Econometría I. 3^o LADE
Tema 4. Estimación con regresores estocásticos

1 Introducción

En los capítulos anteriores hemos abierto la puerta al uso de regresores que no son de tipo determinístico, sino que tienen una naturaleza estocástica. Esto nos permite estudiar las propiedades temporales de las variables, interpretando el comportamiento de las mismas para, por ejemplo, sacar conclusiones con respecto a la influencia de los shocks. Esto nos permite conocer si los shocks tienen una influencia permamente o, por el contrario, si esta desaparece en el tiempo. El conocimiento del proceso estocástico que genera los valores de una variable nos puede ayudar también si estamos interesados en conocer sus valores futuros. Por ejemplo, es sencillo predecir la evolución futura de la variable y si conocemos que esta variable, por ejemplo, sigue un proceso AR(1) como el siguiente:

$$y_t = \phi_1 y_{t-1} + u_t; \qquad |\phi_1| < 1$$

En la medida que conozcamos el valor del parámetro ϕ_1 la obtención de los valores futuros de la variables es una tarea simple. Desafortunadamente, nunca vamos a conocer los parámetros poblacionales, por lo que se hace imprescindible acudir a ténicas de estimación que sean capaces de proporcionarnos una estimación fiable de dicho parámetro.

Si, por otro lado, estamos interesados en conocer la relacióne existentes entres dos o más variables, podemos especificar un modelo como el siguiente:

$$y_t = \beta_1 + \beta_1 x_{2t} + u_t$$

si disponemos de datos medidos como series temporales, hemos visto que las variables presentan correlación y, en general, nos podemos aproximar a su modelización a partir de un modelo ARIMA. De nuevo, al desconocer los valores de los parámetros de posición deberemos estimarlos. Sin embargo, como ocurría en el ejemplo anterior, desconocemos las propiedades de la estimación mco, ya que ahora la matriz de regresore no es fija sino estocástica.

El objetivo de este tema es, en primer lugar, el de estudiar las propiedades de la estimación mco en el caso en el que los regresores del modelo son de tipo estocástico. Esto supone que se incumple una de las hipótesis formuladas. El efecto que tiene la consideración de regresores estocásticos sobre la estimación mco del modelo depende del tipo de regresores estocásticos ante el que nos situemos. Como veremos, existen situaciones en las que las propiedades de la estimación mco son idénticas a las estudiadas en el caso de que los regresores son fijos. Sin embargo, en otras situaciones la estimación mco no es capaz de proporcionarnos estimadores con propiedades óptimas. En este segundo caso, debemos acudir a métodos de estimación alternativos. Algunos de estos métodos serán discutidos también en este tema.

2 Regresores estocásticos independientes de la perturbación

Supongamos que queremos estudiar el siguiente modelo:

$$y = X\,\beta + u$$

donde y es un vector columna de orden T que contiene las observaciones de la variable endógena, X es una matriz de orden (TxK) que contiene las T observaciones de las k variables explicativas del modelo, u es un vector de orden (Tx1) de perturbaciones y β es un vector-columna de orden k de parámetros. Asimismo, suponemos que X es una matriz estocástica, lo que implica que al menos una de las variables explicativas no es determinística sino de naturaleza estocástica. Además, consideramos que la distribución de la matriz X es independiente de la distribución del vector de perturbaciones, por lo que tenemos que $E\,(X'u) = E(X')\,E(u)$. En ese caso, resulta cierto que:

$$E(u/Z) = E(u) = 0$$

$$E(uu'/Z) = E(uu') = \sigma^2\,I$$

Por tanto, este supuesto no varía en gran medida las propiedades de los estimadores mínimo cuadrático ordinarios ya que se mantiene la propiedad de que el vector de perturbaciones sigue teniendo media 0 y matriz de varianzas y covarianzas escalar. Entonces, podemos demostrar en primer lugar la insesgtadez del vector de estimadores mínimo cuadrático ordinarios:

