ner las restricciones de exclusión de las variables que generan el problema de colinealidad.

Los estimadores sin restringir y restringido $\hat{\boldsymbol{\beta}}_{MCO}$ y $\hat{\boldsymbol{\beta}}_R$ pueden compararse de acuerdo con distintos criterios. Una primera posibilidad consistiría en escoger el estimador restringido $\hat{\boldsymbol{\beta}}_R$ si la diferencia entre las matrices de ECM de $\hat{\boldsymbol{\beta}}_{MCO}$ y $\hat{\boldsymbol{\beta}}_R$ es definida positiva. Alternativamente, podría escogerse el estimador restringido si la traza de la primera es superior a la de la última, y el estimador no restringido en caso contrario.

Wallace y Toro-Vizcarrondo (1968 y 1969) demostraron que el estimador restringido es preferible, de acuerdo con el primer criterio, si se satisface la desigualdad:

$$\mu = \frac{(\mathbf{r} - \mathbf{R}\boldsymbol{\beta})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{r} - \mathbf{R}\boldsymbol{\beta})}{2\sigma_u^2} \leqslant \frac{1}{2}$$

Cuando las hipótesis lineales son ciertas, el estadístico que se obtiene al sustituir en μ los valores de $\boldsymbol{\beta}$ y σ_u^2 por sus estimaciones, sigue una distribución $F_{q,T-k}$ multiplicada por $q/2$. Cuando las hipótesis no son ciertas, el estadístico resultante sigue una distribución $F_{q,T-q}$ no centrada, con parámetro de no centralidad igual al valor del estadístico μ. Los autores citados sugieren comparar el valor de μ con los valores de una distribución $F_{q,T-k}$ con parámetro de no centralidad igual a 1/2. Si el valor de μ supera el valor tabulado de esta distribución, ello es señal de que el parámetro de no centralidad es grande (por tanto, superior a 1/2), y el estimador MCO sin restringir se considera preferible. Si el estadístico μ supera el valor de las tablas, pero es inferior a 1/2, se entiende que las restricciones $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ son falsas, pero su imposición genera un estimador que, aun siendo sesgado, tiene un ECM inferior al estimador sin restringir (que es insesgado).

El estimador restringido es preferible de acuerdo con el segundo criterio (comparación de las trazas de las matrices ECM) si $\mu \leqslant q/2$, donde q es el número de restricciones lineales que se contrasta, por lo que habría que comparar esta vez el valor de μ con una distribución $F_{q,T-k}$ con parámetro de no centralidad igual a $q/2$. El conjunto de restricciones lineales se rechaza si el estadístico μ supera al valor de las tablas de la distribución citada.

10.6. ERRORES DE MEDIDA

Una dificultad que subyace a casi todo trabajo empírico en Economía es la imposibilidad de disponer de observaciones muestrales de las variables que se pretende relacionar. Por ejemplo, las variables de Contabilidad Nacional, como el PIB, el stock de capital, o el consumo, no son sino estimaciones de conceptos teóricos que no se observan en la realidad. En otros casos, como ocurre con la Renta Permanente, la inteligencia, o la capacidad de un trabajador para realizar una determinada tarea, no disponemos ni siquiera de estimaciones, por lo que suelen utilizarse *variables proxy*, que aproximan los conceptos que se quieren utilizar. Así, por ejemplo, se utilizan los años de experiencia en el puestro de trabajo como *proxy* de la *habilidad* de un trabajador.

Basándonos en los desarrollos de capítulos previos, cabe esperar que los *errores de medida*, así como el uso de *variables proxy*, introduzcan sesgos en muestras finitas en el estimador MCO. En realidad se trata de una combinación de variables omitidas (si bien no totalmente), a la vez que de inclusión de variables irrelevantes (aunque no completamente). El sesgo será menor: *a*) cuanto más se aproxime la variable que realmente se incluye en el modelo a la variable que teóricamente debió utilizarse, así como *b*) cuanto más independiente sea el error de medida de las restantes variables del modelo.

Consideremos el modelo lineal simple:

$$y_i = \beta x_i + u_i, \qquad i = 1, 2, \ldots, N \qquad [10.10]$$

en el que sólo observamos una medida imperfecta de la variable endógena. Es decir, aunque disponemos de información muestral acerca de x_i, sólo observamos:

$$y_i^* = y_i + v_i, \qquad i = 1, 2, \ldots, N \qquad [10.11]$$

donde suponemos que v_i se distribuye $N(0, \sigma_v^2)$, independiente de u_i y x_i.

Incorporando esta variable al modelo tenemos:

$$y_i^* = \beta x_i + (u_i + v_i) = \beta x_i + \varepsilon_i \qquad [10.12]$$

que, bajo los supuestos mencionados, puede analizarse por los procedimientos habituales, sin que suponga ninguna dificultad. De hecho, el coeficiente β es el mismo en el modelo teórico [10.10] que en el modelo con las variables observables [10.12].

