Econometric Methods for Panel Data

Based on the books by Baltagi: *Econometric Analysis of Panel Data* and by Hsiao: *Analysis of Panel Data*

Robert M. Kunst
robert.kunst@univie.ac.at

University of Vienna
and
Institute for Advanced Studies Vienna

April 18, 2016

Outline

The basic dynamic panel model

Consider the model

$$y_{it} = \delta y_{i,t-1} + x_{it}'\beta + u_{it},$$

with one-way random effects

$$u_{it} = \mu_i + \nu_{it},$$

with $\mu_i \sim iid(0, \sigma_\mu^2)$ and $\nu_{it} \sim iid(0, \sigma_\nu^2)$, μ and ν independent.

For simplicity, we first concentrate on the case $\beta = 0$.

The Nickell bias

Nickell (1981, *Econometrica*) investigated the pooled OLS and fixed-effects estimators for the dynamic panel model. Both of these estimators are biased and (for $N \to \infty$) even inconsistent:

- The pooled OLS estimator shows the Hurwicz bias that is typical for all autoregressions. It is directed toward 0 ($|\mathrm{E}\hat{\delta}| < |\delta|$ usually) and of order T^{-1}. For $T \to \infty$, the bias disappears. For $N \to \infty$, it can be substantial for small T.
- The fixed-effects estimator has an inherent bias of order T^{-1} that is usually negative ($\mathrm{E}\hat{\delta} < \delta$). For $\delta \to -1$ or for $T \to \infty$, the bias disappears. For $N \to \infty$, it can be substantial for small T. This Nickell bias is typically smaller than the Hurwicz bias.

Instrumental variables for dynamic panels

There are two sources of the Nickell bias: the Hurwicz bias and regressor-error correlation effects, such as of the average $T^{-1} \sum y_{i,t}$ with errors $\nu_{i,t}$. Instrumental-variables techniques mainly target the correlation effects.

A valid instrument should be well correlated with $y_{i,t-1}$ but not with the errors (u_t for OLS and ν_t for FE). Functions of lagged observations $y_{i,t-h}$ for $h > 0$ could be instruments.

The IV estimator of Anderson and Hsiao

T.W. ANDERSON and C. HSIAO (1981) suggested taking first differences, which eliminates the effects and the inconsistency caused by the correlation of individual time averages and errors:

$$\Delta y_{i,t} = \delta \Delta y_{i,t-1} + \Delta \nu_{i,t}$$

Regressor and errors are again correlated. $y_{i,t-2}$ and $\Delta y_{i,t-2}$ are valid instruments. Later authors discouraged the usage of $\Delta y_{i,t-2}$ as an instrument. If $\nu_{i,t}$ are uncorrelated over time, a GLS version of the IV estimator may be an appealing estimator.

The idea of the Arellano-Bond estimator

Estimating the model in first differences instead of FE implies that no bias emanates from the time average $T^{-1} \sum y_{i,t}$. Note that

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2}) + \nu_{i,t} - \nu_{i,t-1}, \quad t = 3, \ldots, T,$$

implies an error $\Delta\nu_{i,t}$ that is correlated with the regressor and also autocorrelated. $y_{i,t-s}$ for $s > 1$ can be used as instruments and thus define a consistent estimator. The set of instruments $\{y_{i,t-2}, \ldots, y_{i,1}\}$ increases with increasing t. For this reason, the estimator is usually presented in a GMM framework.

This estimator is inefficient, as the error process is autocorrelated. A GLS step yields a consistent and (under ideal assumptions) efficient estimator, the one-step Arellano-Bond estimator.

The one-step Arellano-Bond estimator

Because of $\mathrm{E}(\Delta\nu_i \Delta\nu_i') = \sigma_\nu^2 \mathbf{G}$ with the $(T-2) \times (T-2)$–matrix

$$\mathbf{G} = \begin{pmatrix} 2 & -1 & 0 & 0 & \dots \\ -1 & 2 & -1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & 0 & -1 & 2 & -1 \\ \dots & 0 & 0 & -1 & 2 \end{pmatrix},$$

the estimator can be written as

$$\begin{aligned} \hat{\delta}_1 &= [(\Delta y_{-1})'\mathbf{W}\{\mathbf{W}'(\mathbf{I}_N \otimes \mathbf{G})\mathbf{W}\}^{-1}\mathbf{W}'(\Delta y_{-1})]^{-1} \\ &\quad \times (\Delta y_{-1})'\mathbf{W}\{\mathbf{W}'(\mathbf{I}_N \otimes \mathbf{G})\mathbf{W}\}^{-1}\mathbf{W}'(\Delta y), \end{aligned}$$

with fitting instrument matrices $\mathbf{W}$.

The instrument matrix $\mathbf{W}$

The instrument matrix $\mathbf{W} = (\mathbf{W}_1', \ldots, \mathbf{W}_N')'$ contains the instruments that proliferate as t grows:

$$\mathbf{W}_i = \begin{pmatrix} y_{i,1} & 0 & 0 & \ldots \\ 0 & y_{i,1}, y_{i,2} & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & 0 & y_{i,1}, \ldots, y_{i,T-2} \end{pmatrix}$$

are matrices of dimension $(T-2) \times (T-2)(T-1)/2$, and $\mathbf{W}$ has dimension $N(T-2) \times (T-2)(T-1)/2$.