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

Tomando esperanzas matemáticas a ambos lados de la igualdad tenemos que

$$E(\hat{\beta}) = \beta + E[(X'X)^{-1}X'u]$$

Dado que la distribuciones de X y u son independientes, se demuestra que:

$$E(\hat{\beta}) = \beta + E[(X'X)^{-1}X']\, E(u) = \beta + E[(X'X)^{-1}X'] \cdot 0 = \beta$$

ya que es cierto que E(u) = 0. Por tanto, el vector de estimadores sigue siendo insesgado. Por otro lado, la matriz de varianzas y covarianzas del vector de estimadores es igual a:

$$Var(\hat{\beta}) = E\left\{[\hat{\beta} - E(\hat{\beta})]\, [\hat{\beta} - E(\hat{\beta})]'\right\} = E[(X'X)^{-1}X'u\; u'X(X'X)^{-1}]$$

Cuando la matriz de regresores es de tipo determinístico, el único componente estocástico de la esperanza anterior es el vector u y la esperanza se tranforma en $(X'X)^{-1}X'E\,(u\; u')\,X(X'X)^{-1} = \sigma^2(X'X)^{-1}$. Sin embargo, esta igualdad no es cierta en el caso de que la matriz X sea estocástica estocástica, por lo que la obtención de la matriz de varianzas y covarianzas del vector de estimadores mco no es tan sencilla. No obstante, podemos llegar a una expresión similar. Para ello debemos tener el siguiente resultado

Lemma 1 *Si $f(x,y)$ es la distribución conjunta de dos variables aleatorias x e y, y si $h(x,y)$ es cualquier función tal que exista $E\,[h(x,y)]$, entonces $E\,[h(x,y)] = E_x\left\{E_{y/x}\,[h(x,y)]\right\}$*

La prueba de este lema es sencilla. Basta con tener en cuenta que $f\,(x,y) = \kappa(y/x)\,g(x)$. Entonces,

$$\begin{aligned}
E\,[h\,(x,y)] &= \iint_{-\infty}^{\infty} h(x,y)\, f(x,y)\, dx\, dy = \iint_{-\infty}^{\infty} h(x,y)\;\; \kappa(y/x)\, g(x) dx\, dy \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} h(x,y)\;\; \kappa(y/x)\;\; dy\right] g(x)dx = \int_{-\infty}^{\infty} E_{y/x}\,[h(x,y)]\;\; g(x)\, dx \\
&= E_x\left\{E_{y/x}\,[h(x,y)]\right\}
\end{aligned}$$

Debemos tener en cuenta que en la expresión anterior $E_{y/x}$ representa el valor esperado en la distribución condicional de y, dada X, mientras que E_x es el valor esperado en la distribución marginal de x. Si aplicamos el resultado del Lema anterior a nuestro caso, podemos expresar la igualdad anterior:

$$E[(X'X)^{-1}X'u\; u'X(X'X)^{-1}] = E_u\{E[(X'X)^{-1}X'u\; u'X(X'X)^{-1}/X]\}$$

Ahora, los componentes de la segunda esperanza están condicionados a X, por lo que esta matriz se considera conocida, por tanto, determinística. Entonces, resulta que:

$$\begin{aligned} E_u\{E[(X'X)^{-1}X'u\,u'X(X'X)^{-1}/X]\} &= E_u\{(X'X)^{-1}X'E(u\,u')X(X'X)^{-1}\} = \\ E_u\{(X'X)^{-1}X'\sigma^2 I\,X(X'X)^{-1}\} &= \sigma^2\,E_u\{(X'X)^{-1}X'\,X(X'X)^{-1}\} = \\ &\quad \sigma^2\,E\left[(X'X)^{-1}\right] \end{aligned}$$

Por tanto, la matriz de varianzas y covarianzas del vector de estimadores adopta una firma similar a la conocida en el caso de regresores no estocásticos. La única diferencia es que ahora aparece la esperanza de la matriz $E\,[(X'X)^{-1}]$ en lugar de la propia matriz $(X'X)^{-1}$.

Por último, la estimación del parámetro de dispersión tiene las siguientes propiedades. En primer lugar, sigue siendo insesgado:

$$\begin{aligned} E(\hat{\sigma}^2) &= E(\frac{\hat{u}'\hat{u}}{T-k}) = E\left[\frac{\left(y - X\,\hat{\beta}\right)'\left(y - X\,\hat{\beta}\right)}{T-k}\right] = \\ &= E\,E\left[(T-k)^{-1}\,E\left(\hat{u}'\hat{u}/X\right)\right] = E\,E\left[(T-k)^{-1}\,\sigma^2\,(T-k)\right] \\ &= \sigma^2 \end{aligned}$$

sin más que aplicar los mismos argumentos que los usados para obtener la matriz de varianzas y covarianzas del vector de estimadores mco y teniendo en cuenta adicionalmente que $tr\,(M) = (T-k)$. Por tanto, hemos demostrado la insesgadez del estimador $\hat{\sigma}^2$, lo que también implica la insesgadez del estimador $\widehat{Var}\,(\hat{\beta})$.

