Endogeneidad y heterogeneidad no observada

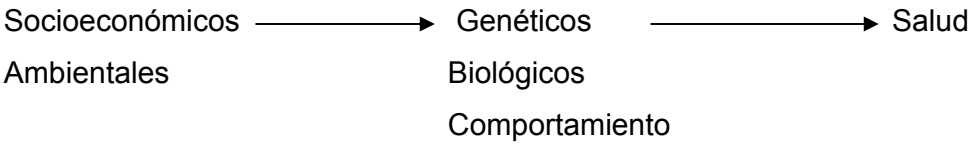
Preparado por Guido Pinto Aguirre (actualizado a Abril 2010)

Menú de hoy:

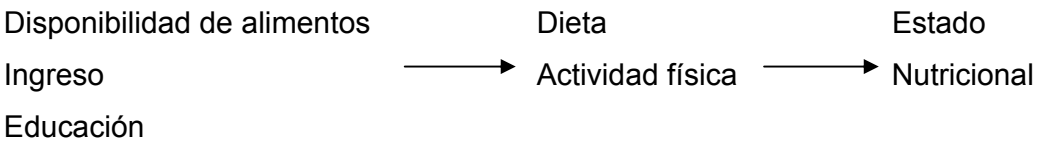
1. **Referencia**
2. **Modelo de regresión clásico y endogeneidad**
3. **Heterogeneidad no observada**
4. **Ejemplo en salud**
5. **Ecuaciones simultaneas**

6.

1. Referencia



Por **ejemplo**:



Cuando los modelos se vuelven complicados, existe la posibilidad que una variable explicativa en una ecuación sea dependiente en otra ecuación.

Estado Nutricional = f (dieta, actividad física)

Dieta= f (características individuales)

2. Modelo de regresión clásico y Endogeneidad

$$Y= \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k + \varepsilon$$

La heterogeneidad (variabilidad) entre observaciones individuales o sujetos es una condición importante para el funcionamiento apropiado de los modelos estadísticos

La variación de una característica individual dada (variable dependiente) en una población o muestra esta explicada por la heterogeneidad existente un conjunto de características observadas (variables explicativas)

Si los modelos fueran perfectos, es decir, si todas las variables posibles podrían ser incluidas en el modelo y además si podrían ser medidas sin error; entonces, las variables independientes incluidas podrían explicar toda la variación en la variable dependiente (el coeficiente de correlación múltiple seria 1)

En este caso todas las observaciones estarían sobre la línea de regresión o el hiperplano de regresión

Sin embargo, en la práctica ningún modelo es perfecto.

Variaciones puramente aleatorias (ruido blanco) en el lado de las variables explicativas requiere la incorporación de un componente denominado “término error” a fin de capturar aquella parte de la variación en la variable dependiente que explicada por la heterogeneidad de las variables explicativas

Sea el modelo:

$Y= \beta_0 + \beta_1 X + \beta_2 Z + \varepsilon$

Supuestos del modelo sobre el modelo y ε’s son los siguientes:

Lineal en los parámetros

$Y= \beta_0 + \beta_1 X + \varepsilon$

Donde ε es no observable

Muestreo aleatorio

Se utiliza una muestra aleatoria de la población bajo estudio.

$$Y_i= \beta_0 + \beta_1 X_i + \varepsilon_i$$

Donde $\{(X_i, Y_i)\ i=1,\ldots, n\}$ es una muestra aleatoria de tamaño n

Exogeneidad estricta

$$E(\varepsilon_i | x_i)= 0, \text{ para todo } i=1,.., n$$

Entonces la variable x se denomina exógena; pero si ε y x están correlacionadas, entonces x se denomina variable endógena.

De manera equivalente podemos decir que $Cov(\varepsilon_i | x_i) = 0$

Heterogeneidad de la variable independiente

En la muestra, la variable independiente X no es igual a una constante. Esto requiere alguna variación en X en la población, es decir,

$$\sum (x_i - x^*) > 0$$

Donde x* es el promedio de x

Homoscedasticidad

$Var(\varepsilon_i | X_i) = \sigma$, para todo i=1,.., n

Multicolinealidad

En la muestra (y en consecuencia en la población) no existe una relación lineal perfecta entre las variables explicativas

Correlación serial

Condicional en X, los errores de dos individuos diferentes o dos periodos diferentes no están correlacionados entre si

$Corr(\varepsilon_i, \varepsilon_j | x_i) = 0$, para todo $i \neq j$

Normalidad

El error poblacional ε es independiente de la variable explicativa x y está distribuido normalmente

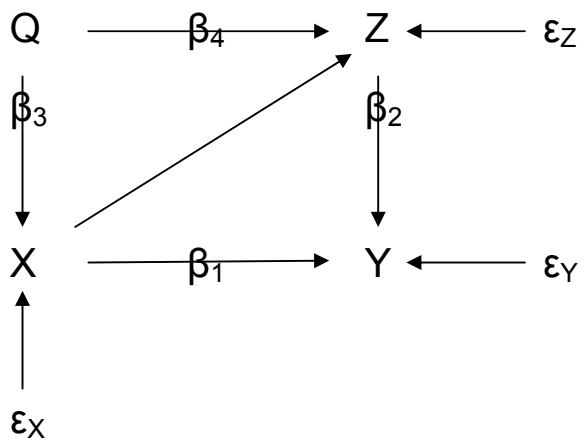
$$\varepsilon \sim N(0, \sigma^2)$$

Si estos supuestos se cumplen, entonces se dice que el modelo es apropiado para los datos disponibles

Una de los supuestos más importantes es el que asume independencia entre ε y las variables explicativas

Pero muchos modelos son más complicados que el presentado arriba. Por ejemplo, podemos considerar el caso en que X & Z están conjuntamente determinadas por otra variable, denominada Q

Esta situación se puede expresar como:



La variable Y está determinada por el efecto combinado de X & Z, las cuales están a la vez determinadas por Q

Es decir, Q no es parte explicita del modelo:

$$Y= \beta_1 X + \beta_2 Z + \varepsilon_Y$$

En consecuencia, la relación implícita puede ser especificada mediante tres ecuaciones:

$$Y= \beta_1 X + \beta_2 Z + \varepsilon_Y$$

$$X= \beta_3 Q + \varepsilon_X$$

$$Z= \beta_4 Q + \varepsilon_Z$$

La variable Z se denomina variable de confusión en la relación X –Y, es decir, parece un factor explicativo de la variable Y correlacionado con la variable X pero que **no es** parte la relación causal entre X & Y

En consecuencia, Z tiene que ser controlada de alguna forma, puesto que ignora su efecto (al dejarla fuera de la ecuación) produciría un estimado sesgado del efecto que tiene X sobre Y

De manera similar, ignorar Q, el cual es determinante de X & Z, conduce a una situación similar a la anterior (confusora)