The two-step Arellano-Bond estimator

If the assumption that $\nu_{it} \sim iid(0, \sigma_\nu^2)$ holds, the one-step estimator is efficient. Otherwise, an idea in the spirit of White-Eicker is to replace $\mathbf{W}'(\mathbf{I}_N \otimes \mathbf{G})\mathbf{W}$ by

$$\hat{\mathbf{V}}_N = \sum_{i=1}^{N} \hat{\nu}_i \hat{\nu}_i',$$

with $\hat{\nu}_i$ containing the residuals from the first step. The resulting estimator

$$\begin{aligned}\hat{\delta}_2 &= \{(\Delta y_{-1})' \mathbf{W} \hat{\mathbf{V}}_N^{-1} \mathbf{W}' (\Delta y_{-1})\}^{-1} \\ &\quad \times (\Delta y_{-1})' \mathbf{W} \hat{\mathbf{V}}_N^{-1} \mathbf{W}' (\Delta y)\end{aligned}$$

is the two-step Arellano-Bond estimator.

The variance of the Arellano-Bond estimator

The asymptotically correct variance estimator for the Arellano-Bond estimator is given by the first part of the estimator expression:

$$\widehat{\mathrm{var}}\hat{\hat{\delta}}_2 = \{(\Delta y_{-1})' \mathbf{W} \hat{\mathbf{V}}_N^{-1} \mathbf{W}' (\Delta y_{-1})\}^{-1}$$

Some authors claim that this estimator can perform poorly. If the ν_{it} are really *iid*, the first portion of the one-step estimator

$$\widehat{\mathrm{var}}\hat{\hat{\delta}}_1 = \hat{\sigma}_\nu^2 [(\Delta y_{-1})' \mathbf{W} \{\mathbf{W}' (\mathbf{I}_N \otimes \mathbf{G}) \mathbf{W}\}^{-1} \mathbf{W}' (\Delta y_{-1})]^{-1}$$

is also consistent for the variance of both the one-step and two-step estimators.

Specification tests for the Arellano-Bond model

The construction uses more instruments ('moment conditions') than are required for exact identification. This implies that a specification test can be based on the overidentifying conditions.

The often reported Sargan test statistic is, under the null of the validity of the restrictions, asymptotically distributed as χ^2 with degrees of freedom corresponding to the number of overidentifying restrictions ($(T-1)(T-2)/2-1$ for the AR(1) without further covariates).

Arellano-Bond with exogenous variables

The variables x in the general dynamic model

$$y_{it} = \delta y_{i,t-1} + x_{it}'\beta + u_{it}$$

and its differenced version

$$\Delta y_{it} = \delta \Delta y_{i,t-1} + \Delta x_{it}'\beta + \Delta \nu_{it}$$

can be *strictly exogenous* (uncorrelated with errors at all leads and lags) or *predetermined* (uncorrelated with current and future errors though correlated with past errors).

- Strictly exogenous variables serve as their own instruments. They are added to the list of instruments at all leads and lags.
- Only lags of predetermined variables can serve as instruments, their number increasing in t in analogy to the Arellano-Bond construction for y_{t-1}.

System GMM: the idea

Additionally to the equation in differences with lagged level terms as instruments (well correlated with lagged differences, though uncorrelated with errors from the differenced equation), it may pay to also consider the equation in levels, with lagged differences as instruments (well correlated with lagged levels, though uncorrelated with errors from the level equation). This option can yield gains in efficiency, particularly if $\delta \uparrow 1$.

An estimator constructed along these lines, with weights on the many moment conditions, is called the **system GMM estimator**. It is quite complex notationally.

GMM or system GMM?

Whereas some authors (Arellano and Bover, 1995, Blundell and Bond, 1998) advertise system GMM, others criticize its performance and argue that it uses too many instruments. A major advantage may be that system GMM has the smallest finite-sample bias among all comparable estimators, under ideal conditions. On the other hand, it requires stronger stationarity assumptions on variables for consistency.

System GMM: the formula

The system GMM estimator can be written as

$$\{\mathbf{W}'\bar{\mathbf{H}}'\mathbf{M}(\mathbf{M}'\bar{\mathbf{H}}\bar{\boldsymbol{\Omega}}\bar{\mathbf{H}}'\mathbf{M})^{-1}\mathbf{M}'\bar{\mathbf{H}}\mathbf{W}\}^{-1}$$
$$\times\mathbf{W}'\bar{\mathbf{H}}'\mathbf{M}(\mathbf{M}'\bar{\mathbf{H}}\bar{\boldsymbol{\Omega}}\bar{\mathbf{H}}'\mathbf{M})^{-1}\mathbf{M}'\bar{\mathbf{H}}y,$$

where $\bar{\mathbf{H}}$ denotes a ‘rotation matrix’ $\mathbf{I}_N \otimes \mathbf{H}$, with the first $T-1$ rows of $\mathbf{H}$ transforming the constant or effect vectors to 0, $\mathbf{W}$ collects the regressor variables, and $\mathbf{M}$ denotes the matrix of instruments. In practice, $\bar{\boldsymbol{\Omega}}$ contains estimates for $\mathbf{I}_N \otimes \boldsymbol{\Omega}$ retrieved from a preliminary step.