Modelos dinámicos de datos de panel.

Siga

Modelos dinámicos de datos de panel (lineales).

- Relajación del supuesto de exogeneidad estricta de los regresores.
- Planteamiento general.
- Consecuencias para los estimadores conocidos hasta ahora (paneles clásicos).
- Identificación.
- Procedimientos de estimación.
- Otros temas relevantes.

Siga

Relajación del supuesto de exogeneidad estricta.

$$Y_{it} = X_{it}\beta + \eta_i + \nu_{it}, i = 1,2,.., N; t = 1,2,..,T.$$

- Modelos clásicos de datos de panel: exogeneidad estricta de los regresores.
- Modelos dinámicos de datos de panel: se relaja el supuesto de exogeneidad estricta.

Atrás

Supuesto de exogeneidad estricta.

$$Y_{it} = X_{it}\beta + \eta_i + \nu_{it}, i = 1,2,.., N; t = 1,2,..,T.$$

[redacted]trás

$$E(Y_{it} \mid X_{it}, \eta_i) = X_{it}\beta + \eta_i .$$

Una vez controlado el efecto de los regresores y de los efectos individuales, el valor esperado de la variable de interés sólo depende de los valores contemporáneos de los regresores y heterogeneidad inobservada.

Ej. salarios en función de la edad.

La edad pasada no influye sobre los salarios actuales y éstos últimos no influyen sobre la edad futura. Esto mismo ocurre con shocks asociados con la edad. Si el salario mínimo depende de la edad, y ésta es estrictamente exógena, los salarios presentes no afectan al valor futuro del shock.

Atrás

Restricciones de momentos secuenciales.

$$Y_{it} = X_{it}\beta + \eta_i + \nu_{it}, i = 1,2,.., N; t = 1,2,..,T.$$

- Al margen de la correlación entre X y η.
- El término de error genérico está incorrrelacionado con los valores contemporáneos y retardados de los regresores (pero correlacionado con los futuros).

$$E(\nu_{it} \mid X_{it}, X_{it-1},....X_{i1}, \eta_i) = 0, \quad t = 1,2,..,T.$$

- Los regresores son <u>secuencialmente exógenos</u>, condicionados a los efectos individuales.

Regresores secuencialmente exógenos.

$$E(Y_{it} \mid X_{it}, X_{it-1}, \ldots, X_{i1}, \eta_i) = X_{it}\beta + \eta_i$$

- Una vez que se controla por los efectos de los regresores y de la heterogeneidad individual inobservada, los valores pasados de los regresores no afectan al valor esperado contemporáneo de Y.
- Permite que haya retroalimentación.

Atrás

Planteamiento general.

- Modelos a considerar.
- Algunos ejemplos introductorios e ilustrativos.

Atrás

Ejemplos.

- Modelos de consumo u oferta de trabajo a lo largo del ciclo vital con hábitos. El valor de δ indica la magnitud de los hábitos.
  - Consumo de tabaco, Becker, Grossman y Murphy (1994).
- Planes de inversión y consumo.
  - Keane y Runkle (1992); Bond y Meghir (1994).
- Dinámica de rentas.

Atrás

Modelos que usaremos.

- Modelo autorregresivo.
- Modelo autorregresivo con regresores exógenos o predeterminados.
- Modelo de Hausman y Taylor (1981).

Atrás

Modelo autorregresivo.

$$Y_{it} = \delta Y_{it-1} + \eta_i + \nu_{it}, i = 1,2,..,N; t = 1,2,..,T.$$

- Simple en su especificación.
- Presenta las características fundamentales de modelos dinámicos.
- $E\,(Y_{it-1}\nu_{is} | \eta_i)=0$, $s \geq t$; $E\,(\nu_{it} | Y_{it-1}, Y_{it-2}, .., Y_{i0}, \eta_i)=0$.
- $E\,(Y_{it} | Y_{it-1}, \eta_i)= \delta Y_{it-1}+ \eta_i$.
- Fuentes de persistencia: het. inobservada y δ.

Atrás

Modelo autorregresivo con regresores exógenos o predeterminados.

$$Y_{it} = \delta Y_{it-1} + \beta x_{it} + \eta_i + \nu_{it}, i = 1,2,..,N; t = 1,2,..,T.$$

- Supuesto: x es un regresor estrictamente exógeno.
- Se cumple:

$$E(\nu_{it} \mid x_i, Y_{it-1}, \ldots Y_{i0}, \eta_i) = 0, \quad t = 1,2,..,T.$$

- Ejemplos: expectativas racionales; Sida.

