

métodos de estimación en modelos con componentes de varianza

nora garro

En este artículo nos referimos a modelos econométricos que hacen uso de datos que se presentan en forma de panel. Dicha información consiste en series de tiempo de un conjunto de datos de corte transversal: se observan N unidades o individuos durante T períodos de tiempo. Las unidades pueden ser empresas, consumidores, áreas geográficas, trabajadores, etc.

Este tipo de información crea importantes oportunidades en la econometría aplicada. En primer lugar, la reunión de datos de corte transversal y series de tiempo, permite contar con información "adicional" en el sentido que no nos restringe a una unidad ni a un período. En segundo lugar, podemos introducir adicionales fuentes de variación sistemática en la variable dependiente, distintas a las que explícitamente son mostradas por las variables independientes observadas. En particular, podemos distinguir variaciones cualitativas entre individuos y entre períodos.

Un modelo de regresión basado en datos presentados en forma de panel puede verse en la página siguiente, donde se consideran T períodos de tiempo, N individuos y K variables.

Las Y 's representan la variable dependiente observada cuyo comportamiento se quiere explicar con las fuentes de variación dadas en el lado derecho de la igualdad. Constituyen un vector de orden $NT \times 1$; el primer subíndice indica la unidad o individuo y el segundo indica el período.

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{1T} \\ Y_{21} \\ Y_{22} \\ \vdots \\ Y_{1T} \\ \vdots \\ Y_{N1} \\ Y_{N2} \\ \vdots \\ Y_{NT} \end{bmatrix} = \begin{bmatrix} X_{111} & X_{211} & \dots & X_{K11} \\ X_{112} & X_{212} & \dots & X_{K12} \\ \vdots & \vdots & & \vdots \\ X_{11T} & X_{21T} & \dots & X_{K1T} \\ X_{121} & X_{221} & \dots & X_{K21} \\ X_{122} & X_{222} & \dots & X_{K22} \\ \vdots & \vdots & & \vdots \\ X_{12T} & X_{22T} & \dots & X_{K2T} \\ \vdots & \vdots & & \vdots \\ X_{1N1} & X_{2N1} & \dots & X_{KN1} \\ X_{1N2} & X_{2N2} & \dots & X_{KN2} \\ \vdots & \vdots & & \vdots \\ X_{1NT} & X_{2NT} & \dots & X_{KNT} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix} + \begin{bmatrix} \alpha_1 \\ \alpha_1 \\ \vdots \\ \alpha_1 \\ \alpha_2 \\ \alpha_2 \\ \vdots \\ \alpha_2 \\ \vdots \\ \alpha_N \\ \alpha_N \\ \vdots \\ \alpha_N \end{bmatrix} + \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_T \\ \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_2 \\ \vdots \\ \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_T \end{bmatrix} + \begin{bmatrix} \eta_{11} \\ \eta_{12} \\ \vdots \\ \eta_{1T} \\ \eta_{21} \\ \eta_{22} \\ \vdots \\ \eta_{2T} \\ \vdots \\ \eta_{N1} \\ \eta_{N2} \\ \vdots \\ \eta_{NT} \end{bmatrix}$$

Las X 's representan las variables independientes observadas. Constituyen una matriz de orden $NT \times K$, donde K es el número de dichas variables usadas en el modelo. El primer subíndice de los X 's en la matriz indica la variable de que se trata (de 1 a K); el segundo la unidad o individuo (de 1 a N); el tercero el período (de 1 a T). Las X 's tienen un efecto sistemático sobre las Y 's medido por los β 's que son coeficientes desconocidos. En la mayoría de los estudios se supone que el coeficiente β para cada variable X es el mismo para todas las unidades y para todos los períodos. Entre las X 's puede haber una cuyo valor sea la unidad o bien una cuyo valor no varíe con el tiempo (como por ejemplo, raza u otro antecedente invariable) o bien una que sea la versión rezagada de la variable dependiente; por último las X 's pueden simplemente variar con el tiempo y las unidades. Asimismo consideramos a las variables X 's como variables estocásticas.

Las variables α 's representan el efecto invariante en el tiempo. Se suponen diferentes para cada individuo. Como ejemplo de tal influencia sistemática podemos citar una medida no observada de ambición o habilidad natural que se supone tiene influencia en la determinación del ingreso o bien puede significar habilidad empresarial la cual se supone tiene un efecto sistemático no observado en el producto de la firma.

Las variables λ 's se refieren al efecto invariante en los individuos. Son efectos cualitativos, sistemáticos y no observables que influyen en Y debido a las características de los períodos. Se supone que difieren para cada período. Por ejemplo, a través de λ podemos distinguir la variación en el ingreso Y debido a la depresión o al auge.

Por último, las variables η 's son los residuos no observados del análisis de regresión. O sea, son los efectos no sistemáticos debido a las variables exógenas no incluidas en la regresión. Como es usual en los modelos de regresión, se supone que las variables η_{it} independientes entre sí y que tienen la misma distribución normal con media cero y varianza σ^2_{η} .

El objetivo usual de los modelos econométricos suele ser la obtención de estimadores eficientes de los coeficientes β 's. En este artículo revisaremos los métodos de estimación que se han desarrollado para tomar en cuenta ese objetivo cuando existen efectos invariantes no observados tales como α y λ .

Básicamente se han desarrollado dos métodos de estimación, el del "efecto fijo" y el del "efecto aleatorio".

En el primero se considera que los α 's y los λ 's son parámetros desconocidos equivalentes a las β 's. Sin embargo, a diferencia de estos últimos, no interesa tanto la estimación de sus valores, sino la proporción de la variabilidad total de las Y 's que es explicada por la variabilidad de las α 's y las λ 's. Por ejemplo, nos interesa saber en que medida la diferencia en la habilidad natural entre las personas explica la diferencia en los ingresos. O sea, que $\sigma^2 \alpha$ y $\sigma^2 \lambda$ estimadores de la varianza de α y λ , respectivamente, proporcionan información más valiosa que $\hat{\alpha}_i$ y $\hat{\lambda}_t$, estimadores de α_i y λ_t , respectivamente.

