 + \frac{1+\alpha}{1-\alpha}\sigma_c^2}{1-\alpha^2}.$$

Finally

Substituting all results yields

$$
\begin{aligned}
\operatorname{plim}\hat{\pi} &= \alpha - 1 + \frac{\sigma_c^2}{1-\alpha} \cdot \frac{1-\alpha^2}{\sigma_u^2 + \frac{1+\alpha}{1-\alpha}\sigma_c^2} = \alpha - 1 + \frac{1-\alpha^2}{1-\alpha} \cdot \frac{\sigma_c^2}{\sigma_u^2 + \frac{1+\alpha}{1-\alpha}\sigma_c^2} \\
&= (\alpha - 1)\left[1 - \frac{1-\alpha^2}{(1-\alpha)^2} \cdot \frac{\sigma_c^2}{\sigma_u^2 + \frac{1+\alpha}{1-\alpha}\sigma_c^2}\right] = (\alpha - 1)\left[1 - \frac{\frac{1+\alpha}{1-\alpha}\sigma_c^2}{\sigma_u^2 + \frac{1+\alpha}{1-\alpha}\sigma_c^2}\right] \\
&= (\alpha - 1)\left[\frac{\sigma_u^2 + \frac{1+\alpha}{1-\alpha}\sigma_c^2 - \frac{1+\alpha}{1-\alpha}\sigma_c^2}{\sigma_u^2 + \frac{1+\alpha}{1-\alpha}\sigma_c^2}\right] = (\alpha - 1)\frac{\sigma_u^2}{\sigma_u^2 + \frac{1+\alpha}{1-\alpha}\sigma_c^2} \\
&= (\alpha - 1)\frac{\frac{1-\alpha}{1+\alpha}}{\frac{1-\alpha}{1+\alpha} + (\sigma_c^2/\sigma_u^2)} \quad \text{q.e.d.}
\end{aligned}
$$

Idea

Having shown that $y_{i,t-2}$ may be a weak instrument for the differenced equation, Blundell and Bond (1998) suggest to use a lagged difference, $\Delta y_{i,t-1}$ for the level equation

$$y_{it} = \alpha y_{i,t-1} + c_i + u_{it}.$$

Note that $\Delta y_{i,t-1}$ is often a good instrument for $y_{i,t-1}$: From the differenced AR(1) model

$$\Delta y_{i,t-1} = \alpha \Delta y_{i,t-2} + \Delta u_{i,t-1}$$

we can see that $\Delta y_{i,t-1}$ appears to be unrelated to c_i and u_{it}. However, this is not generally true. To see this, solve it backwards which yields

$$\Delta y_{i,t-1} = \alpha^{t-3} \Delta y_{i,2} + \sum_{j=0}^{t-4} \alpha^j \Delta u_{i,t-j}, \qquad t-1 = 3, \ldots, T.$$

Hence, $\Delta y_{i,t-1}$ is a valid instrument if u_{it} is white noise and $\Delta y_{i,2}$ does not depend on c_i.

Stationarity assumption

A natural — but somewhat restrictive — assumption is that of stationarity.

The unconditional mean of a stationary AR(1) model without intercept,

$$y_{it} = \alpha y_{i,t-1} + v_{it},$$

is zero, $\mathrm{E}(y_{it}) = 0$. The unconditional variance is, as shown in the proof above,

$$\mathrm{Var}(y_{it}) = \frac{\sigma_u^2 + \frac{1+\alpha}{1-\alpha}\sigma_c^2}{1-\alpha^2} = \frac{\sigma_c^2}{(1-\alpha)^2} + \frac{\sigma_u^2}{1-\alpha^2}.$$

A specification of y_{i1} that achieves this is

$$y_{i1} = \frac{c_i}{1-\alpha} + \frac{u_{i1}}{\sqrt{1-\alpha^2}},$$

where $\mathrm{E}(c_i) = \mathrm{E}(u_{i1}) = 0$, $\mathrm{E}(c_i u_{i1}) = 0$, $\mathrm{Var}(u_{i1}) = \sigma_u^2$, and $\mathrm{E}(u_{i1}u_{it}) = 0$, $t = 2, \ldots, T$.

The Blundell-Bond assumption

Blundell and Bond (1998) make the less restrictive assumption on the first observation

$$y_{i1} = \frac{c_i}{1-\alpha} + u_{i1},$$

where $\mathrm{E}(c_i) = \mathrm{E}(u_{i1}) = 0$, $\mathrm{E}(c_i u_{i1}) = 0$, and $\mathrm{E}(u_{i1} u_{it}) = 0$, $t = 2, \dots, T$, but $\mathrm{Var}(u_{i1})$ is left unrestricted.

What does this assumption mean for the process y_{it}?

- **Instrument validity**: Δy_{i2} is unrelated to c_i, see proof below. Then (see backward solution above) any $\Delta y_{i,t-1}$ is unrelated to c_i and u_{it} and $\Delta y_{i,t-1}$ may instrument $y_{i,t-1}$ in the level equation.
- **Interpretation of the assumption**: Conditional on c_i, the process y_{it} fluctuates around the mean $c_i/(1-\alpha)$. The first observation deviates from this conditional mean by the amount u_{i1}. What we require is that this deviation u_{i1} is uncorrelated with c_i.