Atrás

Modelo de Hausman y Taylor (1981).

$$Y_{it} = X_{it}\beta + Z_i\gamma + \eta_i + \nu_{it},$$

$$i = 1,2,\ldots\ldots, N;$$

$$t = 1,2,\ldots\ldots., T.$$

- Regresores variantes e invariantes a lo largo del tiempo (que varían con los individuos).
- Regresores endógenos o estrictamente exógenos.

Atrás

Consecuencias para los estimadores de paneles clásicos.

- Inconsistencia de los estimadores habitualmente empleados en paneles clásicos.
    - estimador por MCO del modelo completo de panel.
    - estimador *intra* (*within*).
    - estimador por MCG del modelo de panel.
    - estimador en diferencias.

Siga

MCO del panel completo.

- Sesgados e inconsistentes por la correlación entre regresores y término de error: efectos individuales y error genérico.
- No ocurre lo mismo con datos de serie temporal (si el término de error no presenta autocorrelación).

Atrás

Estimador *intra* (efectos fijos).

- Sesgado y, en general, inconsistente.
- El tamaño del sesgo asintótico, Nickell (1981), es una función de 1/T, por lo que para T fijo, y pequeño, el sesgo no desaparece cualquiera que sea el tamaño de la muestra transversal.
- <u>En efecto</u>,
- Nickell(1981) facilita el <u>tamaño del sesgo</u>.

<u>Atrás</u>

$$Y_{it} - \bar{Y}_i = \delta(Y_{it-1} - \bar{Y}_{i-1}) + \nu_{it} - \bar{\nu}_i .$$

donde

$$\overline{Y}_{i-1} = \frac{\sum_{2}^{T} Y_{it-1}}{T-1}; \qquad \overline{\nu}_i = \frac{\sum_{t} \nu_{it}}{T}.$$

Atrás

[redacted]

Atrás

Estimador en diferencias.

- Anderson y Hsiao (1981) sugieren:

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2}) + (\nu_{it} - \nu_{it-1});$$

$$\Delta y_{it} = \delta \Delta y_{it-1} + \Delta \nu_{it}.$$

- El estimador por MCO del modelo en diferencias es sesgado e inconsistente. Su sesgo no desaparece al crecer T.

Atrás

Identificación.

- Lo que se presenta:
  - Identificación de parámetros estructurales en modelos con variables estrictamente exógenas o predeterminadas, pero correlacionadas con los efectos individuales.
- Lo que no se presenta:
  - Propiedades de series temporales de modelos con componentes del error (esquemas autorregresivos).
  - Restricciones de estacionariedad.
  - Identificación en modelos con múltiples efectos individuales o efectos individuales multiplicativos.

Atrás

Identificación de parámetros estructurales.

- Elementos a tener en cuenta:
    - Variables estrictamente exógenas; predeterminadas y heterogeneidad individual inobservada.

1. Regresión estática y exogeneidad estricta.
2. Modelo de ajuste parcial y exogeneidad estricta.
3. Modelo de ajuste parcial y variable predeterminada.
4. Efectos individuales incorrelacionados.

Atrás

Identificación en el caso de regresión estática y variable estrictamente exógena.

$$Y_{it} = \beta x_{it} + \mu_t + \eta_i + \nu_{it}, i = 1,2,.., N; t = 1,2,..,T.$$

- Donde, $(Y_T,..Y_1,x_T....,x_1,\eta_i)$ es un vector iid con momentos finitos hasta el segundo orden.
- Exogeneidad estricta:
  - En niveles, $E(\nu_{it}|\underline{x}_i^T)=0$; t =1,2…..T.
  - En diferencias, $E(\nu_{it}- \nu_{i(t-1)}|X_i^T)=0$; t=2….T.
- Para $T\geq 2$ el parámetro β está identificado.