El método del efecto fijo está asociado con el uso de las variables artificiales y con el análisis de varianza. En el primero los α 's y λ 's son estimados como coeficientes de variables artificiales y en el segundo no son estimados explícitamente pero son tratados como parámetros.

El método del efecto aleatorio considera los α 's y λ 's como variables aleatorias que representan cierta ignorancia al igual que el usual residuo η_{it} . Sin embargo, se trata de una ignorancia "específica" mientras que la que recogen los η_{it} es una ignorancia "general". De tal manera que el residuo del modelo puede descomponerse en tres componentes no observados α , λ y η . En este enfoque en lugar de estimar N α 's y T λ 's se estiman los llamados "componentes de la varianza". $\sigma^2 \alpha$ y $\sigma^2 \lambda$.

El método del efecto aleatorio está asociado con el método de mínimos cuadrados generalizado y el método de máxima verosimilitud.

En este artículo expondremos ambos procedimientos de estimación: el del efecto fijo y el del efecto aleatorio. Ocurre con frecuencia en este tipo de modelo la presencia de correlación entre las X 's y los componentes del residuo. Distinguiremos entre dos tipos de correlación:

- a) debido a la presencia entre las X 's de una variable dependiente rezagada la cual genera un tipo especial de correlación en serie;
- b) debido a la presencia de una o más variables entre las X 's que están correlacionadas con el efecto invariante en el tiempo.

Con objeto de simplificar la exposición no consideraremos el efecto invariante en los individuos λ , porque su inclusión complicaría la presentación sin añadir ninguna generalidad esencial.

Para el caso en que aparece la variable dependiente rezagada consideraremos el método de estimación con variables artificiales, el de mínimos cuadrados generalizado y el de variables instrumentales.

Para el caso de correlación entre X 's y α 's presentaremos el método de la regresión auxiliar y el de la variable instrumental.

Las propiedades de cada método de estimación para el caso de muestra pequeña y grande serán asimismo enunciados.

I. LA PRESENCIA DE UNA VARIABLE DEPENDIENTE REZAGADA

Consideremos el modelo

$$Y_{it} = X_{1t}\beta_1 + \dots + X_{kt}\beta_k + u_{it} \quad \begin{matrix} i = 1 \dots N \\ t = 1 \dots T \end{matrix}$$

donde una de las X 's es $Y_{i,t-1}$ y α es variable aleatoria que forma parte del residuo de u_{it} tal que

$$u_{it} = \alpha_i + \eta_{it} \quad \begin{matrix} i = 1 \dots N \\ t = 1 \dots T \end{matrix}$$

En este caso además del usual problema de correlación entre $Y_{i,t-1}$ y u_{it} a través de η_{it} aparece otra complicación adicional puesto que la covarianza entre el residuo corriente u_{it} y el residuo del período precedente $u_{i,t-1}$ para la misma unidad i no es cero como lo requiere el modelo de regresión simple sino que vale $\sigma^2 \alpha$. Ello implica un problema de autocorrelación de residuos y una causa más para sostener que $Y_{i,t-1}$ está correlacionada con u_{it} a través de su relación con $u_{i,t-1}$.

Las soluciones a este problema son:

- a) Tratar los α 's como parámetros a estimar. O sea, considerarlos coeficientes de variables artificiales y no como parte del residuo.

- b) Reemplazar la variable $Y_{i,t-1}$ por una variable instrumental
- c) Usar una transformación de variables (MCG).

A. Métodos de estimación: variable artificial

El modelo a usar es el siguiente

$$Y = X\beta + D'\alpha + \eta$$

donde X es una matriz de orden $NT \times K$ equivalente a la presentada anteriormente con la salvedad de que una de sus columnas está destinada a la variable Y_{it-1} . O sea.

$$X = \{Y_{-1} : \tilde{X}\}$$

$$NT \times K \quad NT \times 1 \quad NT \times (K-1)$$

β es un vector de coeficientes de orden $K \times 1$, α es un vector de orden $N \times 1$ y por otro lado, D es una matriz de orden $NT \times N$

$$D = \begin{bmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} & & 0 \\ & \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} & \\ & \ddots & \\ & 0 & \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \end{bmatrix}$$

$T \times 1$

donde cada columna D_i representa una variable artificial tal que

$$D_i = \begin{cases} 1 & \text{para la unidad } i \\ 0 & \text{para otras unidades} \end{cases}$$

Por lo tanto, el modelo así constituido permite la obtención de los siguientes estimadores

$$\hat{\beta} = (X'X)^{-1} X'Y$$

$$\hat{\alpha}_i = Y_{i.} - \hat{\beta} X_{i.} \quad i = 1, \dots, N$$

$$\hat{\sigma}^2_{\eta} = \left\{ \sum_{i,t} (Y_{it} - Y_{i.})^2 - \frac{\sum_{k,i,t} (X_{kit} - X_{ki.})^2 (Y_{it} - Y_{i.})}{\sum_{k,i,t} (X_{kit} - X_{ki.})^2} \right\} : NT - N - K$$

Para obtener un estimado de $\sigma\alpha$ utilizamos la fórmula para la varianza de $\hat{\sigma}$ sugerida por Nerlove (ver bibliografía).

$$\hat{\sigma}^2_{\alpha} = \frac{\sum_{i=1}^N (\hat{\alpha}_i - \sum_{i=1}^N \hat{\alpha}_i / N)^2}{N}$$

B. Método de la variable instrumental

Mediante este procedimiento utilizamos una variable instrumental en reemplazo de la variable rezagada $Y_{i,t-1}$. Dicha variable debe estar correlacionada con $Y_{i,t-1}$ pero no correlacionada con los residuos $u_{i,t}$.

En este caso la selección de la variable instrumental no es difícil porque la variable rezagada es una función de las variables independientes rezagadas.

