Capítulo 20. REGRESORES ESTOCASTICOS

20.1. INTRODUCCIÓN 860

CONSECUENCIAS SOBRE LA ESTIMACIÓN MCO 862

DISTRIBUCIÓN CONDICIONADA 866

CASO A. VARIABLE EXPLICATIVA ENDÓGENA CON RETARDO 868

CASO B. VARIABLE EXPLICATIVA NO OBSERVABLE 869

CASO C. VARIABLES EXPLICATIVAS ENDÓGENAS 870

20.2 MODELO DE REGRESIÓN LINEAL CON REGRESORES ESTOCÁSTICOS 871

REGRESORES INDEPENDIENTES DE LA PERTURBACIÓN 872

INCORRELACIÓN CONTEMPORÁNEA 874

CORRELACIÓN CONTEMPORÁNEA 875

20.3. MÉTODO DE VARIABLES INSTRUMENTALES 876

20.4 CONTRASTE DE HAUSMAN 882

CASOS DE ESTUDIO, PREGUNTAS Y PROBLEMAS 884

CASO 20.1. MODELO KEYNESIANO SIMPLE 884

CASO 20.2 MODELO SIMPLE DE MERCADO 893

PREGUNTAS 894

PROBLEMA 20.1. MODELO CON ERRORES DE MEDIDA 896

PROBLEMA 20.2. MODELO CON ENDÓGENA RETARDADA 898

PROBLEMA 20.3 CÁLCULO DEL ESTIMADOR DE VARIABLE INSTRUMENTAL (VI) 899

BIBLIOGRAFÍA 901

Capítulo 20. REGRESORES ESTOCÁSTICOS

En el modelo de regresión lineal se ha expuesto, hasta ahora, que las variables explicativas eran fijas -o no estocásticas- en muestras repetidas. Este supuesto puede ser apropiado para experimentos de laboratorio, en los que el investigador tiene el control sobre las variables explicativas, puede fijar el valor de las mismas y observar los resultados obtenidos para la variable endógena en experimentos repetidos; o para el caso de variables que se construyen artificialmente, como pueden ser las tendencias lineales o las variables ficticias. Pero en el proceso de investigación econométrica descrito hasta ahora, en general, las variables explicativas no están sujetas a control y tanto las variables endógenas como los regresores son el resultado de un determinado sistema económico-social; por lo tanto, ambos tipos de variables, son estocásticos por naturaleza. Si se está analizando la relación entre consumo y renta, donde el parámetro de interés es la propensión marginal a consumir, no se puede suponer que la variable explicativa renta sea fija; ya que, tanto el consumo como la renta, vienen determinados por el mismo sistema económico-social y son aleatorias. Bajo esta nueva situación, se verá si los métodos de inferencia desarrollados aún se pueden aplicar y, en caso contrario, de qué métodos de estimación alternativos se dispone.

20.1. Introducción

En el proceso de investigación econométrica se planteó el modelo:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad\qquad \boldsymbol{\varepsilon} \sim \mathbf{N}(\mathbf{0}, \sigma^2\mathbf{I})$$

donde se supone que los regresores incluidos en $\mathbf{X}$ son variables deterministas.

Se dice que una variable es determinista si sus valores no cambian en distintas muestras. Cuando no puedan repetirse las observaciones -como ocurre a menudo en Economía- no es trivial saber si una variable es o no aleatoria. Por ello es importante analizar si el posible incumplimiento de esta hipótesis afecta las propiedades de la estimación MCO.

Por otra parte, muchos modelos econométricos tienen algunos regresores aleatorios. Por ejemplo:

- modelos de ecuaciones simultáneas, o
- modelos dinámicos

En el proceso de investigación econométrica sólo se reconoce que existe este problema cuando la estructura del modelo *implica* la existencia de regresores estocásticos.

Los sistemas de ecuaciones simultáneas describen el comportamiento de un vector de variables endógenas en función de un vector de variables exógenas.

Los regresores estocásticos surgen del hecho de que la variable endógena de una ecuación puede entrar en otra como variable explicativa.

Ejemplo 20.1 El modelo de consumo de Haavelmo se define como:

$$\mathrm{C_t} = \beta \mathrm{PBI_t} + \varepsilon_t$$
$$\mathrm{PBI_t} = \mathrm{C_t} + \mathrm{I_t}$$

o bien

$$\begin{bmatrix} 1 & -\beta \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \mathrm{C_t} \\ \mathrm{PBI_t} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mathrm{I_t} + \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix}$$

e implica que $\mathrm{PBI_t}$ es un regresor estocástico en la primera ecuación ya que, por la segunda ecuación $\mathrm{PBI_t}$ depende de $\mathrm{C_t}$ y, por la primera ecuación $\mathrm{C_t}$ depende de ε_t

La variable endógena de un modelo de regresión es siempre estocástica. Por ello, en un modelo de regresión dinámico que incluya la variable endógena retardada como variable explicativa, siempre hay regresores estocásticos.

Ejemplo 20.2 El modelo AR(1) es un modelo de regresión con variables explicativas estocásticas:

$$y_t = \rho y_{t-1} + \varepsilon_t$$

Hay otras situaciones en las que es necesario reconocer que el modelo lineal general tiene regresores estocásticos; cabe destacar los de variables medidas con error o modelos que incluyen expectativas como variables explicativas.

Consecuencias sobre la estimación MCO

Sea el método de regresión lineal general en el que se cumplen los supuestos habituales, pero donde ahora la matriz de regresores $\mathbf{X}$ es estocástica. Los coeficientes de regresión se pueden estimar aplicando el criterio MCO:

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}$$

Este estimador ya no es una combinación lineal de las perturbaciones, sino que es una función estocástica no lineal de $\mathbf{X}$ y $\boldsymbol{\varepsilon}$, por lo tanto, sus propiedades dependerán de *la distribución conjunta* de estas. Por ejemplo, si se quiere comprobar si el estimador es insesgado:

$$E(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\beta} + E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}]$$

Para obtener $E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}]$ se debe conocer la distribución conjunta de las variables aleatorias $\mathbf{X}$ y $\boldsymbol{\varepsilon}$.

Bajo el supuesto de regresores fijos, el problema se soluciona fácilmente:

$$E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}] = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\boldsymbol{\varepsilon})$$

y este valor medio es cero, dado que $E(\boldsymbol{\varepsilon}) = \mathbf{0}$.

Cuando los regresores son estocásticos, esta igualdad ya no se cumple y es preciso contar con la distribución conjunta de $\mathbf{X}$ y $\boldsymbol{\varepsilon}$ para derivar propiedades de los estimadores $\widehat{\boldsymbol{\beta}}$, así como las distribuciones de los estadísticos de contraste habituales.

La consecuencia fundamental de la presencia de regresores estocásticos en el modelo lineal general, es que resulta imposible obtener los momentos del estimador mínimo cuadrado ordinario.

Por tanto, cuando hay regresores estocásticos se desconocen:

a) los primeros momentos del estimador mínimo cuadrático ordinario

b) sus propiedades en muestras finitas.

Concretamente, no pueden justificarse la insesgadez del estimador MCO, su eficiencia y la distribución de los estadísticos de contraste bajo la hipótesis nula. Ello obliga a estudiar *propiedades asintóticas*, esto es, las que tendría el estimador en una muestra con infinitas observaciones.

La propiedad asintótica fundamental es la de *consistencia*. Bajo el supuesto de que todas las *variables explicativas están centradas*. Trivialmente:

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} \Rightarrow \text{plim}\,\widehat{\boldsymbol{\beta}} = \text{plim}\,\boldsymbol{\beta} + \text{plim}\left[\left(\frac{1}{T}\mathbf{X}'\mathbf{X}\right)^{-1}\left(\frac{1}{T}\mathbf{X}'\boldsymbol{\varepsilon}\right)\right]$$

aplicando sucesivamente los teoremas de Slutsky y Kinchine:

$$\text{plim}\,\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta} + \text{plim}\left(\frac{1}{T}\mathbf{X}'\mathbf{X}\right)^{-1}\text{plim}\left(\frac{1}{T}\mathbf{X}'\boldsymbol{\varepsilon}\right) = \boldsymbol{\beta} + \Sigma_{\mathbf{XX}}^{-1}\Sigma_{\mathbf{X}\boldsymbol{\varepsilon}}$$

Donde se parte que los momentos de las variables X convergen en probabilidad a una matriz no singular Σ_{XX}. Por lo que el estimador MCO es consistente si y sólo si los regresores *no están correlacionados* con el término de error.

Observación: Los siguientes son resultados fundamentales de teoría asintótica.