Se pueden identificar dos tipos de variables: Q es una variable exógena (sus determinantes no son parte del modelo), mientras que X & Z son variables endógenas (están determinadas por otras variables en el modelo de 3 ecuaciones)

En la primera ecuación frecuentemente asumimos que el termino error ε es puramente aleatorio, es decir, no está correlacionado con X & Z

3. Heterogeneidad no observada

Supongamos ahora que la variable dependiente Y está afectada no sólo por X, Z y variaciones puramente aleatorias sino también por ciertos factores no aleatorios específicos a cada individuo, los cuales no son observables o medibles, tales como factores genéticos de predisposición

En este caso el modelo puede escribirse como:

$$Y= \beta_1 X + \beta_2 Z + \mu H + v_Y$$

Donde v_Y es un residual verdaderamente aleatorio y H contiene las características no observadas para cada individuo y μ es el parámetro que explica la relación entre Y y H. Además

$$\varepsilon_Y= \mu H + v_Y$$

La variable H representa las diferencias **no aleatorias** que existen entre observaciones o individuos pero que no pueden medirse directamente

El efecto de no poder incluir explícitamente H en el modelo depende del supuesto que se haga sobre la relación

que puede existir entre H y las otras variables explicativas del modelo

Si se asume que

$Cov\ (H, X)=0,\ \ Cov\ (H, Z)= 0$

Entonces dejar fuera de la ecuación el término μH conducirá a un modelo subespecificado con un residual más grande.

Sin embargo, no tendrá ningún efecto sobre el tamaño de los coeficientes sino que producirá errores estándar más grandes de lo que deberían ser

Pero si

$Cov\ (H, X) \neq 0,\ \ Cov\ (H, Z) \neq 0$

Entonces, la heterogeneidad no observada estará correlacionad con las variables explicativas, actuando como variable confusora

Consideremos el siguiente sistema de ecuaciones:

$$Y= \alpha_1 X + \alpha_2 Z + \mu_1 H_Y + \varepsilon_Y$$

$$Z= \beta_1 V + \beta_2 W + \mu_2 H_Z + \varepsilon_Z$$

La variable endógena Z es en parte una función de características no observables H_Z, que se encuentra correlacionada con H_Y, es decir, Cov (H_Z, H_Y) ≠ 0

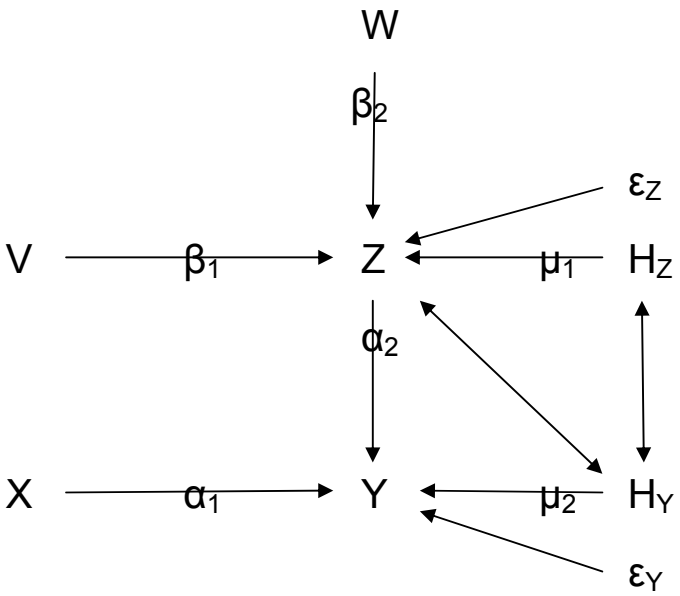
En consecuencia, Z esta correlacionada con H_Y. Entonces, la variable H_Y (aunque no observable) actúa a través del error

$$\varepsilon'_Y = \mu_1 H_Y + \varepsilon_Y$$

y es un factor de riesgo para la variable Y, al mismo tiempo que esta correlacionada con otra variable explicativa Z.

Entonces, H_Y opera como confusora de la relación Z – Y.

La representación gráfica es la siguiente manera:



En este modelo Y & Z son variable endógenas, mientras que X, V & W son exógenas

Cada ecuación que describe el comportamiento de Y & Z tiene asociadas un error puramente aleatorio (ε) y heterogeneidad no observada (H)

Se observa que en lugar de controlar directamente por la heterogeneidad no observada, existe una manera de prevenir que “actué” como confusora

Sabemos que H_Y es una confusora de la relación Z – Y debido a que la correlación con Z no es cero pero que al mismo tiempo es un factor de riesgo para Y

En consecuencia, **si** la asociación entre H_Y y Z puede ser eliminada, entonces dejara de ser confusora

En econometría existen métodos para remover de estos modelos variables confusoras: encontrar una variable denominada instrumento

Una **variable instrumental** debe estar altamente, aunque no completamente, correlacionada con la variable endógena pero no correlacionada con el término error

El instrumento es entonces utilizado para sustituir la variable endógena original

En el sistema anterior, podemos que ambas V y W son posibles candidatas para un instrumento de Z

Cuando se tienen datos longitudinales, una posible solución es utiliza el rezago de la variable endógena como instrumento (variable endógena del periodo anterior)

En econometría uno de los mejores instrumentos es el valor predicho de la variable endógena como función de todas las variables exógenas en el sistema.

Esto se conoce como estimación de la **forma reducida**. Este estimado es entonces utilizado como el instrumento para la variable endógena en la ecuación original

En el caso anterior, se debe sustituir el valor predicho de Z (que ahora no estará correlacionado con H_Y) en lugar del valor medido de Z (que esta correlacionado con H_Y).

Este método se denomina Mínimos Cuadrados dos Etapas

4. Ejemplo

El efecto de la lactancia en el regreso de la fecundabilidad puede ser representado como:

$$M_{ti} = \alpha_1 + \alpha_2 BF_{ti} + \alpha_3 X_{ti} + \mu_{(M)i} + \varepsilon_{(M)ti}$$

M: probabilidad de retorno de la fecundabilidad

BF: medidas de lactancia (frecuencia, intensidad, etc.)

X: otras variables explicativas como paridad, peso, energía, ingesta de alimentos, gasto de energía, actividad sexual, etc.

$\mu_{(M)i}$: término error especifico a cada individuo (i) que no cambia con el tiempo (heterogeneidad no observada)

$\varepsilon_{(M)ti}$: erro puramente aleatorio

En esta ecuación μ puede pensarse como una variable omitida y una fuente potencia de confusión

La relación de comportamiento de la lactancia puede escribirse como:

$$BF_{ti} = \beta_1 + \beta_2 M_{ti} + \beta_3 Y_{ti} + \mu_{(BF)i} + \varepsilon_{(BF)ti}$$

Y: incluye factores como edad de la madre, paridad, anticoncepción

En la primera ecuación, BF y μ pueden estar correlacionadas, que convierte a BF en una variable endógena.