Atrás

$$x_i^{\,t} = (x_{i1}, x_{i2}, x_{i3} \ldots\ldots, x_{it})'.$$

Por lo que,

$$x_i^{\,T} = (x_{i1}, x_{i2}, x_{i3} \ldots\ldots, x_{iT})'.$$

Atrás

Identificación para modelo de ajuste parcial y variable estrictamente exógena.

$$Y_{it} = \delta Y_{it-1} + \beta_0 x_{it} + \beta_1 x_{it-1} + \mu_t + \eta_i + \nu_{it}, i = 1,2,.., N; t = 1,2,.., T.$$

- Exogeneidad estricta:
  - En niveles, $E(\nu_{it}|x_i^T)=0$; t=2,…..T.
  - En diferencias, $E(\nu_{it}- \nu_{i(t-1)}|x_i^T)=0$; t=3,….T.
- Puede haber correlación serial de ν, por lo que $Y_{i(t-1)}$ es una variable endógena.
- Tomar primeras diferencias elimina la correlación del regresando retardado con la heterogeneidad individual, pero no la posible con el término genérico.
- Para $T\geq 3$ los parámetros δ, β_0, β_1 están identificados.
- Ejemplo: Becker-Grossman-Murphy (1994).

Atrás

Consumo de tabaco en USA, con datos de panel de estados.

$$C_{it} = \theta C_{it-1} + \beta\theta C_{it+1} + \gamma p_{it} + \mu_t + \eta_i + \nu_{it+1}, i = 1,2,.., N; t = 1,2,..,T.$$

C: consumo per cápita; p: precios; η: utilidad marginal del dinero, variante entre estados; μ: shocks agregados (correlacionados con precios); ν: errores, desplazamientos no observados de la utilidad a lo largo del ciclo vital. Los precios se consideran estrictamente exógenos, dados el consumo en el periodo anterior y posterior.

Atrás

Para T =3, las restricciones son:

$$E\left[\begin{pmatrix}1\\ x_{i1}\\ x_{i2}\\ x_{i3}\end{pmatrix}\left(\Delta Y_{i3}-\delta\Delta Y_{i2}-\beta_0\Delta x_{i3}-\beta_1\Delta x_{i2}-\Delta\mu_3\right)\right]=0.$$

$$E\left[Y_{i2}-\delta Y_{i1}-\beta_0 x_{i2}-\beta_1 x_{i1}-\mu_2\right]=0.$$

Lo que produce una identificación exacta.

Atrás

Identificación: modelo de ajuste parcial y variable predeterminada.

$$Y_{it} = \delta Y_{it-1} + \beta_0 x_{it} + \beta_1 x_{it-1} + \mu_t + \eta_i + \nu_{it}, i = 1,2,..,N; t = 1,2,..,T.$$

- Exogeneidad secuencial:
  - En niveles, $E(\nu_{it}|x_i^t, y_i^{t-1})=0$; t=2,…..T.
- Los errores contemporáneos están incorrelacionados con valores retardados de la variable Y y con los valores contemporáneos y retardados de x. Esto no elimina la retroalimentación que se puede dar entre los valores retardados de Y y los contemporáneos y futuros de x.
- A diferencia con el modelo con variable exógena, la presencia de una variable predeterminada exige para la identificación que los errores en primeras diferencias presenten autocorrelación de orden 1 y no presenten autocorrelación de ningún otro orden.

Siga

Identificación: modelo de ajuste parcial y variable predeterminada.

- En este caso, el supuesto anterior puede ser contrastado.
- Ej: ecuaciones de Euler para consumidores y empresas; oferta de trabajo femenina y maternidad.
- La exogeneidad secuencial implica que en el modelo en primeras diferencias se cumple:
- $E(\nu_{it} - \nu_{i(t-1)} | x_i^{t-1}, Y_i^{t-2}) = 0$; t=3,….T.
- Para $T \geq 3$ los parámetros δ, β_0, β_1 están <u>identificados</u>

Para T =3, las restricciones son:

$$E\left[\begin{pmatrix} 1 \\ Y_{i1} \\ x_{i1} \\ x_{i2} \end{pmatrix}\left(\Delta Y_{i3} - \delta \Delta Y_{i2} - \beta_0 \Delta x_{i3} - \beta_1 \Delta x_{i2} - \Delta \mu_3\right)\right] = 0.$$

$$E\left[Y_{i2} - \delta Y_{i1} - \beta_0 x_{i2} - \beta_1 x_{i1} - \mu_2\right] = 0.$$

Lo que produce una identificación exacta.