Convergencia en probabilidad: Sea $\{b_n\}$ una sucesión de variables aleatorias definida en el conjunto de números reales. Se dice que $\{b_n\}$ converge en probabilidad a b si y sólo si

$$\exists\delta > 0/ \lim_{n\to\infty} P(\|b_n - b\|) < \delta = 1$$

Otra forma de expresar matemáticamente este concepto es,

$$\text{plim} b_n = b$$

expresión que se lee *el límite en probabilidad de* b_n *es* b.

Estimador consistente: Se dice que $\hat{\beta}_n$ es un estimador consistente de β si

$$\text{plim}\hat{\beta}_n = \beta$$

Una condición suficiente (no necesaria) para que $\hat{\beta}_n$ sea un estimador consistente de β, es que

$$\lim_{n\to\infty} E[\hat{\beta}_n] = \beta \text{ y } \lim_{n\to\infty} \text{Var}[\hat{\beta}_n] = 0$$

Por otra parte, si

$$\text{plim}\,\hat{\alpha}_n = \alpha \text{ y plim}\,\hat{\beta}_n = \beta$$

Entonces,

$\text{plim}\,\gamma = \gamma$ siendo γ una constante (no depende de n)

$\text{plim}\left[\hat{\alpha}_n \pm \hat{\beta}_n\right] = \alpha \pm \beta$

$\text{plim}\left[\hat{\alpha}_n . \hat{\beta}_n\right] = \alpha . \beta$

$\text{plim}\left[\frac{\hat{\alpha}_n}{\hat{\beta}_n}\right] = \frac{\alpha}{\beta}, \text{si } \hat{\beta}_n \neq 0 \text{ y } \beta \neq 0$

$\text{plim}\left[\sqrt{\hat{\alpha}_n}\right] = \sqrt{\alpha} \text{ si } \hat{\alpha}_n > 0 \text{ y } \alpha > 0$

$\text{plim}\,\varphi[\hat{\alpha}_n] = \varphi[\alpha]$ siendo $\varphi[\ \ ]$ una función continua (teorema de Slutsky)

Si m es un momento muestral de orden r y μ es el correspondiente momento poblacional, entonces $\text{plim} m_r = \mu_r$ (teorema de Kinchine)

Estas propiedades se cumplen tanto para escalares como para vectores.

Es importante tener en cuenta que las propiedades de consistencia e insesgadez son distintas e independientes. Concretamente, es posible que un estimador sea:

- sesgado y consistente (por ejemplo, la varianza muestral),
- insesgado e inconsistente (por ejemplo, $y_i = \beta X_i + \varepsilon_i \quad \beta = \frac{y_i}{X_i}$)

• sesgado e inconsistente, (por ejemplo, el estimador anterior más un sesgo deliberado), o

• insesgado y consistente (por ejemplo, el estimador MCO bajo las hipótesis habituales).

Distribución condicionada

Una forma de enfocar este problema es utilizar la distribución de $\mathbf{y}$ condicionada a las $\mathbf{X}$. La función de distribución conjunta $f(\mathbf{y},\mathbf{X};\boldsymbol{\beta},\sigma^2)$ se puede escribir como:

$$f(\mathbf{y},\mathbf{X};\boldsymbol{\beta},\sigma^2) = f(\mathbf{y}/\mathbf{X};\ \boldsymbol{\beta},\sigma^2)\ \ f(\mathbf{X},\boldsymbol{\gamma})$$

Si el interés se centra en los parámetros de la *distribución condicionada* $(\boldsymbol{\beta},\sigma^2)$ y estos no están relacionados con los parámetros de la *distribución marginal*, $\boldsymbol{\gamma}$, se puede olvidar de ella y considerar solo la distribución de $\mathbf{y}$ condicionada a uno de los valores fijos de las variables $\mathbf{X}$.

El modelo de regresión lineal general condicionado a $\mathbf{X}$, se puede escribir como:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

donde:

$$E(\boldsymbol{\varepsilon}/\mathbf{X}) = \mathbf{0}_T$$

$$E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'/\mathbf{X}) = \boldsymbol{\sigma}^2\mathbf{I}_T$$

$$\rho(\mathbf{X}) = k \leq T$$

$$\boldsymbol{\varepsilon}/\mathbf{X} \sim N(\mathbf{0}, \boldsymbol{\sigma}^2 \mathbf{I}_T$$

Se obtienen los siguientes resultados condicionados:

$$E(\widehat{\boldsymbol{\beta}}/\mathbf{X}) = \boldsymbol{\beta} + E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}/\mathbf{X}] = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\boldsymbol{\varepsilon}/X) = \boldsymbol{\beta} \qquad (20.1)$$

De la misma forma:

$$V(\widehat{\boldsymbol{\beta}}/\mathbf{X}) = \boldsymbol{\beta} + E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\,(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'/\mathbf{X})\,\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}] = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'/\mathbf{X})\,\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\sigma}^2\mathbf{I}_T\,\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \boldsymbol{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1} \qquad (20.2)$$

Donde, σ^2se estima a partir de S^2 y se cumple que: $E(S^2/\mathbf{X}) = \sigma^2$. De modo que S^2 es estimador insesgado de σ^2

De esta forma, un estimador insesgado de la varianza condicionada de los estimadores, viene dado por:

$$\hat{V}(\widehat{\boldsymbol{\beta}}/\mathbf{X}) = S^2(\mathbf{X}'\mathbf{X})^{-1}$$

El estimador $\widehat{\boldsymbol{\beta}}$ no es un estimador lineal, sino una función estocástica no lineal en $\mathbf{X}$ e $\mathbf{y}$, por lo que estrictamente hablando no se puede aplicar el teorema de Gauss-Makov y decir que es ELIO. Sin embargo, si se considera la varianza del estimador como condicional a valores dados de $\mathbf{X}$, entonces el estimador es eficiente.

Por otro lado, la distribución de $\widehat{\boldsymbol{\beta}}$ condicionada a los regresores $\mathbf{X}$, es:

$$\widehat{\boldsymbol{\beta}}/\mathbf{X} \sim N[\boldsymbol{\beta}, \boldsymbol{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}]$$

y los estadísticos de contraste, condicionados a $\mathbf{X}$, siguen una distribución t de Student y F de Snedecor, respectivamente.

De esta forma, aunque en principio las variables $\mathbf{X}$ son variables aleatorias, si se condiciona el análisis a valores fijos de estas, los resultados dependen de los valores concretos que tomen en la muestra.

El problema se plantea en contextos en que los regresores son estocásticos y no tiene sentido realizar un análisis condicionado a valores fijos de $\mathbf{X}$. Estas situaciones se pueden presentar en los siguientes tres casos:

Caso A. variables explicativas con retardo

CASO B. variable explicativa no observable

CASO C. variables explicativas endógenas

CASO A. variable explicativa endógena con retardo

Si se tiene el siguiente modelo de regresión:

$$Y_t = \alpha + \beta Y_{t-1} + \mu_t \qquad t = 1, \dots, T \qquad (20.3)$$

En este modelo aparece como regresor la variable dependiente retardada un período. Dado que $Y_1, \dots, Y_T$ son variables aleatorias, el regresor Y_{t-1} es una variable aleatoria. En esta situación, la matriz

$$\mathbf{X} = \begin{bmatrix} 1 & y_{2-1} \\ \vdots & \vdots \\ 1 & y_{T-1} \end{bmatrix}$$

es estocástica. Por otro lado, no se puede realizar el análisis condicionado a valores fijos de Y_{t-1}, $t = 2, \dots, T$, ya que no tendría sentido porque es el propio modelo estocástico el que indica cómo se generan.

CASO B. variable explicativa no observable

Dado el siguiente modelo de regresión:

$$Y_t = \alpha + \beta X_t^* + \mu_t \qquad t = 1, \dots, T \tag{20.4}$$

Donde X^* es una variable no observable porque es difícil de cuantificar o medir. En su lugar, se observa la variable X, tal que:

$$X_t = X_t^* + \varepsilon_t \qquad t = 1, \dots, T$$

ε_t es una variable aleatoria que recoge el error de medida en t. En esta situación, X_t es una variable aleatoria aunque se considere X^* como fija. Por lo tanto, el modelo en términos de X -sustituyendo X_t^* por X_t- queda:

$$Y_t = \alpha + \beta X_t + \omega_t \qquad t = 1, \dots, T \tag{20.5}$$

Donde $\omega_t = \mu_t - \beta\varepsilon_t$ es el término de perturbación que recoge, además de μ_t, el error de medida ε_t.