En consecuencia, los errores de medida en la variable endógena no producen ningún problema importante al estimar por mínimos cuadrados, por lo que en lo sucesivo suponemos que es x_i quien se observa con error:

$$x_i^* = x_i + \omega_i, \qquad i = 1, 2, \ldots, N \qquad [10.13]$$

donde ω_i se distribuye $N(0, \sigma_\omega^2)$, independiente de u_i, de x_i y de y_i.

El modelo en las variables observables es:

$$y_i = \beta x_i^* + (u_i - \beta\omega_i) = \beta x_i^* + \varepsilon_i \qquad [10.14]$$

en el que, a diferencia de lo que ocurría en [10.12], tenemos una dificultad: el término de error compuesto ε_i está correlacionado con la variable explicativa x_i^*, ya que:

$$\begin{aligned}\mathrm{Cov}(\varepsilon_i, x_i^*) &= \mathrm{Cov}(u_i - \beta\omega_i, x_i + \omega_i) = \mathrm{Cov}(u_i, x_i) - \beta\,\mathrm{Cov}(\omega_i, x_i) + \\ &+ \mathrm{Cov}(u_i, \omega_i) - \beta\,\mathrm{Cov}(\omega_i, \omega_i) = 0 - \beta\cdot 0 + 0 - \beta\sigma_\omega^2\end{aligned}$$

donde el primer sumando es nulo por los supuestos habituales del modelo de regresión, mientras que los restantes se deben a las hipótesis efectuadas sobre el error de medida.

Esta correlación hace que el estimador MCO de [10.14] sea sesgado. Si calculamos su límite en probabilidad, tenemos:

$$plím\ \hat{\beta} = \frac{plím\ \frac{1}{N}\sum_{1}^{N} x_i^* y_i}{plím\ \frac{1}{N}\sum_{1}^{N} x_i^{*2}} = \frac{plím\ \frac{1}{N}\Sigma(x_i + \omega_i)(\beta x_i + u_i)}{plím\ \frac{1}{N}\Sigma(x_i + \omega_i)^2} =$$

$$= \beta + \frac{plím\ \frac{1}{N}\sum_{1}^{N}(x_i + \omega_i)(u_i - \beta\omega_i)}{plím\ \frac{1}{N}\sum_{1}^{N}(x_i + \omega_i)^2} = \beta + \frac{-\beta\sigma_\omega^2}{S_x^2 + \sigma_\omega^2} = \frac{\beta}{1 + \sigma_\omega^2/S_x^2}. \tag{10.15}$$

donde S_x^2 denota $S_x^2 = plím\ \frac{1}{N}\sum_{1}^{N} x_i^2$ que suponemos que existe. Para obtener [10.15] hemos utilizado nuevamente las hipótesis acerca del error de medida que antes hicimos.

El resultado general es que, *en presencia de errores de medida, el estimador MCO de β estará sesgado hacia el origen*, tanto si β es positivo como si es negativo. La magnitud del sesgo es tanto mayor cuanto mayor sea la volatilidad del error de medida. Un error de medida en x_i que fuese constante. no produciría ningún sesgo en la estimación del coeficiente β.

En el caso de un modelo de regresión múltiple, se tiene:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$
$$\mathbf{X}^* = \mathbf{X} + \boldsymbol{\omega}$$

donde permitimos que *todas* las variables explicativas se midan con error. Extendiendo los resultados anteriores, se llegaría a:

$$plím\ \frac{\mathbf{X}^{*\prime}\mathbf{X}^*}{N} = \boldsymbol{\Sigma}_{xx} + \boldsymbol{\Sigma}_{\omega\omega}$$

donde $\boldsymbol{\Sigma}_{xx}$ es el límite: $\boldsymbol{\Sigma}_{xx} = plím\ \frac{\mathbf{X}'\mathbf{X}}{N}$, supuesto existente.

Por otra parte:

$$plím\ \frac{\mathbf{X}^{*\prime}\mathbf{y}}{N} = \boldsymbol{\Sigma}_{xx}\boldsymbol{\beta}$$

y, finalmente:

$$plím\ \hat{\boldsymbol{\beta}}_{MCO} = [\boldsymbol{\Sigma}_{xx} + \boldsymbol{\Sigma}_{\omega\omega}]^{-1}\boldsymbol{\Sigma}_{xx}\boldsymbol{\beta} = \boldsymbol{\beta} - [\boldsymbol{\Sigma}_{xx} + \boldsymbol{\Sigma}_{\omega\omega}]^{-1}\boldsymbol{\Sigma}_{\omega\omega}\boldsymbol{\beta}$$

que implica que un solo error basta para generar inconsistencias en todos los coeficientes del modelo.