Consideremos el modelo más simple con una variable independiente y la variable dependiente rezagada en un período. Así.

$$Y_{it} = \beta X_{it} + \gamma Y_{i,t-1} + \alpha_i + \eta_{it} \quad \begin{matrix} i = 1 \dots N \\ t = 1 \dots T \end{matrix}$$

La variable dependiente rezagada puede ser escrita como

$$Y_{i,t-1} = \beta X_{i,t-1} + \gamma Y_{i,t-2} + \alpha_i + \eta_{i,t-1}$$

De tal modo que la ecuación del modelo también es

$$Y_{it} = \beta X_{it} + \gamma (\beta X_{i,t-1} + \gamma Y_{i,t-2} + \alpha_i + \eta_{i,t-1}) + \alpha_i + \eta_{it}$$

Si consideramos el proceso de sustitución de $Y_{t-\tau}$ versiones rezagadas de Y terminaremos con la siguiente expresión:

$$\begin{aligned} Y_{i,t-1} &= \beta X_{i,t-1} + \gamma \beta X_{i,t-2} + \gamma^2 \beta X_{i,t-3} + \dots + \omega \\ &= \sum_{\tau=1}^{\infty} \delta_{\tau} X_{i,t-\tau} + \omega \end{aligned}$$

donde W es una combinación lineal de los sucesivos residuos rezagados. Como usualmente no se cuenta con muchas versiones rezagadas de las X 's, podemos restringir los rezagos a sólo un período. O sea, considérese

$$Y_{i,t-1} = \delta_1 X_{i,t-1} + \omega$$

donde los otros términos de la suma son desechados.

A esta última expresión se le puede aplicar mínimos cuadrados ordinarios con el fin de obtener un estimado consistente de δ_1 . Este estimador no es efi-

ciente porque la matriz de varianza-covarianza de los residuos presentará heteroscedasticidad pero en este caso sólo requerimos consistencia.

De tal manera, la variable instrumental para $Y_{i,t-1}$ es

$$Y_{i,t-1} = \hat{\delta}_1 X_{i,t-1}$$

El modelo original se ha transformado en

$$Y_{it} = \beta X_{it} + \gamma \hat{Y}_{i,t-1} + \alpha_i + \eta_{it}$$

$$Y_{it} = \beta X_{it} + \gamma \hat{Y}_{i,t-1} + u_{it}$$

A partir de este modelo podemos obtener estimadores de β , γ , σ^2_α y σ^2_η aplicando mínimos cuadrados ordinarios. Cabe aclarar que $\hat{\beta}$ y $\hat{\gamma}$ no serán eficientes debido a la presencia de heteroscedasticidad, pero sí serán consistentes.

Si suponemos que la esperanza de α es cero, $E(\alpha) = 0$, y que las α 's no están correlacionadas con las η 's, entonces

$$\begin{aligned} \hat{\sigma}^2_\alpha + \hat{\sigma}^2_\eta &= \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \hat{u}_{it}^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (\alpha_i + \eta_{it})^2 \\ &= \frac{1}{N} \sum_{i=1}^N \alpha_i^2 + \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \eta_{it}^2 \xrightarrow{P} \sigma^2_\alpha + \sigma^2_\eta \end{aligned}$$

Para obtener $\hat{\sigma}^2_\alpha$ consideremos el hecho de que para una unidad dada i

$$\text{COV}(u_{it}, u_{it'}) = E(\alpha_i + \eta_{it})(\alpha_i + \eta_{it'}) = \sigma^2_\alpha$$

Entonces

$$\hat{\sigma}_{\alpha}^2 = \frac{1}{T(T-1)} \sum_{t=1}^T \sum_{t \neq t'}^T (\hat{u}_{it} \hat{u}_{it'})$$

Algunos autores, como Maddala, brinda una fórmula un tanto más complicada que toma en cuenta los diferentes $\hat{\sigma}_{\alpha_i}$ que pueden ser obtenidos para cada unidad. El resultado es el estimador

$$\hat{\sigma}_{\alpha}^2 = \frac{\sum_{i=1}^N \left[\left(\sum_{t=1}^T u_{it} \right)^2 - \sum_{t=1}^T u_{it}^2 \right]}{NT(T-1)} = \frac{\sum_{i=1}^N \hat{\sigma}_{\sigma_i}^2}{N} \xrightarrow{P} \sigma_{\alpha}^2$$

el cual representa un promedio de los $\hat{\sigma}_{\alpha_i}$ obtenidos anteriormente. Este estimador resulta ser el estimador de máxima verosimilitud para $\hat{\sigma}_{\alpha}$, pero representa la desventaja de que puede resultar negativo.

C. Método de los mínimos cuadrados generalizados

Al igual que en el método anterior en este caso se considera que el efecto α es parte de los residuos, variables aleatorias no observadas con efecto sistemático en Y. los supuestos en relación a los α 's son

$$\begin{aligned} E(\alpha) &= 0 \\ E(\alpha_i \alpha_j) &= \begin{cases} \sigma_{\alpha}^2 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \end{aligned}$$

En relación a los η_{it} los supuestos son

$$E(\eta_{it}) = 0$$

$$E(\eta_{it} \eta_{i't'}) = \begin{cases} \sigma^2_{\eta} & \text{si } i = i' \text{ y } t = t' \\ 0 & \text{en otra situación} \end{cases}$$