Atrás

Efectos individuales incorrelacionados.

- La presencia en un modelo de datos de panel de variables no correlacionadas con los efectos individuales añade un conjunto adicional de restricciones de identificación (generalmente T).
- Esto no suele ayudar, sin embargo, a que la identificación sea posible con menor dimensión temporal.

Atrás

Procedimientos de estimación.

- Generalidades.
- Lo que se presentará.
    - Arellano y Bond (1991).
- Lo que no se presentará:
    - Keane y Runkle (1992); Ahn y Schmidt (1995); Arellano y Bover (1995); Blundell y Bond (1998).

Siga

Generalidades.

- Utilizar una transformación que elimine la heterogeneidad individual inobservada: efectos fijos; diferencias; transformación ortogonal.
- Selección de instrumentos para variables endógenas en la ecuación anterior.
- Contrastes de autocorrelación en los errores en niveles y en primeras diferencias.
- Forma de las matrices de varianzas y covarianzas de los términos de error.

Atrás

Enfoque de Arellano y Bond (1991).

- Muy popular (DPD).
- Utilizan las condiciones de ortogonalidad entre los valores retardados de la variable del lado izquierdo y el término de error y las características de la matriz de varianzas y covarianzas del modelo analizado.
- En el análisis se parte del modelo autorregresivo.

Siga

$$Y_{it} = \delta Y_{it-1} + \eta_i + \nu_{it}, i = 1,2,..,N; t = 1,2,..,T.$$

En diferencias

$$\Delta y_{it} = \delta \Delta y_{it\text{-}1} + \Delta \nu_{it}.$$

Si el término de error en niveles era ruido blanco, el de la ecuación en diferencias será un MA(1) no invertible (raíz unitaria).
Para apreciar qué instrumentos pueden ser usados para estimar consistentemente δ, desgranemos la ecuación para distintos valores de t.

$$(y_{i3} - y_{i2}) = \delta(y_{i2} - y_{i1}) + (\nu_{i3} - \nu_{i2}). \quad T = 3.$$

<u>Siga</u>

En la ecuación anterior cabe usar y_1 como instrumento.

$$(y_{i4}-y_{i3})=\delta(y_{i3}-y_{i2})+(\nu_{i4}-\nu_{i3}). \quad T=4.$$

En la ecuación anterior cabe usar y_1 , y_2 como instrumentos.
En general, en la ecuación en diferencias para T = t se pueden usar los retardos de la variable Y hasta t-2.

Además la estructura del término de error en la ecuación en diferencias es

$$E[\Delta\nu_i\Delta\nu_i']=\sigma_\nu^{\ 2}(I_N\otimes G).$$

$$\underset{(T-2)x(T-2)}{G}=\begin{bmatrix} 2 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & & & \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{bmatrix}$$

<u>Siga</u>

Los instrumentos, en notación matricial son

$$\underset{(T-2)x(T-2)}{W_i} = \begin{bmatrix} y_{i1} & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & (y_{i1}\ y_{i2}) & \cdots & \cdots & & 0 \\ \vdots & \vdots & \ddots & & & \\ \vdots & \vdots & \ddots & & & \\ 0 & 0 & \cdots\cdots & & & (y_{i1}\ y_{i2} \cdots y_{T-2}) \end{bmatrix}.$$

Las restricciones de ortogonalidad son:

$$E\left[W'_i\, \Delta v_i\right] = 0.$$

Si aplicamos el método VI:

$$W'\Delta y = W'\Delta y_{-1}\delta + W'\Delta v.$$

Siga

Que tiene la matriz de varianzas y covarianzas antes mostrada. Aplicamos MCG y queda

$$\hat{\delta}_1 = \left[(\Delta y_{-1})'W(W'(I_N \otimes G)W)^{-1}W'(\Delta y_{-1})\right]^{-1}\left[(\Delta y_{-1})'W(W'(I_N \otimes G)W)^{-1}W'(\Delta y)\right].$$

Estimador inicial en AB(1991).
El estimador por MGM, Hansen(1982), es en este caso

$$V_N = \sum_{i=1}^{N} W'_i(\Delta v_i)(\Delta v_i)'W_i.$$

$$\hat{\delta}_2 = \left[(\Delta y_{-1})'W\hat{V}_N^{-1}W'(\Delta y_{-1})\right]^{-1}\left[(\Delta y_{-1})'W\hat{V}_N^{-1}W'(\Delta y)\right].$$

$$a\,\widehat{\mathrm{var}}\,\hat{\delta}_2 = \left[(\Delta y_{-1})'W\hat{V}_N^{-1}W'(\Delta y_{-1})\right]^{-1}.$$

Siga

Como se ve se ha sustituido la matriz de varianzas y covarianzas utilizada en la primera etapa por la “óptima” conforme al método MGM.