10.7. ESTIMACION POR VARIABLES INSTRUMENTALES

La estimación consistente en presencia de errores de medida es posible si se dispone de *instrumentos*. Como vimos en la Sección 9.4, un instrumento es una variable, no incluida en el modelo, que: *a*) esté incorrelacionada, al menos asintóticamente, con el término de error, a la vez que: *b*) está correlacionada con la variable explicativa para la que actúa de instrumento.

Volviendo al modelo [10.14], el sesgo asintótico del estimador MCO surge por la correlación entre ε_i y x_i^*. Supongamos ahora que se dispone de una variable z_i tal que:

$$\underset{N}{plím}\ \frac{1}{N}\sum_1^N z_i\varepsilon_i = 0; \qquad \underset{N}{plím}\ \frac{1}{N}\sum_1^N z_i x_i^* \neq 0$$

Entonces el estimador de variables instrumentales de [10.14] es:

$$\hat{\beta}_{VI} = \frac{\sum_1^N z_i y_i}{\sum_1^N z_i x_i^*} \qquad [10.16]$$

Sin más que utilizar [10.14], se tiene:

$$\underset{N}{plím}\ (\hat{\beta}_{VI} - \beta) = \underset{N}{plím}\ \frac{\frac{1}{N}\sum_1^N z_i\varepsilon_i}{\frac{1}{N}\sum_1^N z_i x_i^*}$$

que, bajo los supuestos anteriores, es igual a cero.

Si hay varias variables explicativas con errores de medida, debemos utilizar una variable instrumental para cada una de ellas, definiendo el estimador como:

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

de lo que [10.16] es un caso particular. Bajo los supuestos anteriores, $\hat{\boldsymbol{\beta}}_{VI}$ es consistente, con matriz de covarianzas asintótica:

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}_{VI}) = \sigma_\varepsilon^2(\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{Z})(\mathbf{X}'\mathbf{Z})^{-1}$$

que, en el caso de [10.14], se reduce a:

$$\operatorname{Var}(\hat{\beta}_{\mathrm{VI}}) = \sigma_\varepsilon^2 \cdot \frac{\sum_1^N z_i^2}{\left(\sum_1^N z_i x_i^*\right)^2}$$

El parámetro σ_ε^2 se estima dividiendo la suma residual por el número de grados de libertad, y los residuos se obtienen mediante $\hat{u}_i = y_i - \mathbf{x}_i'\hat{\boldsymbol{\beta}}_{\mathrm{VI}}$.

10.8. UN CONTRASTE DE ESPECIFICACION PARA ERRORES DE MEDIDA

Bajo errores de medida, el estimador MCO es inconsistente, mientras que el estimador de variables instrumentales es consistente. Si, en realidad, no hubiese errores de medida, ambos estimadores serían consistentes y MCO es, además, eficiente, lo que no ocurre con cualquier estimador de variables instrumentales.

Por consiguiente, para contrastar la existencia de errores de medida puede utilizarse un test tipo Hausman como el considerado en la Sección 9.4.d, comparando el valor numérico de la diferencia $(\hat{\boldsymbol{\beta}}_{\mathrm{MCO}} - \hat{\boldsymbol{\beta}}_{\mathrm{VI}})$ con su matriz de covarianzas.

Por un argumento similar al de la sección mencionada, podemos considerar el estadístico de Wald:

$$W = (\hat{\boldsymbol{\beta}}_{\mathrm{MCO}} - \hat{\boldsymbol{\beta}}_{\mathrm{VI}})' (\hat{\boldsymbol{\Sigma}}_{\mathrm{VI}} - \hat{\boldsymbol{\Sigma}}_{\mathrm{MCO}})^{-1} (\hat{\boldsymbol{\beta}}_{\mathrm{MCO}} - \hat{\boldsymbol{\beta}}_{\mathrm{VI}})$$

que se distribuye asintóticamente como una χ_k^2.

Es aconsejable utilizar en ambos casos el valor de σ_u^2 estimado a partir de los residuos de variables instrumentales. En tales condiciones, el estadístico W puede calcularse:

$$W = \frac{\mathbf{q}'\{[\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}]^{-1} - (\mathbf{X}'\mathbf{X})^{-1}\}^{-1}\mathbf{q}}{\hat{\sigma}_u^2} = \frac{\mathbf{q}'[(\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1} - (\mathbf{X}'\mathbf{X})^{-1}]^{-1}\mathbf{q}}{\hat{\sigma}_u^2}$$

donde las columnas de $\hat{\mathbf{X}}$ son los valores ajustados en la regresión de cada variable en $\mathbf{X}$ sobre el vector de variables en $\mathbf{Z}$, y $\hat{\mathbf{q}}$ es el vector diferencia: $\hat{\mathbf{q}} = \hat{\boldsymbol{\beta}}_{\mathrm{MCO}} - \hat{\boldsymbol{\beta}}_{\mathrm{VI}}$.

10.9. OBSERVACIONES INFLUYENTES

En ocasiones, existen en la muestra observaciones que pueden tener una importancia apreciable en la estimación de mínimos cuadrados. Cuando ello