El modelo (20.5) es equivalente al (20.4), pero donde el regresor X_t, $t = 1, \dots, T$, es una variable aleatoria. En este caso, tampoco se puede hacer un análisis condicionado a valores fijos de X, ya que hipótesis sobre $E(\omega/X)$, $E(\omega\omega'/X)$ no tendrían sentido, dado que ω es función de ε y X.

Observación. Sustituyendo en 20.5 a X_t por su igual $X_t^* + \varepsilon_t$

$$Y_t = \alpha + \beta(X_t^* + \varepsilon_t) + \omega_t$$

$$Y_t = \alpha + \beta X_t^* + \beta\varepsilon_t + \omega_t$$

$$Y_t = \alpha + \beta X_t^* + \mu_t$$

Donde $\omega_t = \mu_t - \varepsilon_t$.

CASO C. variables explicativas endógenas

Si se quiere estimar los parámetros de la siguiente ecuación de demanda de un bien:

$$Q_t = \alpha + \beta P_t + \mu_t \qquad t = 1, \dots, T \qquad (20.6)$$

Donde Q es la cantidad demandada y P es el precio. Dado que en el momento t se observa la cantidad y precio de equilibrio, ambas variables se determinan simultáneamente en el mercado, de modo que, tanto Q como P son variables endógenas. Por ejemplo, si en t se produce un shock en la demanda de este bien debido a un cambio en gustos de los consumidores recogido por μ_t, se generaría un cambio en el momento t, tanto de la cantidad demandada Q_t como del precio. En este contexto, la variable dependiente y el regresor se determinan simultáneamente; por lo que ambas variables Q_t y P_t son aleatorias. Este es otro caso donde la matriz de regresores

$$\mathbf{X} = \begin{bmatrix} 1 & P_1 \\ \vdots & \vdots \\ 1 & P_t \end{bmatrix}$$

es estocástica. Por otro lado, tampoco tiene sentido realizar el análisis condicionado a P_t, $t = 1, \dots, T$, dado que P_t se determina simultáneamente con Q_t.

20.2 Modelo de regresión lineal con regresores estocásticos

En el modelo de regresión lineal:

$$\mathbf{y} = \mathbf{X\beta} + \boldsymbol{\varepsilon} \qquad \boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I_T})$$

donde al menos uno de los regresores es una variable aleatoria, siendo, por lo tanto, la matriz

$$\mathbf{X} = \begin{bmatrix} 1 & x_{k1} \\ \vdots & \vdots \\ 1 & x_{kt} \end{bmatrix}$$

estocástica. Los estimadores y estadísticos, son *función de las variables aleatorias* X_{jt} y ε_t; $j = 2, \dots, k$; $t = 1, \dots, T$, por lo que será importante conocer las características estocásticas de ambos conjuntos de variables aleatorias y cómo se relacionan.

Regresores independientes de la perturbación

Cuando las variables *aleatorias* X_{jt} y ε_t son independientes, para todo $j = 2, \dots, k$ y $t = 1, \dots, T$, la función de densidad marginal de $(X_{2t}, \dots, X_{kt})$ no depende del vector paramétrico $(\boldsymbol{\beta}, \sigma_\varepsilon^2)$ para todo t.

Estimador MCO. Bajo los supuestos habituales sobre las perturbaciones del modelo aún se pueden derivar analíticamente algunas propiedades para muestras finitas del estimador MCO de $\boldsymbol{\beta}$: es insesgado y eficiente dado que su matriz de varianzas y covarianzas alcanza la cota de Cramer-Rao.

Si se toma esperanzas sobre X en las expresiones (20.1) y (20.2), se puede demostrar que el estimador MCO es insesgado y obtener su matriz de covarianzas. Utilizado el resultado $E(a) = E_b[E(a/b)]$:

$$E(\hat{\boldsymbol{\beta}}) = E_{\mathbf{X}}[E(\hat{\boldsymbol{\beta}}/\mathbf{X})] = E_{\mathbf{X}}[\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\boldsymbol{\varepsilon}/\boldsymbol{X})] = \boldsymbol{\beta}$$

$$V(\hat{\boldsymbol{\beta}}) = E_{\mathbf{X}}[(\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{X}'E(\boldsymbol{\varepsilon\varepsilon}'/\mathbf{X})\,\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}] = \boldsymbol{\sigma}^2 E_{\mathbf{X}}(\mathbf{X}'\mathbf{X})^{-1}$$

donde $E_{\mathbf{X}}(\mathbf{X}'\mathbf{X})^{-1}$ es la matriz de covarianzas poblacional de los regresores calculada en la distribución marginal de X.

Distribución del estimador. En este caso, no se conoce la distribución exacta de los estimadores MCO. En particular, no siguen una distribución normal aun suponiendo que X_{jt} siga una distribución normal para todo $j = 2, \dots, k$ y $t = 1, \dots, T$

Esto se debe a que este estimador, es una combinación no lineal de las variables aleatorias X_{jt} y ε_t. Como consecuencia, los estadísticos de contraste no tienen una distribución exacta conocida y, en particular, no se distribuirán como una t de Student y una F de Snedecor, respectivamente.

Propiedades asintóticas. Ahora bien, bajo los supuestos habituales y si además se satisface el supuesto adicional de que:

$$plim\ \frac{\mathbf{X}'\mathbf{X}}{T} = \mathbf{\Sigma}_{\mathbf{XX}}, donde\ \mathbf{\Sigma}\ es\ una\ matriz\ definida\ positiva \qquad (20.7)$$

es posible derivar las siguientes propiedades asintóticas para los estimadores MCO. Utilizando los teoremas de *Mann-Wald y Cramer e*l estimador por MCO de β es consistente, es decir:

1. $plim\hat{\beta}_j = \beta_j \qquad \forall\ j = 1, \dots, k$
2. $\sqrt{T}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \overset{d}{\rightarrow} N(0, \sigma^2\mathbf{\Sigma}^{-1})$
3. Bajo la hipótesis bula $H_0\colon \mathbf{R}\beta = \mathbf{r}$, los estadísticos t y F usuales se distribuyen asintóticamente como $N(0,1)$ y χ_q^2, respectivamente, donde *q* es el número de restricciones. Por lo tanto, se pueden utilizar estas distribuciones asintóticas para aproximar la distribución exacta de los estadísticos si el tamaño de la muestra es grande.

El supuesto de independencia entre los regresores y el término de perturbación, no se satisface en los ejemplos de los Casos A, B y C anteriores. Luego, este supuesto sigue siendo bastante restrictivo en muchas ocasiones.

Incorrelación contemporánea

Si las variables *aleatorias* X_{jt} y ε_t no son independientes, aunque estén incorrelacionadas contemporáneamente, esto es,

$E(X_{jt}\varepsilon_t) = 0 \;\; \forall\, j = 2, \dots, k$ y $t = 1, \dots, T$

No se pueden derivar analíticamente propiedades para muestras finitas de los estimadores:

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

Estimador MCO. En general, $E[(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\varepsilon]$ puede ser distinto de cero, con lo cual $\hat{\boldsymbol{\beta}}$ puede ser sesgado. Por otro lado, el cálculo analítico de la matriz de varianzas y covarianzas es difícil debido a la no linealidad del estimador en $\mathbf{X}$ **y** $\boldsymbol{\varepsilon}$.

Distribución de estimador. No se conoce su distribución exacta. En particular, no siguen una distribución normal aun suponiendo que X_{jt} siga una distribución normal para todo $j = 2, \dots, k$ y $t = 1, \dots, T$. Como consecuencia, *los estadísticos* no tienen una distribución exacta conocida.

Propiedades asintóticas. Bajo los supuestos habituales más el (20.7) y aplicando los teoremas de *Mann-Wald, Slutzky y Cramer*, se pueden demostrar los resultados asintóticos de la sección anterior.

En este contexto, se enmarcaría el Caso A si $\varepsilon_t \sim iid(0, \sigma^2)$. En este caso, Y_{t-1} no es variable aleatoria independiente de ε_t. Sin embargo, si $E(\varepsilon_t\, \varepsilon_s) = 0 \;\forall t \neq s$, entonces se satisface que $E(Y_{t-1}\, \varepsilon_t) = 0 \;\forall t$. Por lo tanto, regresor y perturbación están *contemporáneamente incorrelacionados*.

Correlación contemporánea

Sucede cuando algunos de los regresores están correlacionados contemporáneamente con el término de perturbación, es decir, $E(X_{jt}\varepsilon_t) = 0$ para al menos algún j y para todo t.

Estimador MCO. En este caso, por las mismas razones que en el anterior, no es posible derivar ninguna propiedad en muestras finitas de los estimadores MCO.

