

## Panel dinámico

Roberto Montero Granados

*Universidad de Granada*

*abril, 2010*

Supongamos un modelo lineal del tipo:

$$y_{it} = \eta_i + \sum_{j=1}^n \beta_0 y_{it-j} + \beta_1 X_{it} + v_i + u_{it}$$

Donde para cada unidad  $i$  en el tiempo  $t$  la variable dependiente  $y$  depende de sí misma con uno o varios retardos  $j$  y de un conjunto de variables independientes que están en la matriz  $X$ . Además cada individuo  $i$  tiene un carácter idiosincrático no estocástico  $\eta$  y unos errores idiosincráticos y otros normales ambos i.i.d.  $N(0, \sigma)$

Aplicar MCO a este modelo, o MLG de panel con efectos fijos o aleatorios provoca errores estándar de las estimaciones de los parámetros inconsistentes porque, por construcción, el efecto inobservable ( $\eta_i$ ) está correlacionado con los retardos de la dependiente ( $y_{it-j}$ ).

Para corregir este problema se podrían aplicar variables instrumentales. Anderson e Hsiao (1981,1992) proponen utilizar retardos de la dependiente, tanto en nivel como en diferencias. Arellano y Bond (1991) construyen un estimador basado en el Método Generalizado de los Momentos (GMM), que utiliza variables instrumentales basadas en retardos y diferencias de todas las variables del modelo y que está especialmente propuesto para paneles con muchos individuos y pocos periodos. Las posibles variables instrumentales y sus retardos las obtienen del método desarrollado por Hansen (1982).

Concretamente, el modelo a estimar es ahora

$$y_{it} = \beta_0 y_{i,t-1} + \beta_1 x_{it} + \beta_2 w_{it} + v_i + e_{it}$$

donde  $y$  es la variable dependiente del individuo  $i$  en el momento  $t$ ,  $x$  es un vector de variables exógenas y  $w$  es un vector de variables predeterminadas o endógenas. Pero  $v_i$  está correlacionado con  $y_{i,t-1}$ , para evitarlo lo que se estima también es el modelo en primeras diferencias que queda:

$$\Delta y_{it} = \Delta \beta_0 y_{i,t-1} + \Delta \beta_1 x_{it} + \Delta \beta_2 w_{it} + \Delta e_{it} \quad \text{si } g = 1$$

¿Cómo citar?: Montero. R (2010): *Panel dinámico*. Documentos de Trabajo en Economía Aplicada. Universidad de Granada. España

Ahora como  $\Delta y_{i,t-1}$  también está correlacionado con  $\Delta e_{it}$  se hace necesario utilizar instrumentos de las variables para que la estimación sea insesgada. Arellano y Bond utilizan retardos en la/s variable/s endógenas y en las predeterminadas y diferencias en las variables estrictamente exógenas. La diferencia entre predeterminadas y estrictamente endógenas consiste en que una variable es predeterminada cuando su valor actual está correlacionado con valores pasados de  $e$  o de  $y$ . Una variable es endógena cuando su valor actual está correlacionado con valores actuales y pasados de  $e$  o de  $y$ .

El estimador GMM estima la relación entre dependiente e independientes utilizando la información de ambas ecuaciones, en niveles y en diferencias.

En Stata hay implementados varios estimadores, de una etapa y bietápicos. También hay, de ambos, versiones robustas (vce) (La versión robusta del estimador bietápico es atribuible a Windmijer, 2005) . Además existen dos estimadores el inicial de Arellano-Bond (1991) (implementado en Stata como xtabond) y el estimador sistemático desarrollado por Arellano-Bower (1995) y Blundell-Bond (1998) (implementado como xtpdsys)

Una restricción importante del estimador, que debe corregirse con una correcta modelización, es que no puede existir autocorrelación de segundo orden en las primeras diferencias de los errores. Este se realiza mediante el test de Arellano-Bond (estat abund). Es deseable que las primeras diferencias estén correlacionadas en primero orden, ya que de lo contrario estaría indicando que no existen efectos dinámicos y el estimador GMM no sería adecuado, pero no pueden existir dichas diferencias en segundo orden.

En estat abund la hipótesis nula es que no existe autocorrelación por lo que un pvalor  $< 0.05$  indica que se rechaza la hipótesis nula y que sí existe dicha autocorrelación.

También es conveniente pasar el test de Sargan (estat sargan) de sobreidentificación. En este modelo es conveniente que las ecuaciones estén sobreidentificadas (de hecho el estimador GMM podría interpretarse como una combinación lineal de todas las estimaciones posibles de un modelo sobreidentificado).

En estat sargan la  $H_0$  es que las ecuaciones están correctamente sobreidentificadas por lo que un p valor  $> 0.05$  nos indica que la especificación es correcta.

Si el p valor es  $< 0.05$  pueden estar ocurriendo dos cosas: a) que las ecuaciones no esté suficientemente sobreidentificadas (entonces es conveniente pasar el estimador bietápico) o b) que exista heterocedasticidad, en cuyo caso está demostrado que no se conoce la distribución del test de sargan (que es una Chi cuadrado en presencia de homocedasticidad) y existe una tendencia (según Arellano y bond, 1991) a rechazar la hipótesis nula. En este caso lo conveniente es utilizar estimadores robustos (vce)

Para elegir entre el estimado inicial de Arellano-Bond (1991) (xtabond) o el estimador sistemático Arellano-Bower (1995) y Blundell-Bond (1998) (xtpdsys) no existe una regla fija. En general será preferible el segundo, por ser más reciente, porque utiliza condiciones de momento adicionales y porque es más eficiente y reporta errores menores, pero tiene el inconveniente de que exige que no exista autocorrelación en el error idiosincrático ( $v_i$ ) y que no exista correlación entre el efecto del panel ( $\eta_i$ ) y la primera diferencia de la primera observación de la dependiente ( $y_{i1}-y_{i0}$ ) (Blundell-Bond, 1998; y Blundell-Bond-Windmeijer, 2000). Requisitos para lo que no parece ofrecer ningún test de prueba.

#### Bibliografía:

Anderson, T. W., and C. Hsiao. 1981. Estimation of dynamic models with error components. *Journal of the American Statistical Association* 76: 598–606.

Anderson, T. W., and C. Hsiao. 1992. Formulation and estimation of dynamic models using panel data. *Journal of Econometrics* 18: 47–82.

Arellano, M., and S. Bond. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58: 277–297.

Hansen, L. P. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50: 1029–1054.

Windmeijer, F. 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics* 126: 25–51.

Arellano, M., and O. Bover. 1995. Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics* 68: 29–51.

Blundell, R., and S. Bond. 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87: 115–143.

Blundell, R., S. Bond, and F. Windmeijer. 2000. Estimation in dynamic panel data models: Improving on the performance of the standard GMM estimator. In *Nonstationary Panels, Cointegrating Panels and Dynamic Panels*, ed. B. H. Baltagi, 53–92. New York: Elsevier.