Esta es una fuente de sesgo y también puede existir una situación de confusión

Corrección por Endogeneidad

Una técnica estándar para remover la correlación que existe [Cov (x, ε) ≠ 0] es el uso de una variable instrumental.

Cada variable endógena es reemplazada por una variable sustituta, una variable instrumental (VI)

Esta VI debe estar correlacionado con la variable endógena pero no con el termino error

Para crear una VI se debe desarrollar una ecuación de “forma reducida”, la cual predice los valores de la variable endógena a partir de un conjunto de variables estrictamente exógenas

5. Ecuaciones simultáneas

Naturaleza

Los modelos uniecuacionales desafortunadamente ignoran la interdependencia que existe entre las variables en el mundo real. Los modelos más importantes en las ciencias sociales son de naturaleza simultánea

El ejemplo más clásico de simultaneidad en economía es la oferta y demanda de bienes y servicios. Estudiar la demanda de carne sin analizar la oferta de carne es tomar riesgos de omitir importantes relaciones e incurrir en errores significativos

La preocupación de la econometría por las ecuaciones simultáneas aparece en el momento de su estimación utilizando MCO. Si este es el caso, se producen una serie de dificultades que no se encuentran con modelos uniecuacionales.

En las ecuaciones simultáneas principalmente no se cumple el supuesto clásico sobre la no correlación del término error con todas las variables explicativas. Es decir,

$$\text{cov}\ (\varepsilon_i, X_i) = 0$$

Debido a este incumplimiento, los estimados MCO de los coeficientes son ***sesgados*** en los modelos simultáneos. El procedimiento de estimación denominado Mínimos Cuadrados en Dos Etapas (MC2E) se utiliza en tales modelos en lugar de MCO

Ecuaciones estructurales

¿Qué viene primero, la oferta o la demanda? Esta pregunta es difícil de responder en forma satisfactoria porque oferta y demanda están determinadas conjuntamente; existe una relación causal de ida y vuelta entre las variables.

El mundo económico está lleno de este tipo de efectos de retroalimentación y causalidad en dos direcciones que requieren la aplicación de ecuaciones simultaneas. Además de la oferta y demanda y modelos de macroeconómicos simples, se puede mencionar la causalidad bidireccional que existe entre el tamaño de la población y oferta de alimentos, la determinación conjunta de los salarios y precios

Una ecuación simple econométrica típica se escribe como:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t$$

En cambio, un sistema simultáneo es aquel en el cual Y claramente tiene un efecto sobre al menos una de las Xs en adición al efecto que las Xs tienen sobre Y.

Estas relaciones se modelan al distinguir entre aquellas variables que se determinadas simultáneamente tales como las Ys, que se denominan ***variables endógenas***, y aquellas que no se determinan simultáneamente como las Xs, que se denominan ***variables exógenas***.

En el siguiente sistema:

$$Y_{1t} = \alpha_0 + \alpha_1 Y_{2t} + \alpha_2 X_{1t} + \alpha_3 X_{2t} + \varepsilon_{1t}$$

$$Y_{2t} = \beta_0 + \beta_1 Y_{1t} + \beta_2 X_{3t} + \beta_3 X_{2t} + \varepsilon_{2t}$$

La primera ecuación caracteriza el comportamiento de los consumidores y la segunda ecuación muestra el comportamiento de los oferentes de un producto

Estas relaciones de comportamiento se denominan ***ecuaciones estructurales***, caracterizan la teoría económica que existe detrás de cada variable endógena al expresarlas en términos de variables endógenas y exógenas

Los econometristas deben ver las ecuaciones estructurales como un sistema a fin de percibir las retroalimentaciones involucradas. Por ejemplo, el hecho que las Ys están determinadas conjuntamente significa que un

cambio en Y_1 causará un cambio en Y_2, lo que a su vez ocasionara que Y_1 cambie nuevamente

Se puede comparar este mecanismo de retroalimentación con un cambio en X_1, el cual no causará retroalimentación para que X_1 cambie nuevamente

Los parámetros α y β en las ecuaciones estructurales se denominan ***coeficientes estructurales***; las hipótesis deben realizarse sobre sus signos de la misma manera que se hacen sobre los coeficientes de regresión de los modelos uniecuacionales

Es importante aclarar que una variable es endógena porque está determinada conjuntamente y no sólo porque aparece en ambas ecuaciones. Por ejemplo, X_2 está en ambas ecuaciones pero es de naturaleza exógena porque no está determinada simultáneamente en el mercado de algún producto

¿Cómo se decide si una variable en particular debería ser endógena o exógena? Algunas variables son casi siempre exógenas (como el tiempo) pero muchas otras pueden ser consideradas endógenas o exógenas, dependiendo del número y características de las otras ecuaciones en el sistema

En consecuencia, la distinción entre variables endógenas y exógenas depende de cómo el investigador define el enfoque de la investigación

Muchas veces, variables endógenas rezagadas aparecen en sistemas simultáneos, especialmente cuando las ecuaciones involucradas son ecuaciones de rezagos distribuidos

A fin de evitar la confusión que puede existir cuando se tienen variables rezagadas, se definen las **variables predeterminadas** para incluir todas las variables exógenas y endógenas rezagadas

Predeterminadas significa que las variables exógenas y endógenas rezagadas son determinadas fuera del sistema de ecuaciones especificadas o en un momento anterior al periodo actual.

Variables endógenas que no están rezagadas no son predeterminadas porque son conjuntamente determinadas por el sistema en el periodo de tiempo actual

En consecuencia, los econometristas tienden a hablar en términos de variables endógenas y predeterminadas cuando se discuten sistemas de ecuaciones simultáneas

La especificación de un modelo simple de oferta y demanda puede realizarse en los siguientes términos:

$$Q_{Dt} = \alpha_0 + \alpha_1 P_t + \alpha_2 X_{1t} + \alpha_3 X_{2t} + \varepsilon_{Dt}$$

$$Q_{St} = \beta_0 + \beta_1 P_t + \beta_2 X_{3t} + \varepsilon_{St}$$

$$Q_{Dt} = Q_{St} \text{ (condición de equilibrio)}$$

Donde:

Q_{Dt}: cantidad de la bebida gaseosa demanda en el período t

Q_{St}: cantidad de la bebida gaseosa ofertada en el período de tiempo t

P_t: precio de la bebida gaseosa en el periodo de tiempo t

X_{1t}: dólares en publicidad para la gaseosa en el periodo de tiempo t

X_{2t}: variable exógena del lado de la demanda (ingreso, precio de otras bebidas gaseosas, etc.)