Distribución del estimador y propiedades asintóticas. Se pierden las propiedades asintóticas deseables. No se satisface una de las condiciones del teorema de Mann-Wald, por lo que, en general, el estimador MCO no va a ser consistente, ni va a distribuirse asintóticamente como una normal. Esto lleva a que, bajo la hipótesis nula $H_0\text{: } \mathbf{R}\boldsymbol{\beta} = \mathbf{r}$, los estadísticos t y F, no se distribuyen asintóticamente como una $N(0{,}1)$ y una χ_q^2, respectivamente. Por lo tanto, no se dispone de una distribución asintótica para aproximar la distribución exacta de estos estadísticos, si el tamaño de la muestra es grande.

Estas graves consecuencias, hacen necesario buscar un método de estimación alternativo al de MCO, con el que se obtengan al menos estimadores con propiedades asintóticas deseables y que permita derivar estadísticos con distribuciones asintóticas conocidas para contrastar hipótesis sobre el vector de coeficiente $\boldsymbol{\beta}$.

Este supuesto de correlación contemporánea entre regresor y perturbación, es de gran relevancia en la estimación de muchos modelos econométricos. Por ejemplo, los Casos B y C se enmarcan en este contexto. En el Caso B, el término de perturbación del modelo (20.5) recoge el error de medida ε_t que está correlacionado con X_t; dado que $X_t = X_t^* + \varepsilon_t,\ t = 1, \dots, T$. Aun suponiendo que:

$$E(\mu_t \varepsilon_t) = 0 \text{ y } E(X_t^* \varepsilon_t) = 0$$

$$E(X_t \omega_t) = E[(X_t^* + \varepsilon_t)(\mu_t - \beta \varepsilon_t)] = -\beta E(\varepsilon_t^2) = -\beta V(\varepsilon_t) \neq 0$$

En el Caso C, la variable P_t se determina simultáneamente con Q_t; por lo que si μ_t recoge factores que afectan a Q_t estos afectarán simultáneamente a P_t y $E(P_t \mu_t) \neq 0$.

20.3. Método de Variables Instrumentales

El método de estimación conocido como *método de variables instrumentales* (VI), trata de obtener un estimador consistente de β cuando existen problemas del tipo descripto en la sección anterior; es decir, cuando algunos regresores están correlacionados con el término de perturbación, haciendo que el estimador por MCO no sea consistente.

Se quiere estimar el modelo

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

en donde X es una matriz de variables aleatorias tales que

$$\text{plim } \frac{1}{T}\mathbf{X}'\boldsymbol{\varepsilon} \neq \mathbf{0}$$

El método de variables instrumentales se basa en buscar k variables denominadas instrumentos, Z_{jt} $j = 2,\ldots,k$, que estén por su lado, incorrelacionadas con la perturbación ε_t y por otro, muy correlacionadas con las variables para las que hacen de instrumento, es decir:

1. $E(Z_{jt}\varepsilon_t) = 0 \quad j = 2,\ldots,k;\ \forall t$
2. $plim\ \frac{\mathbf{Z}'\mathbf{X}}{T} = \boldsymbol{\Sigma}_{\boldsymbol{ZX}}, donde\ \boldsymbol{\Sigma}\ finita\ y\ no\ singular$

Hay que tener en cuenta que, para aquellas variables explicativas que no están correlacionadas con el término de perturbación, los mejores instrumentos son ellas mismas.

La matriz de instrumentos $Z\ (Txk)$, se puede construir reemplazando las columnas de X correspondientes a las variables explicativas correlacionadas con la perturbación por las T observaciones de otras variables que satisfagan las condiciones 1 y 2, de forma que el rango de $\mathbf{Z}'\mathbf{X}$ sea completo; es decir, que $\mathbf{Z}'\mathbf{X}$ sea una matriz no singular, ya que el estimador de β de variables instrumentales, se define como:

$$\widehat{\boldsymbol{\beta}}_{VI} = (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{y}$$

En general, es difícil conocer las propiedades del estimador $\widehat{\boldsymbol{\beta}}_{VI}$ para muestras finitas, dado que es un estimador no lineal en las variables aleatorias Z, X y ε. Sin embargo, si se satisfacen las condiciones 1 y 2 y

$$plim\ \frac{\mathbf{Z'Z}}{T} = \mathbf{\Sigma}_{ZZ}, donde\ \mathbf{\Sigma}\ es\ una\ matriz\ finita\ y\ definida\ positiva$$

aplicando el *teorema de Mann-Wald* y el *teorema de Cramer*, se pueden demostrar los siguientes resultados asintóticos:

1. $\widehat{\boldsymbol{\beta}}_{VI}$ es un estimador consistente de $\boldsymbol{\beta}$
2. $\sqrt{T}\left(\widehat{\boldsymbol{\beta}}_{\boldsymbol{VI}} - \boldsymbol{\beta}\right) \overset{d}{\rightarrow} N(\mathbf{0}, \sigma^2 \mathbf{\Sigma}_{\mathbf{XZ}}^{-1} \mathbf{\Sigma}_{\mathbf{ZZ}} \mathbf{\Sigma}_{\mathbf{ZX}}^{-1})$

De esta manera, sea $\boldsymbol{Z}$ una matriz (Txk) de variables instrumentales (VI) tal que:

1) Los instrumentos son independientes del término de error:

$$\text{plim}\ \frac{1}{T}\mathbf{Z}'\boldsymbol{\varepsilon} = \mathbf{0}$$

2) Los instrumentos están correlacionados con los regresores:

$$\text{plim}\ \frac{1}{T}\mathbf{Z}'\mathbf{X} = \frac{\Sigma_{\mathbf{ZX}}}{|\Sigma_{\mathbf{ZX}}|} \neq \mathbf{0}$$

3) Lo instrumentos y su covarianza con $\boldsymbol{y}$ es finita:

$$\text{plim}\ \frac{1}{T}\mathbf{Z}'\mathbf{y} = \Sigma_{\mathbf{Zy}}$$

en donde 3) es una condición de regularidad que evita resultados absurdos. En estas condiciones, la expresión del estimador de $\boldsymbol{\beta}$

mediante variables instrumentales, actuando $\boldsymbol{Z}$ como matriz de instrumentos de $\mathbf{X}$, es:

$$\widehat{\boldsymbol{\beta}}_{\mathbf{VI}} = (\mathbf{Z}'\mathbf{X})^{-\mathbf{1}}\mathbf{Z}'\mathbf{y}$$

y este estimador es consistente, ya que:

$$\widehat{\boldsymbol{\beta}}_{\mathbf{VI}} = \boldsymbol{\beta} + (\mathbf{Z}'\mathbf{X})^{-\mathbf{1}}\mathbf{Z}'\boldsymbol{\varepsilon}$$

Por lo que

$$plim\widehat{\boldsymbol{\beta}}_{\mathbf{VI}} = \boldsymbol{\beta} + \operatorname{plim}\left(\frac{1}{\mathrm{T}}\mathbf{Z}'\mathbf{X}\right)^{-\mathbf{1}}\left(\frac{1}{\mathrm{T}}\mathbf{Z}'\boldsymbol{\varepsilon}\right) = \boldsymbol{\beta}$$

Por otra parte, la matriz de covarianzas del estimador VI es:

$$\operatorname{cov}\left(\widehat{\boldsymbol{\beta}}_{VI}\right) \underbrace{=}_{Asint} \sigma_\varepsilon^2 \frac{1}{T}(\Sigma_{\mathbf{ZX}})^{-\mathbf{1}}\Sigma_{\mathbf{ZZ}}(\Sigma_{\mathbf{ZX}})^{-\mathbf{1}}$$

y puede estimarse mediante la expresión:

$$\widehat{\operatorname{cov}}\left(\widehat{\boldsymbol{\beta}}_{VI}\right) = \hat{\sigma}_\varepsilon^2 \left[\frac{\mathbf{X}'\mathbf{Z}}{T}\right]^{-1} \frac{\mathbf{Z}'\mathbf{Z}}{T} \left[\frac{\mathbf{Z}\mathbf{X}'}{T}\right]^{-1}$$

donde

$$\hat{\sigma}_\varepsilon^2 = \frac{\left(\mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}_{\mathrm{VI}}\right)'\left(\mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}_{\mathrm{VI}}\right)}{T}$$

Donde por ser un método con propiedades asintóticas, no es relevante el ajuste de grados de libertad en el denominador.

La forma de la matriz de covarianzas pone de manifiesto que cuanto mayor sea la correlación entre los instrumentos y las variables explicativas, mayor será la eficiencia del estimador VI.