Además

$$E(\alpha_i \eta_{it}) = 0 \quad \forall i, t$$

Por lo anterior, la matriz de varianza-covarianza de los residuos

$$u_{it} = \alpha_i + \eta_{it}$$

tendrá la siguiente forma:

$$Euu' = \begin{pmatrix} \sigma^2_{\alpha} + \sigma^2_{\eta} & \sigma^2_{\alpha} & \dots & \sigma^2_{\alpha} \\ \sigma^2_{\alpha} & \sigma^2_{\alpha} + \sigma^2_{\eta} & \dots & \sigma^2_{\alpha} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2_{\alpha} & \dots & \sigma^2_{\alpha} + \sigma^2_{\eta} & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma^2_{\alpha} & \dots & \sigma^2_{\alpha} + \sigma^2_{\eta} & \dots & \sigma^2_{\alpha} + \sigma^2_{\eta} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma^2_{\alpha} & \dots & \sigma^2_{\alpha} + \sigma^2_{\eta} & \dots & \sigma^2_{\alpha} + \sigma^2_{\eta} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma^2_{\alpha} & \dots & \sigma^2_{\alpha} + \sigma^2_{\eta} & \dots & \sigma^2_{\alpha} + \sigma^2_{\eta} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma^2_{\alpha} & \dots & \sigma^2_{\alpha} + \sigma^2_{\eta} & \dots & \sigma^2_{\alpha} + \sigma^2_{\eta} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma^2_{\alpha} & \dots & \sigma^2_{\alpha} + \sigma^2_{\eta} & \dots & \sigma^2_{\alpha} + \sigma^2_{\eta} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma^2_{\alpha} & \dots & \sigma^2_{\alpha} + \sigma^2_{\eta} & \dots & \sigma^2_{\alpha} + \sigma^2_{\eta} \end{pmatrix}$$

(T x T)

(NT x NT)

Nótese que en el modelo generalizado y en general en cualquier procedimiento donde los α 's sean considerados parte de los residuos de la regresión, aparece el problema de la heteroscedasticidad. O sea, la matriz de varianza-covarianza no es una matriz diagonal, lo cual es un requerimiento de los mínimos cuadrados ordinarios. En este caso el problema se origina porque la covarianza entre los residuos correspondientes a una unidad particular no es cero sino $\hat{\sigma}\alpha$.

Pero el método generalizado da una solución a la heteroscedasticidad (no así el de la variable instrumental) porque la matriz de varianza-covarianza de los residuos transformados es una matriz diagonal (situación de homoscedasticidad). Así, el método generalizado consiste en aplicar mínimos cuadrados ordinarios a los datos transformados.

Adicionalmente, se elimina el problema de la autocorrelación en serie de los residuos debido a la presencia de la variable rezagada porque la covarianza entre residuos se transforma en cero para la misma unidad.

Analicemos la transformación de variables. Cada bloque de la matriz puede ser llamado Ω y es de orden $T \times T$. Así.

$$E(uu') = \sigma^2 \left\{ \begin{array}{ccc} \Omega & & 0 \\ & \ddots & \\ 0 & & \Omega \end{array} \right\} \quad \text{donde } \Omega = \left\{ \begin{array}{ccc} 1 & & \rho \\ & \ddots & \\ \rho & & 1 \end{array} \right\}$$

donde

$$\sigma^2 = \sigma^2_{\alpha} + \sigma^2_{\eta} \quad \text{y} \quad \rho = \frac{\sigma^2_{\alpha}}{\sigma^2_{\alpha} + \sigma^2_{\eta}}$$

Nótese que Ω puede ser escrito como

$$\Omega = (1-\rho) I_t + \rho ee'$$

donde $e_{T \times 1}$ es un vector de unos, (I_T) es la matriz unidad de orden $T \times T$

Puede demostrarse que

$$\Omega^{-1} = \lambda_1 e e' + \lambda_2 I_T$$

donde

$$\lambda_1 = - \frac{\rho}{(1-\rho)(1-\rho + T\rho)} \quad \lambda_2 = \frac{1}{(1-\rho)}$$

La transformación de variables consiste en multiplicar cada observación por $\Omega^{-1/2}$. Hecho esto podemos aplicar mínimos cuadrados ordinarios a los nuevos datos y obtener.

$$\hat{\beta} = \left[\sum_{i=1}^N [X_i' \Omega^{-1} X_i] \right]^{-1} \left[\sum_{i=1}^N X_i' \Omega^{-1} Y_i \right]$$

donde

$$\begin{aligned} \sum_{i=1}^N X_i' \Omega^{-1} X_i &= \lambda_1 \sum_{i=1}^N X_i' e e' X_i + \lambda_2 \sum_{i=1}^N X_i' X_i \\ \sum_{i=1}^N X_i' \Omega^{-1} Y_i &= \lambda_1 \sum_{i=1}^N X_i' e e' Y_i + \lambda_2 \sum_{i=1}^N X_i' Y_i \end{aligned}$$

El estimador de los β 's así obtenidos resulta ser el estimador de máxima verosimilitud para los β 's. O sea, posee la propiedad de consistencia, es asintóticamente normal y eficiente.

El método de los mínimos cuadrados generalizados requiere de dos etapas.

En la primera requiere una estimación de σ_α y σ_η con el objeto de tener un estimador de Ω , $\hat{\Omega}$, para llevar a cabo la transformación de variables. En la segunda etapa se obtienen los estimadores de los β 's. Cabe señalar que las propiedades asintóticas de los β 's, con Ω o con $\hat{\Omega}$ son las mismas cuando se utilizan estimadores consistentes para σ_α^2 y σ_η^2 en la segunda etapa. Al respecto, el método de las variables artificiales o el de variable instrumental pueden ser utilizados en la primera etapa para estimar σ_α^2 y σ_η^2 .

D. Evaluación de los Métodos

Con el objeto de comparar el modelo de variables artificiales con el de mínimos cuadrados generalizados es conveniente expresar los estimadores de este último método en forma diferente.

Consideremos

$$T_{xx} = \sum_i^N X_i' X_i$$

$$W_{xx} = \sum_{i=1}^N \left[X_i' \left(I_t - \frac{ee'}{t} \right) X_i \right]$$

$$B_{xx} = \sum_{i=1}^N X_i' ee' X_i$$

donde X_i es una matriz de orden $T_x K$.

Estos elementos son, respectivamente, la matriz de variación total de las X 's (T_{xx}), la matriz de la suma de cuadrados y suma de productos entre unidades (B_{xx}) y la matriz de suma de cuadrados y suma de productos dentro de las unidades (W_{xx}). Se pueden obtener expresiones semejantes para T_{xy} , B_{xy} y W_{xy} .