X_{3t}: variable exógena del lado de la oferta (precio de sabores artificiales, u otros factores de producción)

ε_t: término error clásico, cada ecuación tiene su propio error

En este caso, el precio y la cantidad están determinados simultáneamente pero una de las variables endógenas no se encuentra en el lado izquierdo de ninguna de las ecuaciones

Por lo tanto, es incorrecto asumir automáticamente que las variables endógenas son sólo aquellas que aparecen en el lado izquierdo de al menos una de las ecuaciones

En este caso particular, podríamos haber escrito la segunda ecuación con el precio en el lado izquierdo y la cantidad ofertada en el lado derecho, como se hizo en las ecuaciones de la carne de res. Mientras que los coeficientes estimados serían diferentes, las relaciones subyacentes no serían diferentes

Debe haber tantas ecuaciones como variables endógenas. En el ejemplo anterior, las variables endógenas son Q_D, Q_S y P

¿Cuáles serían los signos esperados de los coeficientes del precio en las ecuaciones de demanda y oferta? Se esperaría que sea negativo en la ecuación de demanda y positivo en la de oferta; es decir, $\alpha_1 < 0$ y $\beta_1 > 0$. En efecto, la teoría económica dice que cuanto más alto es el precio mayor será la cantidad demandada pero mayor será la cantidad ofertada.

¿Qué pasaría si por accidente colocamos una variable predeterminada del lado de la oferta en la ecuación de la demanda? Se producirían dificultades en identificar cuál ecuación es de demanda y cuál es de oferta, y los signos

esperados de los coeficientes de la variable endógena P serían ambiguos

Por lo tanto, se debe tener cuidado cuando se especifican las ecuaciones estructurales de un sistema

Incumplimiento del supuesto clásico cov (X_i, ε_i)= 0

Uno de los supuestos clásicos requiere que el término error y cada una de las variables explicativas (independientes) en el modelo no deben estar correlacionados

Si existe tal correlación, entonces el método de estimación de MCO es probable que atribuya a una variable explicativa en particular variaciones en la variable dependiente que en realidad están siendo causadas por variaciones en el término error El resultado serán *estimados sesgados*

Para visualizar cómo las ecuaciones simultáneas incumplen el supuesto de independencia entre el término error y las variables explicativas, se marcan las variables involucradas en el sistema presentado al inicio de la exposición. ¿Qué sucede cuando uno de lo términos de error aumenta, manteniendo constante todo los demás en la ecuación?

$$Y_{1t} = \alpha_0 + \alpha_1 Y_{2t} + \alpha_2 X_{1t} + \alpha_3 X_{2t} + \varepsilon_{1t}$$

$$Y_{2t} = \beta_0 + \beta_1 Y_{1t} + \beta_2 X_{3t} + \beta_3 X_{2t} + \varepsilon_{2t}$$

Si ε_1 aumenta en un período de tiempo, entonces Y_1 también aumentará (primera ecuación). En consecuencia, si Y_1 aumenta, Y_2 también aumentará (segunda ecuación), asumiendo que $\beta_1>0$. Pero si Y_2 aumenta en la segunda ecuación, también aumentará en la primera ecuación donde es una variable explicativa

En consecuencia, un aumento en el término error de una ecuación ocasiona un incremento en una variable explicativa en la misma ecuación: si ε_1 aumenta, Y_1 aumenta y entonces Y_2 aumenta, incumpliéndose el supuesto de independencia entre el término error y las variables explicativas

Por lo tanto, todo lo que se requiere para el incumplimiento del supuesto de independencia mencionado es que las variables endógenas estén determinadas conjuntamente en un sistema de ecuaciones simultáneas

Ecuaciones de forma reducida

Una forma alternativa de expresar un sistema de ecuaciones simultáneas es a través del uso de las ***ecuaciones de forma reducida***, las cuales expresan una variable endógena particular sólo en términos del error y todas las variables predeterminadas (exógenas y endógenas rezagadas) en el sistema simultáneo

Las ecuaciones en forma reducida de las ecuaciones estructurales anteriores serian las siguientes:

$$Y_{1t} = \pi_0 + \pi_1 X_{1t} + \pi_2 X_{2t} + \pi_3 X_{3t} + v_{1t}$$

$$Y_{2t} = \pi_4 + \pi_5 X_{1t} + \pi_6 X_{2t} + \pi_7 X_{3t} + v_{2t}$$

Donde las v_s son los términos de error estocásticos y las π_s son los ***coeficientes de la forma reducida*** debido a que son los coeficientes de las variables predeterminadas en las ecuaciones en forma reducida

Ahora cada ecuación contiene sólo una variable endógena, como variable dependiente, y el mismo número de de variables predeterminadas

Los coeficientes de la forma reducida (π_s) son conocidos como ***multiplicadores de impacto*** porque miden el impacto en la variable endógena del incremento en una unidad en el valor de la variable predeterminada, después de controlar por los efectos de retroalimentación en todo el sistema simultáneo

En el ejemplo de la oferta y demanda de las bebidas gaseosas la especificación de las ecuaciones de forma reducida para el modelo se traduce en sólo dos ecuaciones

puesto que la condición de equilibrio (tercera ecuación) obliga a que la cantidad demandada (Q_D) sea igual a la cantidad ofertada (Q_S):

$$P_t = \pi_0 + \pi_1 X_{1t} + \pi_2 X_{2t} + \pi_3 X_{3t} + v_{1t}$$

$$Q_t = \pi_4 + \pi_5 X_{1t} + \pi_6 X_{2t} + \pi_7 X_{3t} + v_{2t}$$

Aunque el precio (P) nunca aparece en el lado izquierdo de ninguna ecuación estructural, es una variable endógena y debe ser tratada como tal

Existen al menos 4 razones para utilizar las ecuaciones de forma reducida:

Primero, puesto que las ecuaciones de forma reducida no tienen simultaneidad inherente, ahora cumplen el supuesto de independencia entre el término error y las variables independientes. En consecuencia, pueden ser estimadas mediante MCO sin encontrar los problemas discutidos anteriormente.

Segundo, Los coeficientes de la forma reducida estimados de esta manera pueden ser utilizados para obtener los coeficientes estructurales estimados. Es decir, las ecuaciones estimadas de forma reducida se pueden utilizar para obtener los coeficientes estimados α_s y β_s de las ecuaciones estructurales.

Este método de calcular los coeficientes estimados estructurales a partir de los coeficientes en forma reducida se denomina Mínimos Cuadrados Indirectos (MCI). Pero MCI es útil en situaciones muy limitadas.