Por ello, si se tienen k regresores y sólo algunos (por ejemplo, los s primeros) tiene correlación con las perturbaciones, la matriz de instrumentos $\boldsymbol{Z}$ se construirá:

• sustituyendo las s primeras columnas de $\mathbf{X}$ por los correspondientes instrumentos, y

• manteniendo los restantes regresores, que actuarán como instrumentos de sí mismos.

Para contrastar hipótesis del tipo $H_0: \mathbf{R\beta} = \mathbf{r}$, se utiliza el estadístico:

$$F = \frac{\left(\mathbf{R}\hat{\beta}_{VI} - \mathbf{r}\right)' [\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{Z})^{-1}(\boldsymbol{Z}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}\left(\mathbf{R}\hat{\beta}_{VI} - \mathbf{r}\right)}{\hat{\sigma}_{VI}^2}$$

Este estadístico se distribuye asintóticamente como una χ_q^2, donde q es el número de restricciones.

Ejemplo 20.3 Se considera de nuevo el modelo de consumo de Haavelmo, dado por:

$$\begin{aligned} (1) \quad & \mathrm{C_t} = \beta \mathrm{PBI_t} + \varepsilon_t \\ (2) \quad & \mathrm{PBI_t} = \mathrm{C_t} + \mathrm{I_t} \end{aligned}$$

Para ver que $PIBt$ es un regresor estocástico de (1), correlacionado con el término de error, basta sustituir (1) en (2):

$$(3) \qquad \mathrm{PBI_t} = \beta \mathrm{PBI_t} + \varepsilon_t + \mathrm{I_t} \Leftrightarrow \mathrm{PBI_t} = \frac{\mathrm{I_t}}{1-\beta} + \frac{\varepsilon_t}{1-\beta}$$

por tanto

$$E\left[\left(\mathrm{PBI_t} - \frac{\mathrm{I_t}}{1-\beta}\right)\varepsilon_t\right] = \frac{\sigma_\varepsilon^2}{1-\beta} \neq 0$$

y el estimador MCO de (1) es inconsistente

Para estimar consistentemente (1) puede usarse el estimador VI, usando It como instrumento de $PIBt$. La expresión del estimador quedaría:

$$\hat{\beta}_{\mathrm{VI}} = \frac{\sum \mathrm{I_t C_t}}{\sum \mathrm{I_t PIB_t}}$$

Ejemplos 20.4 Sea el modelo dinámico:

$$(1) \qquad y_t = \rho y_{t-1} + \beta X_t + \varepsilon_t$$

con los siguientes términos de error:

$$(2.a) \qquad \varepsilon_t = a_t$$
$$(2.b) \qquad \varepsilon_t = a_t - \theta a_{t-1}$$
$$(2.c) \qquad \varepsilon_t = \varphi \varepsilon_{t-1} a_t$$

En este caso:

• La estimación del modelo (1) por MCO es consistente si el error es (2.a)

• La estimación del modelo (1) por MCO es inconsistente si el error es (2.b) o (2.c), ya que y_{t-1} tiene covarianza no nula con a_t

- El modelo (1)-(2.b) puede estimarse consistentemente usando x_{t-1} como instrumento de y_{t-1} y la variable x_t como instrumento de sí misma.

Alternativamente, y_{t-2} es un instrumento válido de y_{t-1}.

- El modelo (1)-(2.c) puede estimarse consistentemente usando x_{t-1} como instrumento de y_{t-1} y la variable x_t como instrumento de sí misma. No hay ningún instrumento válido para estimar por VI el modelo si $\beta = 0$.

20.4 Contraste de Hausman

Cuando en un modelo de regresión lineal general los regresores son estocásticos, es necesario añadir la siguiente hipótesis complementaria al modelo para garantizar la consistencia de la estimación MCO de los coeficientes de regresión:

> Los regresores no están correlacionados con el término de perturbación de forma que, bajo ciertas condiciones de regularidad, se cumple que $\frac{\mathrm{X}'\varepsilon}{\mathrm{T}} = \mathbf{0}$.

Como se ha visto en los apartados anteriores, este supuesto garantiza que el estimador MCO de los coeficientes de regresión β, es consistente. Existen casos en los cuales esta hipótesis no se satisface; por ejemplo, si algún regresor está medido con error, si se omiten variables relevantes, si hay problema de simultaneidad, etc.

Hausman ha desarrollado un procedimiento para contrastar el cumplimiento de esta hipótesis. Este contraste se puede interpretar también, en términos generales, como un contraste de mala especificación de la parte sistemática del modelo.

El mecanismo de contraste es el siguiente. La hipótesis nula, es:

$$\mathrm{H}_0\colon \mathrm{plim}\, \frac{\mathbf{X}'\boldsymbol{\varepsilon}}{\mathrm{T}} = \mathbf{0}$$

En el modelo de regresión uniecuacional, el estadístico de contraste, se basa en la diferencia de los estimadores de los coeficientes de regresión: $\widehat{\boldsymbol{\beta}}_{MCO}$ y $\widehat{\boldsymbol{\beta}}_{VI}$. Bajo la H_0 y suponiendo que se cumplen los supuestos básicos sobre la perturbación, se puede demostrar, bajo ciertas condiciones de regularidad, que:

$\widehat{\boldsymbol{\beta}}_{MCO}$ y $\widehat{\boldsymbol{\beta}}_{VI}$ son consistentes.

$\widehat{\boldsymbol{\beta}}_{MCO}$ es asintóticamente eficiente.

Las distribuciones asintóticas, son:

$$\sqrt{T}\left(\widehat{\boldsymbol{\beta}}_{MCO} - \boldsymbol{\beta}\right) \overset{d}{\to} N(\mathbf{0}, \mathbf{V_1})$$

$$\sqrt{T}\left(\widehat{\boldsymbol{\beta}}_{\boldsymbol{VI}} - \boldsymbol{\beta}\right) \overset{d}{\to} N(\mathbf{0}, \mathbf{V_2})$$

Donde $\mathbf{V_1}$ y $\mathbf{V_2}$ son una matrices definidas positivas.

Bajo la hipótesis alternativa, solo es consistente el estimador $\widehat{\boldsymbol{\beta}}_{VI}$. Por lo tanto, si los regresores y la perturbación están correlacionados, ambos estimadores tenderán a diferir dado que $\widehat{\boldsymbol{\beta}}_{VI}$ es consistente y converge a $\boldsymbol{\beta}$; mientras que, $\widehat{\boldsymbol{\beta}}_{MCO}$ no es consistente y convergerá a un valor distinto de $\boldsymbol{\beta}$.

El estadístico de contraste es:

$$\mathrm{H} = T\left(\widehat{\boldsymbol{\beta}}_{MCO} - \widehat{\boldsymbol{\beta}}_{VI}\right)'\left(\widehat{\mathbf{V}}_{\mathbf{2}} - \widehat{\mathbf{V}}_{\mathbf{1}}\right)^{-\mathbf{1}}\left(\widehat{\boldsymbol{\beta}}_{MCO} - \widehat{\boldsymbol{\beta}}_{VI}\right)$$

Donde $\widehat{\mathbf{V}}_{\mathbf{1}}$ y $\widehat{\mathbf{V}}_{\mathbf{2}}$ son estimadores consistentes de $\mathbf{V}_{\mathbf{1}}$ y $\mathbf{V}_{\mathbf{2}}$, respectivamente. Bajo H_0 el estadístico H se distribuye asintóticamente como una χ^2 con k grados de libertad. Se rechaza la H_0 con un nivel de significación α, si $\mathrm{H} > \chi^2_{\alpha;k}$.

CASOS DE ESTUDIO, PREGUNTAS Y PROBLEMAS

Caso 20.1. Modelo Keynesiano Simple

Hasta ahora se estudió exclusivamente la estimación de relaciones lineales únicas de variables económicas. La mayor parte de los estudios económicos basan sus teorías en modelos con varias ecuaciones, en forma de sistemas de relaciones económicas. Cuando una relación es parte de un sistema, algunos regresores serán estocásticos y no serán independientes de las perturbaciones; entonces la estimación clásica por mínimos cuadrados será inconsistente y se deberán desarrollar procedimientos especiales para estimaciones consistentes.