Como hemos visto, este primer caso de regresores estocásticos no supone grandes cambios con respecto al caso de regresores determinísticos. Si imponemos la restricción de que todos los resultados están condicionados a un valor dado de la matriz X, todos los resultados conocidos se mantienen. Así, aunque no lo demostramos explícitamente, el vector de estimadores sigue siendo ELIO (BLUE) o, imponiendo normalidad en las perturbaciones, eficiente.Finalmente, también se podría demostrar que los métodos de contraste de hipótesis siguen siendo válidos.

Desafortunadamente, éste no es el caso de regresores estocásticos más habitual cuando trabajamos con datos de series temporales. En estas circunstancias, es habitual que la especificación del modelo incluya retardos de la variable endógena. En este tipo de circunstancias, no podemos admitir la independencia entre el vector de perturbaciones y la matriz de regrerores. Aunque, esta dependencia tiene una forma particular tal que algunas de las propiedades que observabamos en el caso de regresores fijos se mantiene. En la siguiente sección comprobamos los efectos que esta dependencia produce sobre las propiedades de los estimadores del modelo.

3 Regresores parcialmente dependientes de la perturbación

Éste es un caso muy frecuente en el análisis econométrico cuando las observaciones disnponibles de las variables son medidas como una serie temporal. Imaginemos que queremos estudiar el comportamiento de una variable como el consumo agregado de una economía. Como variables explicativas podemos utilizar el nivel' de Renta de la Economía y, además, el consumo retardado del período anterior. La motivación para la inclusión de esta segunda variable viene dada porque las decisiones sobre el consumo en los periodos anteriores pueden determinar el consumo del presente periodo. Un ejemplo sería aquel en el que el consumidor decide ahorrrar un parte de la renta del periodo $t-1$ para destinarla al consumo en el periodo t. En estas condiciones, la especificación del modelo queda de la siguiente manera:

$$C_t = \beta_1 + \beta_2 R_t + \beta_3 C_{t-1} + u_t$$

donde C_t es el nivel del consumo agregtado del período t e R_t es la renta disponible en el período t y u_t es una perturbación que cumple las hipótesis básicas.

Como vemos, si estamos caracterizando la variable consumo como la variable endógena, debe tener propiedades estocásticas. En consecuencia, C_{t-1} también debe ser considerada como una variable aleatoria y, por ello, estamos en el caso de regresores estocásticos. Sin embargo, no podemos admitir que ahora la matriz de regresores se distribuya de forma independiente con respecto al vector de perturbaciones, ya que claramente el consumo del periodo t depende de la perturbación del periodo t y de todas las perturbaciones de periodos anteriores. Sin embargo, no depende de las perturbaciones futuras. En realidad, este resultado ya era conocido por cuanto el modelo anterior no es sino un proceso AR(1) al que le hemos agregado la variable Renta. Por tanto, se incumple uno de los supuestos sobre los que se construye el MLG. Esto conlleva que la distribución conjunta de X y u no tenga siempre una media igual a cero, ya que ahora $E(X'u)$ no es necesariamente igual a 0. Esto implica que el vector de estimadores mínimo cuadrático ordinario no es insesgado en muestras finitas ya que:

$$E(\hat{\beta}) = \beta + E[(X'X)^{-1}X'u] \neq \beta$$

Como consecuencia de este resultado, el vector p dejará de ser ELIO y eficiente en muestras finitas, por lo que la estimación mínimo cuadrática ordinaria deja de tener propiedades óptimas.