La matriz VN se obtiene con los residuos correspondientes a la primera estimación.

Ambos estimadores son equivalentes en el caso que se cumpla que el término de error genérico en la ecuación inicial, en niveles, tenga una distribución iid N(0, σ_{ν}^{2}).

Atrás

$$y_{it}^{+} = \left[\frac{T-t}{T-t+1} \right]^{1/2} \left[y_{it} - \frac{1}{T-t} \left(y_{i,t+1} + \ldots + y_{iT} \right) \right] \quad t = 1,2,\ldots T-1.$$

- El contraste se reduce al de H_0: $\gamma=0$.
- Debido a la presencia de heteroscedasticidad y autocorrelación de forma desconocida la matriz de varianzas y covarianzas a utilizar es la sugerida por White (1980).

Ahn, S.C. y P. Schmidt (1995), “Efficient Estimation of Models for Dynamic Panel Data”. *Journal of Econometrics*, 68, págs.5-28.

Anderson, T.W. y C. Hsiao (1981), “Estimation of Dynamic Models with Error Components”. *Journal of the American Statistical Association*, 76, págs. 598-606.

Anderson, T.W. y C. Hsiao (1982), “Formulation and Estimation of Dynamic Models using Panel Data”. *Journal of Econometrics*, 18, págs. 47-82.

Arellano, M.(1993), “On the testing of correlated effects with panel data”. *Journal of Econometrics*, 59, págs. 87-97.

Arellano, M. y S. Bond (1991), “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations”. *Review of Economic Studies*, 58, págs. 277-297.

Arellano, M. y O. Bover (1993), “Another Look at the Instrumental Variables Estimation of Error Components Models”. *Journal of Econometrics*, 68, págs. 29-52.

Arellano, M. y B. E. Honore (2002) “Panel Data: Some recent Developments”. En E. Leamer y J.J. Heckman (eds.) *Handbook of Econometrics*. North Holland.

Baltagi, B.H. (1995), *Econometric Analysis of Panel Data*. John Wiley.

Becker, G., M. Grossman y K. Murphy (1994), “An Empirical Analysis of Cigarrette Addiction”. *American Economic Review*, 84, págs. 396-418.

Blundell, R. y S. Bond (1998), “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models”. Journal of Econometrics, 87, págs. 115-143.

Bond, S. y C. Meghir (1994), “Dynamic Investment Models and the Firm´s Financial Policy”. *Review of Economic Studies*, 61, págs. 197-222.

Bover, O. y M. Arellano (1997), “Estimating dynamic limited dependent variable models from panel data”. *Investigaciones Económicas*, 21, págs. 141-165.

Chamberlain, G. (1984), “Panel data”. Cap. 22 en Z. Griliches y M. Intriligator (eds.) *Handbook of Econometrics*. North Holland. Págs. 1247-1318.

Holtz-Eakin, D. (1988), “Testing for Individual Effects in Autorregresive Models”. *Journal of Econometrics*, 39, págs. 297-307.

Holtz-Eakin, D.; W. Newey y H.S. Rosen (1988), “Estimating Vector Autorregresions with Panel Data”. *Econometrica*, 56, págs. 1371-1395.

Keane, M.P. y D.E. Runkle (1992), “On the Estimation of Panel-Data Models with Serial Correlation when Instruments are not Strictly Exogenous”. *Journal of Business and Economic Statistics*, 10, págs. 1-9, con discusión.

Nickell, S. (1981), “Biases in Dynamic Models with Fixed Effects”. *Econometrica*, 49, págs. 1417-1426.

White, H. (1986), “Instrumental variables analogs of generalized least squares estimators”. En R.S. Mariano (ed.) *Advances in Statistical Analysis and Statistical Computing*, vol. 1. JAI Press. Nueva York, págs. 173-277.