Proof: Instrument validity

Start again from the AR(1) model for $t = 2$ from which we subtract y_{i1} from both sides:

$$\Delta y_{i2} = (\alpha - 1) y_{i1} + c_i + u_{i2}.$$

Substituting

$$y_{i1} = \frac{c_i}{1-\alpha} + u_{i1}$$

yields

$$\Delta y_{i2} = (\alpha - 1)\left(\frac{c_i}{1-\alpha} + u_{i1}\right) + c_i + u_{i2} = u_{i2} + (\alpha - 1) u_{i1}.$$

Hence, Δy_{i2} is unrelated to c_i.

It solely depends on the white noise disturbances u_{i2} and u_{i1}.

Blundell-Bond: Instrument relevance

Consider again the AR(1) model with $T = 3$,

$$y_{i3} = \alpha y_{i2} + c_i + u_{it}$$

for which we use $\Delta y_{i,2}$ as instrument (and hence have only one observation per individual). Then the first-stage regression is

$$y_{i2} = \kappa \Delta y_{i,2} + r_i$$

Blundell and Bond (1998) claim that the plim of the POLS estimator is

$$\operatorname{plim} \hat{\kappa} = \frac{1}{2} \, \frac{1}{1+\alpha}, \qquad |\alpha| < 1,$$

which does not tend to zero as α tends to 1.

All moment conditions

The Blundell-Bond estimator uses the known AB moment conditions for the differenced equation

$$\mathrm{E}[y_{i1}\Delta u_{it}] = \cdots = \mathrm{E}[y_{i,t-2}\Delta u_{it}] = 0, \qquad t = 3, \ldots, T.$$

It adds the $T-2$ moment conditions for the level equation

$$\mathrm{E}[\Delta y_{i,t-1} v_{it}] = 0, \qquad t = 3, \ldots, T.$$

System of level and differenced equations

We have moment conditions with respect to the level equation

$$y_{it} = \alpha y_{i,t-1} + v_{it}$$

and with respect to the differenced equation

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \Delta u_{it}.$$

To jointly use all moment conditions, we stack the two equations

$$\begin{pmatrix} \Delta y_{i3} \\ \vdots \\ \Delta y_{iT} \\ y_{i3} \\ \vdots \\ y_{iT} \end{pmatrix} = \begin{pmatrix} \Delta y_{i2} \\ \vdots \\ \Delta y_{i,T-1} \\ y_{i2} \\ \vdots \\ y_{i,T-1} \end{pmatrix} \alpha + \begin{pmatrix} \Delta u_{i3} \\ \vdots \\ \Delta u_{iT} \\ v_{i3} \\ \vdots \\ v_{iT} \end{pmatrix}$$

or compactly

$$\mathbf{y}_i^+ = \mathbf{y}_{i,-1}^+ \alpha + \mathbf{v}_i^+.$$

Instrument matrices

Collecting all instruments for the differenced equation yields

$$\mathbf{W}_i = \begin{bmatrix} \mathbf{y}_{i,1:1} & & & 0 \\ & \mathbf{y}_{i,1:2} & & \\ & & \ddots & \\ 0 & & & \mathbf{y}_{i,1:T-2} \end{bmatrix}$$

Collecting the instruments for the level equation yields

$$\overline{\mathbf{W}}_i = \begin{bmatrix} \Delta y_{i2} & & & 0 \\ & \Delta y_{i3} & & \\ & & \ddots & \\ 0 & & & \Delta y_{i,T-1} \end{bmatrix}$$

Putting the instruments together yields

$$\mathbf{W}_i^+ = \begin{bmatrix} \mathbf{W}_i & 0 \\ 0 & \overline{\mathbf{W}}_i \end{bmatrix}$$

GMM estimator

The moment conditions

$$\mathrm{E}(\mathbf{W}_i' \Delta \mathbf{u}_i) = \mathbf{0} \qquad \text{and} \qquad \mathrm{E}(\overline{\mathbf{W}}_i' \mathbf{v}_i) = \mathbf{0}$$

can be written as

$$\mathrm{E}(\mathbf{W}_i^{+\prime} \mathbf{v}_i^{+}) = \mathbf{0}.$$

- The two-step GMM estimator is of the same form as before. It is often called the Blunell-Bond system estimator.
- To make the GMM estimator feasible, we need a consistent one-step estimator to generate residuals from which a the covariance matrix of $\mathbf{v}_i^{+}$ can be estimated.
- A possible first step is to use the AB one-step estimator.
- Adding strictly exogenous or predetermined regressors proceeds exactly as discussed above.

Discussion

- Blundell and Bond (1998) show that their system GMM estimator produces dramatic efficiency gains over the basic first-difference AB estimator as $\alpha \to 1$ or σ_c^2/σ_u^2 increases.
- In fact, for $T = 4$ and $\sigma_c^2/\sigma_u^2 = 1$, the asymptotic variance ratio of the AB estimator to the BB system estimator is 1.75 for $\alpha = 0$ and increases to 3.26 for $\alpha = 0.5$ and 55.4 for $\alpha = 0.9$.
- While things improve for first-difference GMM as T increases, with short T and persistent series, the Blundell and Bond findings support the use of the extra moment conditions.
- These results are reviewed and corroborated in Blundell and Bond (2000) and Blundell, Bond Windmeijer (2000). In fact, the system GMM estimator not only improves the estimation precision but also reduces the finite sample bias.