Quizá el modelo más familiar en los libros de economía es el sistema keynesiano simple

$$(1)\quad C_t = \alpha + \beta Y_t$$

$$(2)\quad Y_t = C_t + I_t$$

Donde,

C : Consumo;

Y : Renta;

I : Inversión

t : Unidades de observaciones (temporales o transversales), $\forall t = 1, \ldots, T$

La interpretación típica de este modelo es que (1) representa la ecuación de comportamiento de los consumidores y que (2) es una condición de equilibrio que iguala el ahorro $(Y - C)$ a la inversión; y que la inversión es autónoma. Esto es, dada una inversión el modelo determina los valores de equilibrio del consumo y de la renta. Se ve que tanto el consumo como la renta dependen de la inversión. Es decir, si se resuelve el sistema sin tener en cuenta los subíndices:

$$(3)\quad C = \frac{\alpha}{1-\beta} + \frac{\beta}{1-\beta} I,$$

$$(4)\quad Y = \frac{\alpha}{1-\beta} + \frac{1}{1-\beta} I$$

Hasta aquí el modelo es exacto y, por tanto, obviamente incongruente con una descripción empírica de la economía. Una formulación econométrica del sistema es,

$$(5)\quad C_t = \alpha + \beta Y_t + \varepsilon_t$$

$$(6)\quad Y_t = C_t + I_t$$

Donde $\boldsymbol{\varepsilon}$ es el vector de orden $T \times 1$ que representa a la perturbación aleatoria con,

$$(7)\quad E(\varepsilon_t) = 0,\quad E(\varepsilon_t^2) = \sigma^2,\quad E(\varepsilon_t \varepsilon_s) = 0 \text{ para todo } s, t;\ s \neq t$$

Para mantener la idea que la inversión es autónoma determinada fuera del sistema, se supone que,

$$(8)\quad I_t \text{ y } \varepsilon_t \text{ son independientes } (t = 1, \ldots, T;\ s = 1, \ldots, T).$$

Se tiene ahora la dependencia explícita de $C \text{ e } Y \text{ sobre } I \text{ y } \varepsilon$, resolviendo el sistema

$$(9)\quad C_t = \frac{\alpha}{1-\beta} + \frac{\beta}{1-\beta} I_t + \frac{1}{1-\beta} \varepsilon_t$$

$$(10)\quad Y_t = \frac{\alpha}{1-\beta} + \frac{1}{1-\beta} I_t + \frac{1}{1-\beta} \varepsilon_t$$

Dada una muestra de observaciones conjuntas sobre $C, Y \text{ e } I$, el interés se basa en estimar los parámetros de la función consumo (5). Ahora bien, en esa ecuación el regresor y la perturbación no son estadísticamente independientes, ni temporal ni contemporáneamente. Se puede encontrar la covarianza de $\mathbf{Y} \text{ y } \boldsymbol{\varepsilon}$ multiplicando (10) por ε_t y tomando esperanzas:

$$(11)\quad E(Y_t\varepsilon_t) = \frac{\alpha}{1-\beta}E(\varepsilon_t) + \frac{1}{1-\beta}E(I_t\varepsilon_t) + \frac{1}{1-\beta}E(\varepsilon_t^2)$$
$$= \frac{1}{1-\beta}\sigma^2 \neq 0, \text{ utilizando (7) y (8).}$$

Así, para la estimación de (5) MCO no produciría estimaciones consistentes.

Si se considera esto explícitamente, el estimador clásico de β en (5) es

$$(12)\quad b = \frac{\sum(C-\overline{C})(Y-\overline{Y})}{\sum(Y-\overline{Y})^2} = \frac{\sum(\alpha+\beta Y+\varepsilon-\alpha-\beta\overline{Y})(Y-\overline{Y})}{\sum(Y-\overline{Y})^2}$$
$$= \frac{\sum\beta(Y-\overline{Y})^2+\varepsilon(Y-\overline{Y})}{\sum(Y-\overline{Y})^2} = \beta + \frac{\sum\varepsilon(Y-\overline{Y})}{\sum(Y-\overline{Y})^2} =$$
$$= \beta + \frac{\sum\varepsilon(Y-\overline{Y})/T}{\sum(Y-\overline{Y})^2/T}$$

Ahora bien, $\sum\varepsilon(Y-\overline{Y})/T$ es la *covarianza muestral*, así que bajo condiciones generales

$$(13)\quad P\lim\frac{\sum\varepsilon(Y-\overline{Y})}{T} = E(\varepsilon_t Y_t) = (1-\beta)^{-1}\sigma^2$$

Similarmente, $\sum(Y-\overline{Y})^2/T$ es la *varianza muestral*, que bajo condiciones generales

$$(14)\quad P\lim\frac{\sum(Y-\overline{Y})^2}{T} = E[Y-E(Y_t)]^2 = \sigma_{yy}$$

Entonces,

$$(15)\quad P\lim b = \beta + \frac{(1-\beta)^{-1}\sigma^2}{\sigma_{yy}}$$

Por lo que el estimador MCO no sería consistente. Realmente está clara la dirección del sesgo asintótico si se emplea la información económica de que la propensión marginal a consumir está entre cero y uno: con $0<\beta<1, P\lim b>\beta.$

También es informativa una expresión distinta del sesgo. Como I y ε no están correlacionados (10) implica

$$\text{(16)} \quad Var(Y)=\sigma_{YY}=E[Y-E(Y_t)]^2=(1-\beta)^{-2}\left\{E[I-E(I)]^2+E(\varepsilon^2)\right\}$$
$$=(1-\beta)^{-2}(\sigma_{ii}+\sigma^2)$$

Introduciendo esto para σ_{yy} en (15), se encuentra que

$$\text{(17)} \quad P\lim b=\beta+\frac{(1-\beta)^{-1}\sigma^2}{(1-\beta)^{-2}(\sigma_{ii}+\sigma^2)}=\beta+(1-\beta)\frac{\sigma^2}{\sigma_{ii}+\sigma^2}$$

Nuevamente con $0<\beta<1, \mathrm{P}\lim \mathrm{b}>\beta.$ Más aún, $\mathrm{P}\lim \mathrm{b}-\beta$ será grande cuando la varianza de las perturbaciones es grande en relación con la varianza de la inversión. Una interpretación heurística del resultado es que la regresión clásica MCO del consumo sobre la renta da crédito a la renta debido al efecto de las perturbaciones puesto que éstas están correlacionadas positivamente con la renta.

Otra forma de mirar el resultado es considerar que MCO puede suministrar estimaciones consistentes, cuando los parámetros en las relaciones son los de la esperanza condicionada del regresando dados los regresores. Pero no es este el caso en (5), ya que

$$\text{(18)} \quad \mathrm{E}(\mathrm{C_t}\mid \mathrm{Y_t})=\alpha+\beta \mathrm{Y_t}+\mathrm{E}(\varepsilon_t\mid \mathrm{Y_t})$$

Pero, $\mathrm{E}(\varepsilon_t\mid \mathrm{Y_t})\neq \mathrm{E}(\varepsilon_t)=0$

Aunque esta forma de mirar el resultado recuerda que los MCO deberían ser apropiados para resolver las relaciones (9) y (10), porque (9) cae bajo el modelo de regresión lineal estocástico independiente, se observa que

$$(19)\quad \mathrm{E}(\mathrm{C_t}\,|\,\mathrm{I_t}) = \frac{\alpha}{1-\beta} + \frac{\beta}{1-\beta}\mathrm{I_t} + \mathrm{E}\left(\frac{1}{1-\beta}\varepsilon_{\mathrm{t}}\,|\,\mathrm{I_t}\right) = \\ = \frac{\alpha}{1-\beta} + \frac{\beta}{1-\beta}\mathrm{I_t}$$

Estimando por MCO

$$(20)\quad \mathrm{C_t} = \pi_0 + \pi_1 \mathrm{I_t} + \nu_{\mathrm{t}}$$

Donde

$$(21)\quad \pi_0 = \frac{\alpha}{1-\beta},\quad \pi_1 = \frac{\beta}{1-\beta},\quad \nu_{\mathrm{t}} = \frac{1}{1-\beta}\varepsilon_{\mathrm{t}}$$

designando las estimaciones MCO de π_0, π_1 como $\mathrm{p}_0, \mathrm{p}_1$, los cuales son consistentes, ya que

$$(22)\quad \mathrm{P}\lim \mathrm{p}_0 = \pi_0 = \frac{\alpha}{1-\beta},\quad \mathrm{P}\lim \mathrm{p}_1 = \pi_1 = \frac{\beta}{1-\beta}$$

Continuando con el modelo Keynesiano, (21) sugiere que se denomine a la estimación de β por $\hat{\beta}$, definido por $\mathrm{p}_1 = \hat{\beta}/(1-\hat{\beta})$, esto es

$$(23)\quad \hat{\beta} = \frac{\mathrm{p}_1}{1+\mathrm{p}_1}$$

$\hat{\beta}$ es consistente ya que

$$(24)\quad \text{P}\lim\hat{\beta}=\frac{\text{P}\lim \text{p}_1}{\text{P}\lim(1+\text{p}_1)}=\frac{\beta/(1-\beta)}{1+\beta/(1-\beta)}=\beta$$

Igualmente (21) sugiere que se denomine a la estimación de α por $\hat{\alpha}$, definido por $p_0=\hat{\alpha}/(1-\hat{\beta})$, esto es

$$(25)\quad \hat{\alpha}=\text{p}_0(1-\hat{\beta})$$

Por lo tanto $\hat{\alpha}$ es consistente

$$(26)\quad P\lim\hat{\alpha}=P\lim p_0 P\lim(1-\hat{\beta})=\frac{\alpha(1-\beta)}{(1-\beta)}=\alpha$$

Debe observarse que aunque p_0,p_1 son insesgados $\hat{\alpha},\hat{\beta}$, que son funciones no lineales de p_0,p_1, no son insesgados. Aunque sí consistentes y, por tanto, insesgados asintóticamente.