A pesar de ello, la facilidad en el uso de este tipo de estimadores ha sido un gran estímulo a la hora de continuar usándolos. Además, a pesar de perder sus buenas propiedades en un escenario de muestras finitas, asintóticamente mantienen intactas estas propiedades. Para comprobar este extremo debemos comenzar por considerar la siguiente especificación:

$$Y = Z\,\beta + u$$

donde Z es la matriz de regresores que incluye tanto los diversos retardos de la variable endógena $z_t' =(\mathrm{y}_{t-1}, y_{t-2,}..., y_{t-p,} x_t')$ Del mismo modo, mantenemos el supuesto de que el vector de perturbaciones tiene media 0 y matriz de varianzas y covarianzas σ^2I.. Por último, continúa siendo cierto que

$$plim\, Z'Z/T = \Sigma_{zz}$$

donde Σ_{xx} es una matriz no singular y, dadas las características de la matriz Z y del vector u, es cierto que:

$$plim\, Z'u/T = 0.$$

Bajo estas premisas, puede demostrarse la consistencia del vector de estimadores mínimo cuadrático ordinarios:

$$\begin{aligned} p\lim \hat{\beta} &= p\lim \left[\beta + (Z'Z)^{-1} Z'u\right] = \beta + p\lim \left[(Z'Z)^{-1} Z'u\right] \\ &= \beta + p\lim \left[(Z'Z)^{-1}/T\right] p\lim\,[Z'u/T] \;=\; \beta \;+\; \Sigma_{zz}^{-1}\,0 \\ &= \beta \end{aligned}$$

De forma similar, podemos demostrar la consistencia del estimador mco de la varianza de la perturbación. Para ello, basta tener en cuenta que:

$$p\lim \tilde{\sigma}^2 = p\lim\left(\frac{\hat{u}'\hat{u}}{T-k}\right) = p\lim\left(\frac{T-k}{T-k}\;\sigma^2\right) = \sigma^2$$

Tan sólo nos resta demostrar que asintóticamente también es posible el uso de los contrastes de hipótesis habitualmente utilizados. Para ello, debemos tener en cuenta los resultados del teorema de Mann-Wald

3.1 Mann-Wald Theorem.

Before when deriving the limiting distribution of the least-squares estimates, we required a result of the form

$$T^{-1/2}\Sigma_{t=1}^{T} X_t u_t \rightarrow^d N\left(0, \sigma^2 A\right)$$

where

$$p\lim T^{-1}\Sigma_{t=1}^{T} X_t X_t' = A.$$

The result used was valid for the case where the X's are independent of the errors u_t. We now consider a generalization that still requires the X's and the u_t to be uncorrelated contemporaneously but allows correlation at some leads and lags thereby allowing lagged dependent variables as regressors.

Lemma 2 *(Mann-Wald Theorem): Consider a k dimensional vector of regressors Z_t and a sequence of errors $u_t \sim i.i.d.(0, \sigma^2)$ such that*

$$p\lim T^{-1}\Sigma_{t=1}^{T} Z_t Z_t' = Q$$

with Q a positive definite matrix; and

$$E\left[Z_t u_t\right] = 0 \quad \textit{for all } t.$$

Then,

$$p\lim T^{-1}\Sigma_{t=1}^{T} Z_t u_t = 0,$$

and

$$T^{-1/2}\Sigma_{t=1}^{T} Z_t u_t \rightarrow^d N\left(0, \sigma^2 Q\right).$$

We can use the above result to analyze the limiting distribution of the least-squares estimates in the following regression model:

$$y_t = \gamma Z_t + u_t$$

with $\varepsilon_t \sim i.i.d.\,(0, \sigma^2)$. We have,

$$\begin{aligned}
\sqrt{T}\left(\hat{\gamma} - \gamma\right) &= \left(T^{-1}\Sigma_{t=1}^{T} Z_t Z_t'\right)^{-1}\left(T^{-1/2}\Sigma_{t=1}^{T} Z_t u_t\right) \\
&\rightarrow^d Q^{-1} \cdot N\left(0, \sigma^2 Q\right) \\
&= N\left(0, \sigma^2 Q^{-1}\right).
\end{aligned}$$

En resumen, a pesar de la existencia de un cierto grado de correlación entre la matriz de regresores y el vector de perturbaciones aleatorias, la estimación OLS mantiene todas sus propiedades, al tiempo que la inferencia estadística habitualmente utilizada sigue siendo válida asintóticamente.