Tercero, la interpretación de los coeficientes en forma reducida como multiplicadores de impacto significa que tienen significado económico y una aplicación útil. Por ejemplo, si se quiere comparar un aumento en los gastos de gobierno con una reducción de impuestos en términos del impacto por dólar en el primer año, estimados de los multiplicadores (π_s) permitirían tal comparación

Cuarto, las ecuaciones de forma reducida tienen un rol crucial en la técnica de estimación más frecuentemente utilizada para ecuaciones simultáneas. Esta técnica se denomina Mínimos Cuadrados en dos Etapas (MC2E)

Obtención de los coeficientes de la forma reducida

Sean las ecuaciones incluidas en el modelo de las bebidas gaseosas:

$$Q_{Dt} = \alpha_0 + \alpha_1 P_t + \alpha_2 X_{1t} + \alpha_3 X_{2t} + \varepsilon_{Dt}$$

$$Q_{St} = \beta_0 + \beta_1 P_t + \beta_2 X_{3t} + \varepsilon_{St}$$

$$Q_{Dt} = Q_{St} \text{ (condición de equilibrio)}$$

Se puede utilizar la condición de equilibrio para igualar las ecuaciones anteriores y despejar el precio:

$$\alpha_0 + \alpha_1 P_t + \alpha_2 X_{1t} + \alpha_3 X_{2t} + \varepsilon_{Dt} = \beta_0 + \beta_1 P_t + \beta_2 X_{3t} + \varepsilon_{St}$$

$$\alpha_1 P_t - \beta_1 P_t + \alpha_2 X_{1t} + \alpha_3 X_{2t} + \varepsilon_{Dt} = \alpha_0 + \beta_0 + \beta_2 X_{3t} + \varepsilon_{St}$$

$$(\alpha_1 - \beta_1) P_t = - \alpha_2 X_{1t} - \alpha_3 X_{2t} - \varepsilon_{Dt} + \alpha_0 + \beta_0 + \beta_2 X_{3t} + \varepsilon_{St}$$

$$(\alpha_1 - \beta_1) P_t = (\alpha_0 + \beta_0) - \alpha_2 X_{1t} - \alpha_3 X_{2t} + \beta_2 X_{3t} + (\varepsilon_{St} - \varepsilon_{Dt})$$

$$P_t = (\alpha_0 + \beta_0)/ (\alpha_1 - \beta_1) - \alpha_2/(\alpha_1 - \beta_1) X_{1t} - \alpha_3/(\alpha_1 - \beta_1) X_{2t} + \beta_2/(\alpha_1 - \beta_1) X_{3t} + (\varepsilon_{St} - \varepsilon_{Dt})/(\alpha_1 - \beta_1)$$

La anterior expresión puede escribir:

$$P_t = \pi_0 + \pi_1 X_{1t} + \pi_2 X_{2t} + \pi_3 X_{3t} + v_{1t}$$

Reemplazando P_t en la ecuación de oferta o demanda se tiene que:

$$Q_t = \beta_0 + \beta_1 (\pi_0 + \pi_1 X_{1t} + \pi_2 X_{2t} + \pi_3 X_{3t} + v_{1t}) + \beta_2 X_{3t} + \varepsilon_{St}$$

$$Q_t = \beta_0 + \beta_1 \pi_0 + \beta_1 \pi_1 X_{1t} + \beta_1 \pi_2 X_{2t} + \beta_1 \pi_3 X_{3t} + \beta_1 v_{1t} + \beta_2 X_{3t} + \varepsilon_{St}$$

$$Q_t = \beta_0 + \beta_1 \pi_0 + \beta_1 \pi_1 X_{1t} + \beta_1 \pi_2 X_{2t} + \beta_1 \pi_3 X_{3t} + \beta_1 v_{1t} + \beta_2 X_{3t} + \varepsilon_{St}$$

$$Q_t = (\beta_0 + \beta_1 \pi_0) + \beta_1 \pi_1 X_{1t} + \beta_1 \pi_2 X_{2t} + (\beta_1 \pi_3 + \beta_2) X_{3t} + (\beta_1 v_{1t} + \varepsilon_{St})$$

$$Q_t = \pi_4 + \pi_5 X_{1t} + \pi_6 X_{2t} + \pi_7 X_{3t} + v_{2t}$$

En consecuencia, las ecuaciones de forma reducida son:

$$P_t = \pi_0 + \pi_1 X_{1t} + \pi_2 X_{2t} + \pi_3 X_{3t} + v_{1t}$$

$$Q_t = \pi_4 + \pi_5 X_{1t} + \pi_6 X_{2t} + \pi_7 X_{3t} + v_{2t}$$

Sesgo de simultaneidad producido por MCO

Todos los supuestos clásicos deben cumplirse para que los estimados MCO sean MELI. Cuando se incumple uno de los supuestos, se debe identificar cual de las propiedades no se mantiene

Cuando se aplica MCO directamente a las ecuaciones estructurales de un sistema simultáneo (denominado ahora Mínimos Cuadrados Directos- MCD) produce estimados de los coeficientes sesgados. Este sesgo se llama sesgo de ecuaciones simultáneas o ***sesgo de simultaneidad***

El sesgo de simultaneidad se refiere al hecho que en un sistema simultáneo, los valores esperados de los coeficientes estructurales estimados (b_s) mediante MCO no son iguales a los verdaderos β_s

Estos coeficientes estimados son también ***inconsistentes***. Es decir, los valores esperados de las b_s no

se aproximan a los verdaderos β_s aún cuando el tamaño de la muestra es bastante grande

Por lo tanto, se enfrenta el problema que en un sistema simultáneo:

$$E(b) \neq \beta$$

¿Por qué existe este sesgo de simultaneidad? Se debe recordar que en un sistema de ecuaciones simultáneas, los términos de error (ε_s) tienden a estar correlacionados con las variables endógenas (Y_s) cada vez que las Y_s aparecen como variables explicativas

Se puede seguir el significado de esta correlación (asumiendo coeficientes positivos) a través del siguiente sistema de ecuaciones estructurales:

$$Y_{1t} = \beta_0 + \beta_1 Y_{2t} + \beta_2 X_{1t} + \varepsilon_{1t}$$

$$Y_{2t} = \alpha_0 + \alpha_1 Y_{1t} + \alpha_2 Z_t + \varepsilon_{2t}$$

Puesto que no se puede observar el término error (ε_1) y de desconoce cuando ε_{1t} se encuentra encima del promedio, aparecerá como si cada vez que Y_1 esta encima del promedio, Y_2 también.