Tras una serie de manipulaciones algebraicas se puede llegar a la siguiente expresión para los estimadores MCG del método de los mínimos generalizados.

$$\hat{\beta}_{\text{MCG}} = \left[W_{xx} + \theta B_{xy} \right]^{-1} \left[W_{xy} + \theta B_{xy} \right]$$

donde

$$\theta = 1 + \frac{\lambda_1 T - 1 - \rho}{\lambda_2} = \frac{1 - \rho}{1 - \rho + \rho T} = \frac{\sigma_\eta^2}{\sigma_\eta^2 + T \sigma_\alpha^2}$$

Nótese que W_{xx} y W_{xy} dan los estimadores $\hat{\beta}_{VA}$'s del método de variables artificiales y B_{xx} y B_{xy} dan los estimadores $\hat{\beta}$'s de mínimos cuadrados ordinarios. Por lo tanto, los estimadores $\hat{\beta}_{\text{MCG}}$ representan un compromiso entre los estimadores del método de las variables artificiales y el método de mínimos cuadrados ordinarios. La ponderación está dada por el parámetro θ , que es una función de los componentes de la varianza y de T .

Numerosos autores (Amemiya, Maddala, Nerlove, Baltaghi, entre otros) han demostrado a través de estudios Monte Carlo para muestras pequeñas, que el método de las variables artificiales producen sesgos hacia abajo para el estimador del coeficiente de la variable y sesgos hacia arriba para el resto de los β 's. Dichos sesgos se reducen cuando T y la diferencia entre unidades ($\sigma^2 \alpha$) aumenta. Asimismo, para muestras pequeñas el método de mínimos cuadrados generalizado, en cuya primera etapa se utilice el método de variables artificiales, tiene el menor sesgo y el menor error cuadrado medio. Por su parte, Taylor demostró que para muestras pequeñas, la varianza de los β 's (con $\sigma^2 \alpha$ y $\sigma^2 \eta$ estimados en una primera etapa) aumentan no más del 17% por encima del límite inferior Cramer-Rao. La cifra cae drásticamente cuando aumenta N ó T .

Consideremos el caso en que aumenta el tamaño de la muestra, por incrementos en N o en T o en ambos.

Supongamos que N está fijo y T tiende a infinito. Entonces θ tiende a cero y $\hat{\beta}_{\text{MCG}} = \hat{\beta}_{VA}$. Este resultado no es sorprendente pues el sentido común nos

dice que cuando poseemos gran cantidad de información para cada individuo a través del tiempo, la variabilidad entre dichos individuos tiende a dominar los estimadores. En la práctica, sin embargo, no es éste el caso usual porque T está limitado y es generalmente pequeño ($N > T$).

Si T está fijo y N tiende a infinito, se pierde el interés específico en cada α_i pero los α 's adquieren relevancia como valores de una variable aleatoria. En ese sentido nos interesan más las características poblacionales de los α 's y el método de mínimos cuadrados generalizados es el recomendable.

Si ambos, N y T , tienden a infinito, Amemiya ha demostrado que tanto el método de las variables artificiales como el de mínimos cuadrados generalizados dan el mismo resultado. O sea, que para este tipo de muestra grande, el tratamiento de α como variable aleatoria o como parámetro fijo son esencialmente equivalentes. En términos prácticos, sin embargo, el uso de variables artificiales puede implicar demasiadas variables adicionales.

Con respecto al método de la variable instrumental, los estudios Monte Carlo demostraron que las propiedades de los estimadores son erráticas. Para muestras grandes se descubrió que para valores poblacionales de β diferentes de cero, los estimadores resultaron consistentes pero cuando son diferentes de cero, hay inconsistencia. Además, los estimadores β son asintóticamente menos eficientes a medida que los valores poblacionales de β son menores.

Podemos hacer inferencias adicionales sobre la pertinencia de los métodos, si suponemos diferentes valores de θ . Así, para N y T fijos, a medida que disminuye la proporción del total de la varianza debida a los efectos individuales α_i , entonces.

$$\rho \longrightarrow 0$$

$$\theta \longrightarrow 1$$

y

$$\hat{\beta}_{\text{MCG}} = T_{xx}^{-1} T_{xy}$$

o sea los estimadores de método de mínimos cuadrados generalizados se aproximan a los estimadores del método de mínimos cuadrados ordinarios.

En el caso en que la varianza total se debe principalmente al componente

individual σ_{α_i} entonces.

$$\rho \longrightarrow 1$$

$$\theta \longrightarrow 0$$

y el método de mínimos cuadrados generalizados se iguala con el método de las variables artificiales. O sea,

$$\hat{\beta}_{\text{MCG}} = \hat{\beta}_{\text{VA}}$$

A modo de información diremos que el método de máxima verosimilitud puede ser utilizado en modelos con una variable rezagada. Fijando valores para ρ entre cero y uno se logran resultados mediante procesos iterativos. La dificultad consiste en que las ecuaciones normales son altamente no lineales para β , σ^2_{α} y σ^2_{η} . Además, Nerlove encontró un comportamiento pobre de este método en el caso de muestras pequeñas.

A partir del análisis realizado a lo largo de este punto, podemos concluir que en modelos económicos cuyos datos se presentan en forma de panel, o sea, una serie de tiempo de series de corte transversal, el método más eficaz para estimar componentes de la varianza es el de mínimos cuadrados generalizados, puesto que las propiedades de sus estimadores son buenas, tanto en pequeñas como en grandes muestras (no sesgo y consistencia). Para la primera etapa se recomienda el uso del método de variables artificiales ya que produce estimadores consistentes de los componentes de la varianza.

II. LA PRESENCIA DE CORRELACION ENTRE LAS X's Y EL EFECTO INVARIANTE EN EL TIEMPO

Consideremos las variables X's como variables estocásticas tales que la matriz de variables independientes representa los valores observados en la muestra. Asimismo tratemos al efecto α como una variable aleatoria.