En resumen, el Modelo Keynesiano Simple demuestra que cuando una relación es una de las muchas de un sistema simultáneo, las estimaciones clásicas MCO de sus coeficientes serán generalmente inconsistentes. La razón subyacente es que algunos regresores están determinados conjuntamente con el regresando y, por tanto, son dependientes de la perturbación contemporánea. Se ha visto también que se pueden obtener estimaciones consistentes mediante una especie de procedimiento indirecto mínimo cuadrático. Sin embargo, se verá que esta última alternativa no es por lo general aprovechable.

Por supuesto que sí es aprovechable el método de variables instrumentales; no es difícil demostrar que los estimadores $\hat{\alpha},\hat{\beta}$ son los estimadores de variables instrumentales de α,β en (5), donde I, que es independiente de las perturbaciones, se utiliza como instrumento para Y. Sin embargo, no siempre será tan simple encontrar una variable instrumental legítima.

Alternativamente, se podría haber mirado (10) y observado que:

$$(27)\quad E(Y_t \mid I_t) = \frac{\alpha}{1-\beta} + \frac{1}{1-\beta} I_t + E\left(\frac{1}{1-\beta}\varepsilon_t \mid I_t\right) = \\ = \frac{\alpha}{1-\beta} + \frac{1}{1-\beta} I_t$$

Y estimado por MCO

$$(28)\quad Y_t = \delta_0 + \delta_1 I_t + \omega_t$$

Donde

$$(29)\quad \delta_0 = \frac{\alpha}{1-\beta},\quad \delta_1 = \frac{1}{1-\beta},\quad \omega_t = \frac{1}{1-\beta}\varepsilon_t$$

Entonces las estimaciones MCO designadas por d_0, d_1 serán consistentes:

$$(30)\quad P\lim d_0 = \delta_0 = \frac{\alpha}{1-\beta},\quad P\lim d_1 = \delta_1 = \frac{\beta}{1-\beta}.$$

Entonces se podría haber considerado los estimadores $\tilde{\alpha}, \tilde{\beta}$ definidos por $d_1 = 1/(1-\tilde{\beta})$ y $d_0 = \tilde{\alpha}/(1-\tilde{\beta})$ esto es,

$$(31)\quad \tilde{\beta} = 1 - \frac{1}{d_1},\quad \tilde{\alpha} = \frac{d_0}{d_1}$$

Y ver que eran consistentes;

$$(32) \quad P\lim\tilde{\beta} = \beta, \quad P\lim\tilde{\alpha} = \alpha.$$

Sin embargo, no hay por qué hacer esto en el presente modelo; utilizando (6) se puede demostrar que $\tilde{\beta} = \hat{\beta}$ y que $\tilde{\alpha} = \hat{\alpha}$.

Con algunos datos de Haavelmo para la economía de Estados Unidos referidos a consumo, renta e inversión, se calculan los siguientes momentos alrededor de la media:

	C	Y	I
C	35.887	47.585	11.698
Y		64.993	17.408
I			5.710

La estimación mínimo cuadrática clásica inconsistente de β en (5) es entonces

$$(33) \quad \hat{\beta} = \frac{m_{cy}}{m_{yy}} = \frac{47.585}{64.993} = 0{,}732$$

La estimación mínimo cuadrática clásica consistente para π_1 en (20) es

$$(34) \quad p_1 = \frac{m_{ci}}{m_{ii}} = \frac{11.698}{5.710} = 2{,}048$$

De esto se puede deducir una estimación consistente de β a través de (23)

$$(35) \quad \hat{\beta} = \frac{p_1}{1 + p_1} = \frac{2,048}{3,048} = 0,672$$

Se observa para esta muestra que $\hat{\beta} > \beta$, lo que no es sorprendente puesto que $Plim\, b > \beta = Plim\hat{\beta}$.

De esta forma se puede tomar la estimación mínimo cuadrática clásica de δ_1 en (28):

$$(36) \quad d_1 = \frac{m_{yi}}{m_{ii}} = \frac{17.408}{5.710} = 3,048$$

Y de esto deducir una estimación consistente de β a través de (31)

$$(37) \quad \tilde{\beta} = 1 - \frac{1}{d_1} = 1 - \frac{1}{3,048} = 0,672 = \hat{\beta}$$

Y también, para la estimación de la variable instrumental de β en (5) se ve que

$$(38) \quad b^* = \frac{m_{ic}}{m_{iy}} = \frac{11.698}{17.408} = 0,672 = \hat{\beta}.$$

Caso 20.2 Modelo simple de mercado

Para una segunda demostración, se considera el modelo de la Oferta y Demanda para una mercancía en particular -con una perturbación que permite desplazamientos aleatorios en las curvas de oferta y demanda-.

(39) Demanda $q_t = \alpha + \beta p_t + \varepsilon_t$

(40) Oferta $q_t = \gamma + \delta p_t + \varepsilon^*{}_t$

Si en la ecuación de demanda, un regresor p_t fuera independiente de la perturbación ε_t, entonces cuando la ecuación de demanda recibe una perturbación positiva q_t en (39) debería elevarse en la cantidad de la perturbación. Entonces, en (40), esto haría a q_t independiente de p_t y ε_t dando lugar a que $\varepsilon_t = \varepsilon^*{}_t$.

Aunque las perturbaciones de la demanda y de la oferta pueden estar correlacionadas, sin embargo, es absurdo pensar que sean idénticas. Se concluye que p_t y ε_t no son independientes, el precio está determinado conjuntamente por la cantidad y por los desplazamientos aleatorios de la ecuación de demanda.

Preguntas

1-El método de variables instrumentales se utiliza:

a) cuando hay heterocedasticidad

b) cuando además de heterocedasticidad hay autocorrelación

c) se puede aplicar cuando estimamos una ecuación estructural aislada de un modelo de ecuaciones simultáneas

d) se aplica cuando se tienen instrumentos estadísticos adecuados: medias, covarianzas, límites de probabilidad, etc.

2 -El estimador de variables instrumentales:

a) es consistente

b) los residuos se pueden expresar como $Y - X\beta_{VI} = \varepsilon + X(Z'X)^{-1}Z'\varepsilon$

c) para obtener la matriz de varianzas-covarianzas de los estimadores hay que estimar previamente el estimador de la varianza de las perturbaciones

d) ninguna de las anteriores

3.- El método de variables instrumentales:

a) Se debe emplear para solucionar los problemas de multicolinealidad, empleando instrumentos de las exógenas correlacionadas

b) Se debe emplear cuando hay exógenas retardadas en el segundo miembro de la ecuación de regresión

c) Se debe emplear cuando el coeficiente de determinación no es un buen instrumento para medir la bondad del ajuste.

d) Se debe emplear como alternativa consistente a MCO cuando se piensa que alguna/s explicativa/s de la ecuación son endógenas y por tanto no independientes de la perturbación

4.- Si una de las variables explicativas del modelo de regresión es la propia endógena retardada y las perturbaciones presentan autocorrelación, es un caso de:

a) Regresores estocásticos dependientes de las perturbaciones a lo largo de toda la muestra

b) Regresares estocásticos dependientes de las perturbaciones pasadas, pero no de las contemporáneas ni futuras

c) Regresores estocásticos, pero no se sabe si dependientes o independientes de las perturbaciones

d) Regresores correlacionados entre sí. Un caso, por tanto, de multicolinealidad

5.- Si hay variables endógenas retardadas como explicativas en un modelo de regresión cuyas perturbaciones son ruido blanco

a) Los estimadores MCO son sesgados e inconsistentes

b) $plim\left(\frac{X'\varepsilon}{T}\right) = 0$

c) $plim\left(\frac{X'X}{T}\right)^{-1} = 0$

d) Ninguna de las anteriores

Problema 20.1. Modelo con errores de medida

Dada el siguiente modelo experimental con errores de medida de las dos variables que intervienen, Y^* y X^*:

$$Y_t^* = \alpha + \beta X_t^* + \mu_t$$

Con $\alpha = 1$ y $\beta = 0.5$

$$Y_t = Y_t^* + \varepsilon_t \qquad \varepsilon_t \sim N(0;\ 1)$$

$$X_t = X_t^* + \omega_t \qquad \omega_t \sim N(0;\ 2.25)$$

Este es el Caso B de la parte teórica, es decir hay dependencia contemporánea y no contemporánea entre X y μ. En la simulación se miden con error tanto Y^* y X^*. Sin embargo, los errores de medida de la endógena no afectan a la aleatoriedad de los regresores.