Como resultado, el procedimiento de estimación de MCO tenderá a atribuir a Y_2 incrementos en Y_1 causados en realidad por el término error ε_1, en consecuencia se sobreestimará β_1. ***Esta sobreestimación es sesgo de simultaneidad***

Si el término error es anormalmente negativo, Y_{1t} es menor de lo que hubiese sido de otra manera, causando que Y_{2t} sea menor de lo que hubiese sido, y el procedimiento de MCO atribuirá el decremento en Y_1 a Y_2, nuevamente causando la sobreestimación de β_1 (induciendo a un sesgo hacia arriba)

Se debe recordar que la causalidad entre Y_1 y Y_2 va en ambas direcciones porque las dos variables son interdependientes

Como resultado, cuando β_1 es estimado por MCO no puede interpretarse como el impacto de Y_2 en Y_1, manteniendo constante X. En lugar, ahora b_1 mide una mezcla de los efectos de las dos variables endógenas, una sobre la otra

Adicionalmente, β_2 se supone que es el efecto de X sobre Y_1, manteniendo constante Y_2. ¿Cómo se puede mantener constante Y_2 cuando se produce un cambio en Y_1? Como resultado existe un sesgo potencial en todos los coeficientes estimados en un sistema simultáneo

El incumplimiento del supuesto clásico: cov (ε_i, X_i)= 0, siempre se traducirá en un sesgo en la estimación de β_1

Mínimos Cuadrados en dos Etapas (MC2E)

Existe una variedad de técnicas econométricas de estimación disponibles que ayudan a resolver el sesgo y evitar la inconsistencia inherente en la aplicación de MCO a las ecuaciones simultáneas; sin embargo, la alternativa a MCO más utilizada se denomina Mínimos Cuadrados en 2 Etapas (MC2E)

MCO produce sesgos en la estimación de ecuaciones simultaneas porque tales ecuaciones no cumplen el supuesto **cov (X_i, ε_i)= 0**, de modo que una solución al problema es explorar maneras de evitar el incumplimiento del supuesto

Se podría hacer esto si se encuentra una variable que: (1) es una buena proxy de la variable endógena, y (2) no está correlacionada con el término error

Si se substituye esta nueva variable por la variable endógena que aparece como variable explicativa, la nueva variable no estará correlacionada con el término error y se cumplirá el supuesto cov (X_i, ε_i)= 0

Sea el siguiente sistema:

$$Y_{1t} = \beta_0 + \beta_1 Y_{2t} + \beta_2 X_t + \varepsilon_{1t}$$

$$Y_{2t} = \alpha_0 + \alpha_1 Y_{1t} + \alpha_2 Z_t + \varepsilon_{2t}$$

Si se puede encontrar una variable altamente correlacionada con Y_2 pero que no esté correlacionada con ε_1, entonces podríamos substituir esta nueva variable por Y_2 en el lado derecho de la primera ecuación en el sistema anterior, por lo tanto se cumpliría el supuesto mencionado más arriba

Esta nueva variable se llama **variable instrumental**. Una variable instrumental reemplaza una variable endógena (cuando es una variable explicativa). Entonces, es una buena proxy para la variable endógena y es independiente del término error

Puesto que no existe una causalidad conjunta entre la variable instrumental y cualquier variable endógena, el uso de la variable instrumental evita el incumplimiento del supuesto cov $(X_{i,} \varepsilon_i)= 0$

¿Cómo se construye o encuentra una variable con estas características? MC2E proporciona una respuesta aproximada

MC2E es un método de creación sistemática de variables instrumentales para reemplazar las variables endógenas donde aparecen como variables explicativas en sistemas de ecuaciones simultáneas

El procedimiento tiene dos etapas:

Primera etapa: *correr MCO en las ecuaciones de forma reducidas para cada una de las variables endógenas que aparecen como variables explicativas en el sistema de ecuaciones estructurales*

Puesto que las variables predeterminadas (exógenas y endógenas rezagadas) no están correlacionadas con el error de la forma reducida, los estimados de MCO de los coeficientes de la forma reducida (p_s) son insesgados

Estos p_s pueden utilizarse para calcular estimados de las variables endógenas. Debido a que las π_s no están correlacionadas con las ε_s, este procedimiento sólo produce variables instrumentales aproximadas que proporciona estimados de los coeficientes de las ecuaciones estructurales (β_s) que son consistentes (para muestras grandes) pero sesgados (para pequeñas muestras)

$$\hat{Y}_{1t} = p_0 + p_1 X_t + p_2 Z_t$$

$$\hat{Y}_{2t} = p_3 + p_4 X_t + p_5 Z_t$$

Las $\hat{Y}_s$ son utilizadas como variables substitutas (proxies) en las ecuaciones estructurales del sistema simultáneo

Segunda etapa: substituir los $\hat{Y}_s$ de la forma reducida en lugar de los Y_s que aparecen en sólo en el lado derecho de las ecuaciones estructurales y luego estimar estar ecuaciones estructurales revisadas mediante MCO

Esto es, la segunda etapa consiste en estimar las siguientes ecuaciones con MCO:

$$Y_{1t} = \beta_0 + \beta_1 \hat{Y}_{2t} + \beta_2 X_t + \varepsilon_{1t}$$

$$Y_{2t} = \alpha_0 + \alpha_1 \hat{Y}_{1t} + \alpha_2 Z_t + \varepsilon_{2t}$$

Debe notarse que las variables dependientes siguen siendo las variables endógenas originales, sin embargo las substituciones se realizan solamente para las variables endógenas que aparecen en el lado derecho de las ecuaciones estructurales

Si las ecuaciones del segundo estado anteriores son estimadas con MCO, los errores estándar resultantes serán incorrectos, así que se debe utilizar el procedimiento de estimación de MC2E existente en los paquetes econométricos

Propiedades de los MC2E

Este procedimiento tiene varias propiedades:

Primero, los estimados de MC2E son todavía sesgados pero ahora son consistentes. Es decir, para muestras pequeñas el valor esperado de b producido por MC2E no es igual al verdadero β. Pero para muestras cada vez más grandes, el valor esperado de b se aproxima al verdadero β

A medida que el tamaño muestral crece, la varianza de los estimados MCO y MC2E decrecen. Estimados de MCO se vuelven precisos pero de los números equivocados, mientras que los estimados MC2E se vuelven precisos de los números correctos

En consecuencia, a medida que aumenta el tamaño de la muestra, la técnica de MC2E es mejor para la estimación de las ecuaciones simultáneas

Segundo, el sesgo producido por MC2E para pequeñas muestras es de signo opuesto al sesgo producido por MCO

Se debe recordar que el sesgo en MCO es positivo, indicando que un b producido por MCO para un sistema simultáneo es probable que se más grande que el verdadero β

Para MC2E, el sesgo esperado es negativo y, por lo tanto, es probable que el valor de b estimado, producido por MC2E, sea menor que el verdadero β

Los estimados por MC2E, para un conjunto de datos dados, pueden ser más grandes que aquellos producidos por MCO, pero se puede mostrar que es probable que la mayoría de los estimados por MC2E sean menores que los correspondientes estimados por MCO. Para grandes muestras, el sesgo en MC2E es pequeño

Tercero, si el ajuste de la ecuación de forma reducida es bastante pobre, entonces MC2E no funcionará muy bien

Se debe recordar que se supone que variable instrumental es una muy buena aproximación (proxy) para la variable endógena. Puesto que el ajuste (medido por R^2) de la ecuación de la forma de reducida es pobre, entonces la variable instrumental ya no está altamente correlacionada con la variable endógena original y no hay razón para esperar que MC2E sea efectivo

Como el R^2 de la ecuación de la forma reducida aumenta, la utilidad de MC2E también aumentará

Cuarto, si las variables predeterminadas están altamente correlacionadas, entonces MC2E no funcionará bien