4 Regresores totalmente dependientes de la perturbación

Existen situaciones en las que la presencia de regresores estocásticos en la especificación del modelo no se incluye en los dos casos anteriormente analizado. En ese tipo de situaciones, la matriz de regresores depende contemporáneamente del vector de perturbaciones. Por ejemplo, supongamos que tenemos el siguiente modelo:

$$y = X\beta + u$$

pero ahora la relación entre la matriz de regresores y el vector de perturbaciones es la siguiente:

$$\operatorname*{p\lim}_{T\to\infty} T^{-1}X'u \neq 0.$$

En estas condiciones, la estimacioón mco es inconsistente. Para demostrarlo basta con considerar que:

$$\operatorname*{p\lim}_{T\to\infty}\left(\hat{\beta}-\beta\right) = \underset{\downarrow \atop A^{-1}}{\operatorname{p\lim}\left(\frac{X'X}{T}\right)^{-1}} \underset{\downarrow \atop c}{\operatorname{p\lim}\left(\frac{X'u}{T}\right)} \neq 0.$$

Existen diversos casos en los que se puede presentar este tipo de comportamiento:

a) Modelos dinámicos con perturbación correlacionada

$$y_t = \alpha y_{t-1} + x_t'\beta + u_t$$

donde

$$u_t = \rho u_{t-1} + e_t \quad e_t \sim i.i.d.(0, \sigma^2).$$

Es evidente que ahora, $\operatorname{plim} T^{-1}y_{t-1}u_t \neq 0$ por cuanto y_t depende ahora de todas las perturbaciones, las anteriores, la actual y las futuras.

b) Simultaneous equations model.

If it has an endogenous variable as a regressor, then it will not generally be independent of the disturbance.

c) Models with errors-in-variables.

Supongamos que la variable y viene generada por el siguiente modelo:

$$y_i = \alpha + \beta X_i^* + e_i$$

donde X_i^* no es observable y viene generada por

$$X_i = X_i^* + w_i$$

donde X_i sí que la podemos observar, siendo $w_i \sim i.i.d.\,(0, \sigma_w^2)$ e incorrelacionada con X_i^*. En estas circunstancias, tenemos que:

$$\begin{aligned} y_i &= \alpha + \beta\,(x_i - w_i) + e_i \\ &= \alpha + \beta X_i - \underbrace{(\beta w_i - e_i)}_{u_i} \end{aligned}$$

Por lo que existe correlación entre X_i y u_i.

Como hemos visto, en este tipo de modelos la estimación mco es sesgada, por lo que tenemos que acudir a otro tipo de estimación.

4.1 Estimación por variables instrumentales

Supongamos que tenemos que estimar un modelo donde el vector de perturbaciones está correlacionado contemporáneamente con la matriz de regresores. En estas circunstancias, al estimación mco no presentan propieades óptimas ni en muestras finitas, ni asintóticamente.Una solución es el uso de una serie de instrumentos que estén incorrelacionados con la perturbación del modelo. Esto se llama estimación por variables instrumentales. Si designamos w_t al conjunto de k "variables instrumentales " en el periodo t, el conjunto de instrumentos debe cumplir las siguiente propiedades:

a) $E\,[w_t u_t] = 0$, los instrumentos deben estar incorrelacionados contemporáneamente con la perturbación del modelo.

b) Los instrumentos deben estar altamente correlacionados con las variables explicativas.

Entonces, supongamos que se cumple que:

$$\left.\begin{array}{l} \operatorname*{p\lim}_{T\to\infty} T^{-1}W'X = Q_{WX} \\ \operatorname*{p\lim}_{T\to\infty} T^{-1}W'W = Q_{WW} \end{array}\right\} \text{matrices definidas positivas.}$$

El vector de estimadores por variables instrumentales se define de la siguiente manera:

$$\hat{\beta}_{IV} = (W'X)^{-1}\,W'y.$$

Las propiedades del vector de estimadores $\hat{\beta}_{IV}$ son las siguientes:

a) **Consistencia**

$$\mathrm{p}\lim\left(\hat{\beta}_{IV}-\beta\right)=\mathrm{p}\lim\underset{\underset{Q_{WX}}{\downarrow}}{\left(\frac{W'X}{T}\right)^{-1}}\mathrm{p}\lim\underset{\underset{0}{\downarrow}}{\left(\frac{W'u}{T}\right)}=0.$$

b) **Normalidad asintótica**

A partir del Teorema de Mann-Wald tenemos que

$$T^{-1/2}W'u\rightarrow^{d}N\left(0,\sigma^{2}Q_{WW}\right).$$

Entonces,

$$\begin{aligned}\sqrt{T}\left(\hat{\beta}_{IV}-\beta\right) &= \left(\frac{W'X}{T}\right)^{-1}\left(\frac{W'u}{T^{1/2}}\right)\\ &\rightarrow^{d} Q_{WX}^{-1}N\left(0,\sigma^{2}Q_{WW}\right)\\ &= N\left(0,\sigma^{2}Q_{WX}^{-1}Q_{WW}Q_{WX}^{-1}\right).\end{aligned}$$