En este caso, puede suceder que entre algunas X 's y α exista correlación. Por ejemplo, en una ecuación de salarios, la variable X_k referida a años de escolaridad puede estar correlacionada con α , habilidad natural; o bien en una función de producción X_k , tamaño de la firma, puede estar correlacionada con α , habilidad empresarial.

Los métodos de estimación disponibles en este caso son:

1. Usar una ecuación auxiliar que tome en cuenta la correlación y aplicar el método de mínimos cuadrados generalizados.
2. Usar variables instrumentales para las X 's correlacionadas con α y aplicar el método de mínimos cuadrados generalizados.
3. Usar variables artificiales

Presentaremos los dos primeros procedimientos dado que el tercero es similar al que se presentó en el punto anterior.

A. Ecuación auxiliar

La ecuación del modelo es, como antes

$$Y = D'\alpha + X\beta + \eta$$

a la cual agregamos la relación explícita entre las X 's y α

$$D'\alpha = K(X\pi + \omega)$$

donde π es un vector de orden $k \times 1$ y ω representa el vector de residuos de orden $NT \times 1$. Dichos ω 's se suponen independientes e igualmente distribuidos

$$\omega_{it} \sim N(0, \sigma^2 \omega)$$

la matriz K es de orden $NT \times NT$ y tiene la siguiente estructura:

$$K = \left(\begin{array}{ccc} \begin{pmatrix} 1/T & \dots & 1/T \\ \vdots & & \vdots \\ 1/T & \dots & 1/T \end{pmatrix} & 0 & \\ & (T \times T) \begin{pmatrix} 1/T & \dots & 1/T \\ \vdots & & \vdots \\ 1/T & \dots & 1/T \end{pmatrix} & \\ & & \ddots \\ & 0 & \begin{pmatrix} 1/T & \dots & 1/T \\ \vdots & & \vdots \\ 1/T & \dots & 1/T \end{pmatrix} \end{array} \right)$$

(NT x NT)

o sea

$$K = I_n \oplus J_T$$

donde

$$J_T = \frac{\mathbf{e} \mathbf{e}'}{T}$$

siendo I_N una matriz unidad de orden $N \times N$ y e_T un vector unidad de orden $T \times 1$. Nótese que $K = K^2$.

Reemplazado α en la primera ecuación del modelo obtenemos

$$Y = X\beta + K(X\pi + \omega) + \eta$$

donde el residuo es $\epsilon = K\omega + \eta \sim (0, T\sigma^2_\omega K + \sigma^2_\eta I_{Nt}) \sim (0, \tilde{\Omega})$

La matriz de varianza-covarianza no será diagonal y presentará bloques para cada unidad i de la forma

$$\left(\begin{array}{cccc} \sigma^2_\alpha + \sigma^2_\omega & \sigma^2_\omega & \dots & \sigma^2_\omega \\ \sigma^2_\omega & \sigma^2_\alpha + \sigma^2_\omega & \dots & \sigma^2_\omega \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2_\omega & \sigma^2_\omega & \dots & \sigma^2_\alpha + \sigma^2_\omega \end{array} \right)$$

(T x T)

A partir de aquí puede visualizarse con claridad entre modelos que consideren α como variable aleatoria o como efecto fijo. Así

$$E(Y/X) = X\beta + (KX)\pi = X(\beta + K\pi)$$

en el modelo donde α es considerado variable aleatoria parte de los residuos y

$$E(Y/X) = X\beta + D'\alpha$$

es el modelo donde α es considerado como efecto fijo.

Como ya dijimos al comienzo de este punto, hablamos del primer modelo.

B. Método de los mínimos cuadrados generalizados con la ecuación auxiliar

Como se recordará, el método de los mínimos cuadrados generalizado consiste en realizar un cambio de variables utilizando estimadores de σ^2_η y σ^2_ω obtenidos en una primera etapa. Con dichos estimadores puede calcularse

$$\tilde{\Omega} = T\sigma^2_\omega K + \sigma^2_\eta I_{NT}$$

y

$$\tilde{\Omega}^{-1} = \gamma_1 TK + \gamma_2 I_{NT}$$

donde

$$\gamma_1 = - \frac{\sigma^2_\omega}{\sigma^2_\eta (\sigma^2_\eta + T\sigma^2_\omega)} \quad \gamma_2 = \frac{1}{\sigma^2_\eta}$$

Un método disponible para la primera etapa es el de variables artificiales.

Aquí presentaremos el método de estimación intra-grupo como una interesante alternativa para obtener

$$\begin{aligned} (Y_{it} - Y_{i\cdot}) &= (X_{it} - X_{i\cdot})\beta + (\alpha_i - \alpha_\cdot) + (\eta_{it} - \eta_{i\cdot}) \\ &= (X_{it} - X_{i\cdot})\beta + (\eta_{it} - \eta_{i\cdot}) \end{aligned}$$

$$i = 1 \dots N$$

$$t = 1 \dots T$$

El punto en los subíndices indica que se ha tomado el promedio en el tiempo de los valores de la variable correspondiente.
donde los α 's desaparecen puesto que

$$\alpha_i = \alpha \quad \forall i$$

De tal manera que se puede aplicar mínimos cuadrados ordinarios y obtener

$$\hat{\sigma}^2_{\eta} = \frac{1}{NT - N} \sum_i \sum_t (\eta_{it} - \eta_{i\cdot})^2$$

Este estimador es consistente pues

$$\begin{aligned} \text{plim}_{N \rightarrow \infty} \hat{\sigma}^2_{\eta} &= \text{plim}_{N \rightarrow \infty} \frac{1}{N(T-1)} \sum_i \sum_t (\eta_{it} - \eta^2_{it\cdot} - 2\eta_{it} \eta_{i\cdot}) = \\ &= \frac{1}{N(T-1)} NT \sigma^2_{\eta} + \text{plim}_{N \rightarrow \infty} \frac{1}{N(T-1)} \sum_i \sum_t \frac{\eta^2_{it}}{T^2} + \text{plim}_{N \rightarrow \infty} \frac{1}{N(T-1)} 2 \end{aligned}$$

$$\begin{aligned} \sum_i \sum_t \frac{\eta^2_{it}}{T} &= \frac{1}{N(T-1)} [NT \sigma^2_{\eta} + \frac{NT^2}{T^2} \sigma^2_{\eta} + \frac{NT}{T} \sigma^2_{\eta}] \\ &= \frac{N(T-1) \sigma^2}{N(T-1)} = \sigma^2_{\eta} \end{aligned}$$