Los datos de ambas variables endógena y exógena contienen errores aleatorios de medida, distribuidos normalmente.

A partir de valores fijados de X^* y de simulaciones de ambos errores, se han obtenido valores de Y^* y de Y, y se ha estimado el modelo de regresión de Y en función de X, con 100 replicaciones para cada tamaño muestral entre $T = 10$ y $T = 100$. La tabla siguiente proporciona los resultados de las estimaciones

T	$E(\hat{\alpha})$	$\hat{\sigma}^2_{\hat{\alpha}}$	$ECM(\hat{\alpha})$	$E(\hat{\beta})$	$\hat{\sigma}^2_{\hat{\beta}}$	$ECM(\hat{\beta})$
10	1.56	0.59		0.399	0.014	
30	1.5	0.13		0.402	0.005	
50	1.49	0.09		0.402	0.003	
80	1.52	0.06		0.395	0.002	
100	1.55	0.04		0.393	0.001	

T Tamaño muestral
$E(\alpha)$ valor medio del estimador de α obtenido con las 100 muestras de tamaño T
σ^2_{α} Varianza de las 100 estimaciones de α obtenidas con las muestras de tamaño T
$ECM(\alpha)$ Error cuadrático medio de las 100 estimaciones de α obtenidas con las muestras de tamaño T
Los resultados para β tienen la misma interpretación.

Se pide:

1. Calcular el error cuadrático medio
2. Los estimadores de los parámetros son insesgados o sesgados. ¿El sesgo se mantiene aun en muestras grandes?
3. Cuando aumenta el tamaño de la muestra ¿los estimadores son más precisos? ¿Por qué?
4. ¿El ECM de los estimadores tiende a cero cuando aumenta el tamaño de la muestra? Entonces los estimadores MCO, en el caso de errores de medida ¿son consistentes o inconsitentes?

Problema 20.2. Modelo con endógena retardada

Se realiza la simulación de un modelo que contiene la variable endógena rezagada como explicativa. Los errores del modelo son homocedásticos y no autocorrelacionados. El diseño del experimento es el siguiente:

$Y_t = \alpha + \beta_1 X_t + \beta_2 Y_{t-1} + \mu_t$; con

$$\alpha = 1$$
$$\beta_1 = 0.5$$
$$\beta_2 = 0.5$$
$$\mu \sim N(0; 0.05)$$

Los resultados se muestran en la tabla siguiente:

T	$E(\hat{\alpha})$	$\hat{\sigma}^2_{\hat{\alpha}}$	$ECM(\hat{\alpha})$	$E(\hat{\beta}_1)$	$\hat{\sigma}^2_{\hat{\beta}_1}$	$ECM(\hat{\beta}_1)$	$E(\hat{\beta}_2)$	$\hat{\sigma}^2_{\hat{\beta}_2}$	$ECM(\hat{\beta}_2)$
10	1.11	0.22		0.89	0.13		0.10	0.13	
30	1.23	0.14		0.69	0.04		0.31	0.04	
50	1.11	0.06		0.58	0.02		0.42	0.02	
80	1.09	0.05		0.55	0.01		0.45	0.01	
100	1.05	0.04		0.53	0.01		0.47	0.01	

Con las mismas consignas del problema anterior, cómo interpreta los resultados en este caso.

Problema 20.3 Cálculo del estimador de variable instrumental (VI)

Suponga el siguiente modelo de dos ecuaciones simultáneas:

$$Y1_t = \alpha_0 + \alpha_1 Y2_t + \alpha_2 X1_t + \alpha_3 X2_t + \mu_{1t}$$

$$Y2_t = \beta_0 + \beta_1 Y1_t + \beta_2 X3_t + \mu_{2t}$$

$Y1$ e $Y2$ son las dos variables endógenas del modelo y por tanto, estocásticas (dependen de las perturbaciones). Las X se suponen exógenas, o al menos independientes de las perturbaciones para todas las observaciones. Se está interesado en estimar consistentemente la primera ecuación. Como uno de los regresares es $Y2$, que es endógena del modelo, los estimadores de MCO son inconsistentes. Emplear VI, tomando como instrumento de $Y2$ a $X3$, ya que esta última variable causa a $Y2$.

La muestra es de tamaño 14 se encuentra en la tabla 20.1

El estimador de $\boldsymbol{\beta}$ por variables instrumentales es

$$\boldsymbol{\beta} = (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Y}$$

La matriz $\mathbf{X}$ contiene las variables $\mathbf{Y2}$, $\mathbf{X1}$ y $\mathbf{X2}$

$$\mathbf{X} = \begin{bmatrix} 1 & 84{,}61691 & 3 & 54 \\ 1 & 47{,}58897 & 6 & 21 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 35{,}98463 & 5 & 25 \end{bmatrix}$$

y la matriz $\mathbf{Z}$ las variables $\mathbf{X3}$, $\mathbf{X1}$ y $\mathbf{X2}$

$$\mathbf{Z} = \begin{bmatrix} 1 & 77 & 3 & 54 \\ 1 & 28 & 6 & 21 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 27 & 5 & 25 \end{bmatrix}$$

Donde $\mathbf{X3}$ es el instrumento de $\mathbf{Y2}$, siendo $\mathbf{X1}$ y $\mathbf{X2}$ instrumentos de sí mismas.

Tabla 20.1

	Y1	Y2	X1	X2	X3
1	152,001	84,61691	3	54	77
2	67,46811	47,58897	6	21	28
3	118,165	73,37928	9	32	39
4	231,1687	130,7999	13	96	81
5	202,366	110,4916	7	64	73
6	130,8377	69,70934	45	22	64
7	138,2222	66,55147	3	55	29
8	64,03722	13,43419	1	21	24
9	187,0715	107,7802	9	74	27
10	270,6937	169,057	8	95	53
11	113,4228	44,05012	6	32	21
12	129,2074	71,50718	4	48	10
13	211,6246	110,3704	4	65	85
14	48,7462	35,98463	5	25	27

Cómo elegir adecuadamente los instrumentos: Para aplicar el método de VI. la elección de los instrumentos Z es una decisión clave. En el ejemplo anterior, se tomó como instrumento de $Y2$ a $X3$, porque esta variable explica a $Y2$. Los otros dos regresares $X1$ y $X2$, y la constante, jugaron el papel de instrumentos de sí mismos.

En un modelo de ecuaciones simultáneas, hay una posibilidad sensata de elegir instrumentos. Consiste en hacer primero una regresión de la variable explicativa que es endógena de otra ecuación (en el ejemplo, $Y2$) contra todas las variables predeterminadas del modelo en su conjunto, es decir, contra todas las exógenas y las endógenas retardadas que aparecen en cualquier ecuación del modelo. Se calculan y se guardan en memoria los valores ajustados de esta primera regresión, la que se llama primera etapa. En una segunda fase, se aplica el método de VI, utilizando como instrumento ese valor ajustado. Este método de estimación es el llamado de mínimos cuadrados en dos etapas (MC2E).

Con los datos del problema realice la estimación por MC2E utilizando Eviews.

Bibliografía

◦ Creel, Michael. *Econometrics* Barcelona: Dept. of Economics and Economic History, Universitat Autònoma de Barcelona, 2005.

◦ Fernandez Sainz, A.I - Gonzalez Casimiro, P - Regules Castillo, M - Moral Zuazo, M.P - Esteban Gonzalez, MV. *Ejercicios De Econometría* McGraw Hill, 2005.

◦ Jerez, Miguel. *Econometría Ii* Madrid: Universidad Complutense de Madrid, 2004.

◦ Murillo Fort, C. - González López-Valcárcel, B. "Manual De Econometría," Universidad Pompeu Fabra – Universidad de Las Palmas de GC