Por tanto, la matriz de varianzas y covarianzas asintótica del vector de estimadores instrumentales se puede aproximarse por la siguiente expresión:

$$\hat{\sigma}^{2}\left(W'X\right)^{-1}W'W\left(W'X\right)^{-1}.$$

donde es sencillo comprobar que si W y X están escasamente correlacionaas, entonces Q_{WX} is "pequeño" y, como consecuencia, la varianza del estimador instrumental es "elevada".

Por último, no podemos decir nada sobre la existencia de una vector de estimadores instrumentales que tenga mínima varianza, por cuanto podemos definir infinitos instrumentos que cumplan con los requisitos anteriores. Este es el principal problema de este uso, por cuanto es difícil conocer cuál es el conjunto de instrumentos más adecuado.

4.2 Caso especial

Supongamos que la variable y_t viene generado por el siguiente modelo:

$$y_t=\beta x_t+u_t$$

donde x_t es una variable no aleatoria y, no obstante, utilizamos la estimación instrumental a partir del instrumento w_t. En estas circunstancias, $\hat{\beta}_{OLS}$ y

$\hat{\beta}_{IV}$ son insesgadas y sus matrices de varianzas y covarianzas son:

$$\begin{aligned} Var\left(\hat{\beta}_{OLS}\right) &= \sigma^2 \left(X'X\right)^{-1}, \\ Var\left(\hat{\beta}_{IV}\right) &= \sigma^2 \left(W'W\right)\left(W'X\right)^{-2}. \end{aligned}$$

Entonces, el coeficiente de correlación simple entre la variable explicativa y el instrumento se puede definir de la siguiente manera:

$$r_{XW}^2 = \frac{\left(W'X\right)^2}{\left(W'W\right)\left(X'X\right)}.$$

Como consecuencia, la eficiencia relativa de la estimación por variables instrumentales viene dada por el siguiente cociente:

$$\frac{Var\left(\hat{\beta}_{OLS}\right)}{Var\left(\hat{\beta}_{IV}\right)} = r_{WX}^2.$$

Por tanto, de la anterior expresión se deduce que cuanto mayor es el grado de correlación entre los regresores y los instrumentos, mayor es la eficiencia de la estimación por variables instrumentales. Este resultado, que sólo ha sido demostrado para un caso particular, también es válido en el caso general.

4.3 El estimador GIVE (generalized instrumental variable estimator)

COmo hemos comentado, el principal problema del uso de estos estimadores es la selección del conjunto de instrumentos más adecuado. En este sentido, debemos tener en cuenta que no existe ninguna razón que nos lleve a restringir la aplicación de esta técnica a un único conjunto de instrumentos. Al contrario, si podemos definir distintos conjuntos de intrumentos, parece adecuado pensar en que una combinación de ellos nos puede ayudar a incrementar la eficiencia de la estimación. Una solución es definir une estimador que recoja la información procedente de varios instrumentos a partir del uso de una media ponderada. Los pesos deberían ser elegidos de forma que la correlación con respecto a la matriz X se maximice.

Entonces, supongamos que $\underset{\ell \times 1}{W}$ es una vector de instrumentos, donde $\ell > k$. El instrumento óptimo, que vendría definido por W^*, será:

$$W^* = P_W X = W\left(W'W\right)^{-1} W'X.$$

A partir de esta información, el nuevo vector de estimadores instrumentales queda definido de la siguiente manera:

$$\begin{aligned}\hat{\beta}_{GIVE} &= \left(W^{*\prime}X\right)^{-1}W^{*\prime}y \\ &= \left(X'P_WX\right)^{-1}X'P_Wy\end{aligned}$$

y su distribución asintótica queda determinada por esta expresión:

$$\sqrt{T}\left(\hat{\beta}_{GIVE}-\beta\right)\to^{d} N\left(0,\sigma^{2}\mathrm{p}\lim\left(\frac{X'P_WX}{T}\right)^{-1}\right).$$