Con el objeto de obtener un estimador consistente para σ^2_{ω} utilizaremos el modelo inter-grupo

$$\begin{aligned}
 Y_{i\cdot} &= \sum_{s=1}^K X_{si\cdot} \beta_s + \alpha_i + \eta_{i\cdot} \\
 &= \sum_{s=1}^K X_{si\cdot} \beta_s + \sum_{s=1}^K X_{si\cdot} \pi_s + (\omega_{i\cdot} + \eta_{i\cdot}) \\
 &= \sum_{s=1}^K X_{si\cdot} (\beta_s + \pi_s) + (\omega_{i\cdot} + \eta_{i\cdot}) \\
 &\quad i = 1 \dots N
 \end{aligned}$$

del cual a partir del método de los mínimos cuadrados ordinarios puede obtenerse un estimador de los parámetros $(\beta_s + \pi_s)$. Asimismo podemos calcular

$$\begin{aligned}
 S^2 &= \frac{\sum_{i=1}^N \left[Y_{i\cdot} - \sum_{s=1}^K X_{si\cdot} (\beta_s + \pi_s) \right]^2}{N} = \frac{\sum_{i=1}^N (\omega_{i\cdot} + \eta_{i\cdot})^2}{N} \\
 &= \frac{\sum_{i=1}^N \hat{\omega}_{i\cdot}^2}{N} = \frac{\sum_{i=1}^N \sum_{T=1}^T \hat{\eta}_{it}^2}{NT^2}
 \end{aligned}$$

que tiende en probabilidad a

$$\sigma^2_{\omega} + \frac{\sigma^2_{\eta}}{T}$$

Entonces un estimador consistente para σ^2_{ω} será

$$\hat{\sigma}^2_{\omega} = S^2 - \frac{\hat{\sigma}^2_{\eta}}{T}$$

Con los estimadores $\hat{\sigma}^2_{\eta}$ y $\hat{\sigma}^2_{\omega}$ obtenidos en la primera etapa podemos desarrollar la segunda etapa que brindará los siguientes resultados.

$$\hat{\beta}_{MCG} = (X' \hat{\Omega}^{-1} X)^{-1} X' \hat{\Omega}^{-1} Y$$

$$\hat{\pi}_{MCG} = (X' K \hat{\Omega}^{-1} K X)^{-1} X' K \hat{\Omega}^{-1} Y$$

Por lo tanto, la estimación de los coeficientes de las variables X's en el modelo de regresión con correlación entre las variables independientes y los residuos es igual a

$$(\hat{\beta}_{MCG} + \hat{\pi}_{MCG})$$

Es interesante destacar que para muestras grandes, Mundlack demostró que

$$\hat{\beta}_{MCG} = \hat{\beta}_{\omega}$$

siendo $\hat{\beta}_{\omega}$ el vector de coeficientes de las X's cuando se utiliza el método intra-grupo de estimación y

$$\hat{\pi}_{MCG} = \hat{\beta}_B - \hat{\beta}_W$$

donde $\hat{\beta}_B$ es el vector de coeficientes estimados de las X's cuando se utiliza el método inter-grupo.

Por lo tanto, el coeficiente $\hat{\beta}_B$ es sesgado pues

$$E(\hat{\beta}_B) = E(\hat{\beta}_W + \hat{\pi}) = \beta + \pi$$

Siendo el método de los mínimos cuadrados generalizado un compromiso entre dos métodos inter- e intra- grupo se sigue que si se ignora el problema de la correlación en la segunda etapa, los estimadores de los coeficientes de X's serán sesgados, puesto que $\hat{\beta}_B$ lo es.

C. Variable instrumental en la segunda etapa del método de mínimos cuadrados generalizados

Mediante este procedimiento, se realiza la transformación de variables multiplicando cada observación por $\Omega^{-1/2}$ o más comúnmente $\hat{\Omega}^{-1/2}$, que como hemos visto se obtiene con estimadores de σ^2_η y σ^2_α

De tal manera que al finalizar la primera etapa el modelo es

$$\hat{\Omega}^{-1/2} Y_{it} = \hat{\Omega}^{-1/2} X_{it} \beta + \hat{\Omega}^{-1/2} \alpha_i + \hat{\Omega}^{-1/2} \eta_{it} \quad \begin{matrix} i = 1 \dots N \\ t = 1 \dots T \end{matrix}$$

donde no existe el problema de heteroscedasticidad provocado en el modelo original por la presencia del efecto invariante en el tiempo.

Sin embargo, subsiste el problema de la correlación entre X's y α 's. Para solucionarlo pueden utilizarse variables instrumentales. Si todas las X's estuvieran correlacionadas con los α 's, se requerirían variables instrumentales externas. Pero no es el caso común. Usualmente, podemos usar las X's no correlacionadas como instrumentos para las correlacionadas.

Supongamos que del total K de variables independientes, existen K_1 , que pueden ser consideradas como instrumentos. Si formamos la submatriz X^1 de

orden $NT \times K$, se cumple que

$$\text{plim}_{N \rightarrow \infty} X' \alpha = \{0\}_{K' \times 1}$$

A continuación planteamos las regresiones de cada una de las restantes $K_2 = K - K_1$ variables independientes con las K_1 variables instrumentales. Así

$$X_j = X^1 \psi_j + \zeta$$

donde X_j es un vector de orden $NT \times 1$ para la variable X_j correlacionada con α ; ψ es un vector de orden $K_1 \times 1$ que representa las pendientes de la regresión y ζ es un vector de orden $NT \times 1$ que representa los residuos. Hay j regresiones de este tipo para $j = 1, 2, \dots, K_2$

Aplicando mínimos cuadrados ordinarios obtendremos

$$\hat{X}_j = X^{1'} \hat{\psi}_j$$

que es el instrumento para las K_j , $j = 1 \dots K_2$

Por lo tanto, el modelo para la segunda etapa resulta

$$\hat{\Omega}^{-1/2} Y_{it} = \hat{\Omega}^{-1/2} X^1 \beta^1 + \hat{\Omega}^{-1/2} \hat{X}^2 \beta^2 + \hat{\Omega}^{-1/2} \alpha + \hat{\Omega}^{-1/2} \eta$$

donde X^2 es la matriz de orden $NT \times K_2$ de las variables correlacionadas estimadas; β^1 y β^2 vectores de orden $K_1 \times 1$ y $K_2 \times 2$ de las pendientes de la regresión.

Si se aplica el método de mínimos cuadrados ordinarios se obtienen estimadores eficientes para los β 's.

Recapitulando sobre el problema de la correlación entre algunas X 's y α 's, el método de regresión más apropiado para enfrentarlo es el de los mínimos cuadrados generalizados, siempre que se obtengan estimadores consistentes en la primera etapa. Tanto la primera como la segunda etapa deben de tener

en cuenta el problema de la correlación. Para ello se utiliza el método de la regresión auxiliar o el de la variable instrumental.

Cabe señalar que las propiedades de los estimadores del método de mínimos cuadrados generalizados para muestras pequeñas hasta 1980 no se habían aún estudiado ■

BIBLIOGRAFIA

* Son referencias claves

- * 1.- Amemiya, T., *The estimation of the variances in a variance-component model*. International Economic Review 12, 1971, pp. 1-13.
- 2.- Amemiya, T., *A note on the estimation of Balestra-Nerlove Models*. Technical Report No. 4. Institute for Mathematical Studies in the Social Sciences, Stanford University, agosto 14, 1967.
- * 3.- Baltagi, Badi, *An experimental study of alternative testing and estimation procedures in a two-way error component model*. Journal of Econometrics 17, 1981, pp. 21-49.
- 4.- Bass, F. and Wittink, D., *Pooling issues and methods in regression analysis, some further reflections*. Journal of Marketing Research, vol. XV, mayo 1978, pp. 277-279.
- 5.- Fuller, W. A. and Battese, G. E., *Estimation of linear models with crossed-error structure*. Journal of Econometrics 2, 1974, pp. 67-78.
- * 6.- Griliches, Z., *Estimating the returns to schooling: some econometric problems*. Econometrica, 45, 1977, pp. 1-22.
- * 7.- Hausman, J. and Taylor, W., *Panel data and unobservable individual effects*. Econometrica, vol. 49, No. 6, nov. 1981.
- 8.- Henderson, Ch., *Comment on "The use of error components models in combining cross section with time series data"*. Econometrica, vol. 39, No. 2, marzo, 1971, pp. 397-401.
- 9.- Henderson, C., *Estimation of variance and covariance components*. Biometrics, 9, junio. 1943, pp. 226-252.
- 10.- Houthakker, P. K. and Verleger, Jr. and Sheehan, D. P. *Dynamic demand analysis for gasoline and residential electricity demand*. American Journal of Agricultural Economics, mayo 1974, pp. 412-418.
- * 11.- Izan, Haji y. *To pool or not to pool?* Journal of Econometrics 13, 1980, pp. 391-402.
- 12.- Johnson and Oksanen, *Estimation of the demand for alcoholic beverage*. Review of Economics and Statistics, 1977, 59, pp. 113-117.

- 13.- Maddala, C., *The use of variance components models in pooling cross section and time-series data*. Econometrics, marzo 1971, pp. 341-358.
- * 14.- Maddala, G. S. and Mount, T. D., *A comparative study of alternative estimators for variance components models*. Journal of the American Statistical Association, junio, 1973, pp. 324-328.
- 15.- Maddala, G. S., *Econometrics*. McGraw-Hill, Inc., 1977 Ch. 14.
- * 16.- Mundlak, Y., *On the pooling of time series and cross section data*. Econometrica, 1978, 46, pp. 69-86.
- * 17.- Nerlove, M. and Balestra, P., *Pooling cross-section and time-series data in the estimation of a dynamic model: the demand for natural gas*. Econometrica, julio 1966, pp. 585-612.
- * 18.- Nerlove, M., *Further evidence on the estimation of dynamic economic relations from a time series of sections*. Econometrica, marzo 1971, pp. 359-382.
- * 19.- Nerlove, M., *A note on error components models*. Econometrica, vol. 39, No. 2, marzo 1971, pp. 383-396.
- 20.- Nerlove, M., *Experimental evidence on the estimations from a time series of cross sections*. Economic Studies Quartely, dic. 1967, pp. 42-74.
- * 21.- Nickell, Stephen, *Biases in dynamic models with fixed effects*. Econometrica, vol. 49, p. 6, nov. 1981.
- 22.- Sahai, Hardeo, *A Bibliography of variance component*. International Statistic Revier 47, 1979, pp.177-222.
- 23.- Swamy, P., *Statistical inference in random coeficient regression model*. Springer-Verlag, Berlin. 1971.
- 24.- Swamy. P. and Arora, S., *The exact finite sample propoities of the estimators of coeficients in the error components regression models*. Econometrica, vol. 40, p. 2, marzo 1972.
- * 25.- Taylor, William, *Small sample considerations in estimation from panel data*. Journal of Econometrics. 1980, pp. 203-223.
- 26.- Triyedi, P., *Properties of small samples and collateral information*. Journal of Econometrics, 1980, vol. 12, pp. 301-318.
- 27.- Wallace, T. D. and Hussain A., *The use of error component models in combining cross-section and time series data*. Econometrica, January 1969, pp. 55-72.
- 28.- Zarembka, P., *Frontiers in Econometrics*. Academic Press, N. York and London, 1974. Ch. 5.