La primera etapa de MC2E incluye variables explicativas de ecuaciones estructurales diferentes en la misma forma de

ecuación reducida. Como resultado, es posible la existencia de multicolinealidad severa entre variables explicativas en las ecuaciones de forma reducida provenientes de diferentes ecuaciones estructurales

Cuando esto sucede, un $\hat{Y}$ producido por una ecuación de forma reducida puede estar altamente correlacionado con las variables exógenas en la ecuación estructural; en consecuencia, la segunda etapa de MC2E también mostrará un alto grado de multicolinealidad y las varianzas de los coeficientes estimados serán altas

Así, cuanto más alta sea el coeficiente de correlación simple entre las variables predeterminadas (o cuando más altos sean los factores de inflación de la varianza) menos precisos serán los estimados provenientes de MC2E

Quinto, el uso de la prueba t para contrastar hipótesis es más preciso utilizando estimadores MC2E que utilizando estimadores MCO

A pesar que la prueba t no es exacta para los estimadores MC2E, es suficientemente precisa bajo muchas circunstancias. En contraste, el sesgamiento de los estimadores MCO en sistemas simultáneos implica que sus estadísticos t no son los suficientemente precisos como para contar con ellos para propósitos de contraste de hipótesis

Esto significa que puede ser apropiado el uso de MC2E aún cuando las variables predeterminadas estén altamente correlacionadas

En resumen, MC2E siempre producirá mejores estimadores de los coeficientes de un sistema que aquellos provenientes de MCO. La mayor excepción a esta regla general se origina cuando el ajuste de la ecuación de la forma reducida en cuestión es bastante pobre para una muestra pequeña

Ejemplo de Mínimos Cuadrados en 2 Etapas

El modelo Keynesiano simple de una economía es especificado mediante el siguiente sistema:

$$Y_t = C_t + I_t + G_t + XN_t$$

$$C_t = \beta_0 + \beta_1 YD_t + \beta_2 C_{t-1} + \varepsilon_{1t}$$

$$I_t = \beta_3 + \beta_4 Y_t + \beta_5 r_{t-1} + \varepsilon_{2t}$$

$$r_t = \beta_6 + \beta_7 Y_t + \beta_8 M_t + \varepsilon_{3t}$$

$$YD_t = Y_t - T_t$$

Donde:

Y_t = Producto Interno Bruto (PIB) en el año t

C_t = Consumo personal total en el año t

I_t = Inversión doméstica privada bruta total en el año t

G_t = Compras gubernamentales de bienes y servicios en el año t

XN_t = Exportaciones netas de bienes y servicios (exportaciones menos importaciones) en el año t

T_t = Impuestos en el año t

r_t = tasa de interés (producida en el ámbito comercial) en el año t

M_t = Oferta monetaria en el año t

YD_t = Ingreso disponible en el año t

Todas las variables están en términos reales (medidas en billones de dólares de 1987) excepto la tasa de interés, la cual está medida en términos nominales (porcentaje). Los datos van de 1964 a 1994

Las ecuaciones anteriores son ecuaciones estructurales del sistema, pero sólo las ecuaciones dos, tres y cuatro son estocásticas (comportamiento) y requieren ser estimadas. Las otras dos son identidades

¿Cuáles variables son endógenas y predeterminadas?

Las endógenas son aquellas que están conjuntamente determinadas por el sistema: Y_t, C_t, YD_t y I_t

Para ver por qué estas cuatro variables están simultáneamente determinadas, se debe observar que sí se cambia una de ellas y se sigue este cambio a través del sistema, el cambio volverá a la variable causal original

Por ejemplo, si I_t aumenta por alguna razón, esto causará que Y_t aumente, lo cual retroalimentará a I_t. Están simultáneamente determinadas

¿Qué pasa con las tasas de interés? ¿Es r_t una variable endógena? La respuesta es que, hablando estrictamente, r_t no es endógena en este sistema porque r_{t-1} (no r_t) aparece en la ecuación de la inversión y r_t aparece sólo una vez en el sistema entero

En consecuencia, no hay retroalimentación simultánea a través de la tasa de interés en este modelo simple. Porque no hay simultaneidad, no hay sesgo de simultaneidad y MCO puede ser utilizado para estimar la ecuación de la tasa de interés. En esencia, esta ecuación no está en el sistema simultáneo.

¿Cuáles son las variables predeterminadas? Si la ecuación de la tasa de interés no está en el sistema simultáneo, entonces las variables predeterminadas son G_t, XN_t, T_t, C_{t-1} y r_{t-1}, pero no M_t, porque la ecuación para r_t no es parte del sistema

Para resumir, el sistema simultáneo tiene 4 ecuaciones estructurales, 4 variables endógenos y 5 variables predeterminadas

¿Cuál es el contenido económico de las ecuaciones estructurales estocásticas?

La *función de consumo* es una ecuación de rezagos distribuidos, cercana a la hipótesis del ingreso permanente de Milton Friedman: el consumo está basado no en el ingreso presente sino en la percepción que tiene el consumidor sobre sus ingresos a lo largo de su vida. En consecuencia, cambios en el ingreso transitorio no afectarían el consumo. Puesto que parece razonable hipotetizar que las percepciones del ingreso permanente están basadas en niveles pasados de ingreso, un modelo de consumo permanente simple y otro más sofisticado tienen ecuaciones similares. Se espera que β_1 y β_2 sean positivos

La *función de inversión* incluye un multiplicador y componente del costo del capital. El multiplicador β_4 mide el estímulo a la inversión que es generado por un incremento en el PIB. En un modelo Keynesiano, se esperaría que β_4 sea positivo. Por otro lado, cuanto más alto el costo del capital, se esperaría menos inversión (manteniendo constante los efectos del multiplicador), principalmente porque la tasa esperada de retorno de las inversiones del capital marginal no es suficiente para cubrir los altos costos del capital. Entonces, se espera

que β_5 sea negativa. Toma tiempo planificar e iniciar proyectos de inversión, a pesar que la tasa de interés esta rezagada en un año

La *ecuación de la tasa de interés* es una función de preferencia por liquidez, resuelta para la tasa de interés bajo el supuesto de equilibrio en el mercado de trabajo. En tal situación, un aumento en el PIB incrementará las transacciones de demanda por dinero, manteniendo constante la oferta de dinero, empujando las tasas de interés hacia arriba, de modo que se esperaría que β_7 sea positiva. Si la oferta monetaria aumenta, manteniendo constante el PIB, se esperaría que disminuyan las tasas de interés, de modo que β_8 sería negativa

Se debe recordar que el modelo keynesiano simple tiene precios constantes como supuesto subyacente

Aplicando el método de MC2E se tiene:

Etapa 1: Aún cuando el sistema tiene cuatro variables endógenas, sólo dos de ellas aparecen en el lado derecho de las ecuaciones estocásticas, de modo que sólo dos ecuaciones de forma reducida son estimadas automáticamente por los programas de computación de MC2E. También puede realizarse las estimaciones de manera “manual”, es decir, utilizando MCO en las ecuaciones de forma reducida, para obtener los siguientes resultados:

```
. reg YD G XN T C_1 r_1

      Source |       SS       df       MS              Number of obs =      31
-------------+------------------------------           F(  5,    25) = 2379.17
       Model |  13636091.6     5  2727218.32           Prob > F      =  0.0000
    Residual |  28657.1974    25   1146.2879           R-squared     =  0.9979
-------------+------------------------------           Adj R-squared =  0.9975
       Total |  13664748.8    30  455491.627           Root MSE      =  33.857

------------------------------------------------------------------------------
          YD |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           G |  -.5475457   .2008534    -2.73   0.012     -.961211   -.1338804
          XN |  -.6295365    .164692    -3.82   0.001    -.9687261    -.290347
           T |  -.3431788   .1869255    -1.84   0.078     -.728159    .0418014
         C_1 |   1.237215   .0578294    21.39   0.000     1.118113    1.356316
         r_1 |  -2.033521    4.08872    -0.50   0.623     -10.4544    6.387356
       _cons |   511.6134   86.24527     5.93   0.000      333.988    689.2389
------------------------------------------------------------------------------

Durbin-Watson d-statistic (6, 31) = 2.094304

.
. reg Y G XN T C_1 r_1

      Source |       SS       df       MS              Number of obs =      31
-------------+------------------------------           F(  5,    25) = 3948.38
       Model |  22629912.5     5   4525982.5           Prob > F      =  0.0000
    Residual |  28657.1974    25   1146.2879           R-squared     =  0.9987
-------------+------------------------------           Adj R-squared =  0.9985
       Total |  22658569.7    30  755285.656           Root MSE      =  33.857

------------------------------------------------------------------------------
           Y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           G |  -.5475457   .2008534    -2.73   0.012     -.961211   -.1338804
          XN |  -.6295365    .164692    -3.82   0.001    -.9687261    -.290347
           T |   .6568212   .1869255     3.51   0.002      .271841    1.041801
         C_1 |   1.237215   .0578294    21.39   0.000     1.118113    1.356316
         r_1 |  -2.033521    4.08872    -0.50   0.623     -10.4544    6.387356
       _cons |   511.6134   86.24527     5.93   0.000      333.988    689.2389
------------------------------------------------------------------------------

Durbin-Watson d-statistic (6, 31) = 2.094304
```

Las dos formas reducidas anteriores tienen excelentes coeficientes de ajuste global pero seguramente tienen multicolinealidad severa, debido a que tenemos series de tiempo

No es necesario realizar ninguna prueba e multicolinealidad en las ecuaciones de forma reducida ni se

considera eliminar variables como r_{t-1}, que son estadística y/o teóricamente irrelevantes

El objetivo de la primera etapa de MC2E no es generar ecuaciones de forma reducida estimadas que tengan sentido económico o estadístico, sino general instrumentos de utilidad ($\hat{Y}$, $\hat{Y}D$) para substituirlos por las variables endógenas en la segunda etapa

Se estiman las variables instrumentales para las 31 observaciones al reemplazar en las ecuaciones de forma reducida anteriores las 5 variables predeterminadas

Etapa 2: Se substituyen $\hat{Y}$ y $\hat{Y}D$ por las variables endógenas en el lado derechos de las ecuaciones estructurales, es decir:

$$C_t = \beta_0 + \beta_1 \hat{Y}D_t + \beta_2 C_{t-1} + \varepsilon_{1t}$$

$$I_t = \beta_3 + \beta_4 \hat{Y}_t + \beta_5 r_{t-1} + \varepsilon_{2t}$$

$$r_t = \beta_6 + \beta_7 \hat{Y}_t + \beta_8 M_t + \varepsilon_{3t}$$

Si se utiliza MCO en las ecuaciones de la segunda etapa anteriores se obtienen los siguientes resultados de MC2E:

```
. reg C YDhat C_1

      Source |       SS       df       MS              Number of obs =      31
-------------+------------------------------           F(  2,    28) = 4771.90
       Model |  12338712.8     2  6169356.39           Prob > F      =  0.0000
    Residual |  36199.8081    28  1292.85029           R-squared     =  0.9971
-------------+------------------------------           Adj R-squared =  0.9969
       Total |  12374912.6    30  412497.086           Root MSE      =  35.956

------------------------------------------------------------------------------
           C |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       YDhat |   .4416379   .1811383     2.44   0.021     .0705928     .812683
         C_1 |   .5403089   .1919243     2.82   0.009     .1471698     .933448
       _cons |  -24.73011   41.09578    -0.60   0.552     -108.911    59.45078
------------------------------------------------------------------------------
Durbin-Watson d-statistic(3, 31)=  1.485931


. reg I Yhat r_1

      Source |       SS       df       MS              Number of obs =      31
-------------+------------------------------           F(  2,    28) =  130.65
       Model |  593561.453     2  296780.727           Prob > F      =  0.0000
    Residual |  63606.3789    28  2271.65639           R-squared     =  0.9032
-------------+------------------------------           Adj R-squared =  0.8963
       Total |  657167.832    30  21905.5944           Root MSE      =  47.662

------------------------------------------------------------------------------
           I |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        Yhat |   .1638917   .0101534    16.14   0.000     .1430935    .1846899
         r_1 |  -5.623454   3.180208    -1.77   0.088    -12.13781    .8909062
       _cons |   33.90486   42.08179     0.81   0.427    -52.29579    120.1055
------------------------------------------------------------------------------
Durbin-Watson d-statistic(3, 31)=  1.011444

. reg r Yhat M

      Source |       SS       df       MS              Number of obs =      31
-------------+------------------------------           F(  2,    28) =   13.58
       Model |  109.479692     2  54.7398459           Prob > F      =  0.0001
    Residual |  112.833913    28  4.02978259           R-squared     =  0.4925
-------------+------------------------------           Adj R-squared =  0.4562
       Total |  222.313604    30  7.41045348           Root MSE      =  2.0074

------------------------------------------------------------------------------
           r |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        Yhat |   .0084868   .0016449     5.16   0.000     .0051173    .0118563
           M |  -.0244502   .0047231    -5.18   0.000     -.034125   -.0147754
       _cons |  -12.66314   4.011918    -3.16   0.004    -20.88118   -4.445101
------------------------------------------------------------------------------
Durbin-Watson d-statistic(3, 31)=  .6297558
```

Si se hubiesen estimado las ecuaciones estructurales directamente con MCO en lugar de MC2E, se habrían obtenidos los siguientes resultados:

```
. reg C YD C_1

      Source |       SS       df       MS              Number of obs =      31
-------------+------------------------------           F(  2,    28) = 6719.46
       Model |    12349183     2  6174591.52           Prob > F      =  0.0000
    Residual |  25729.5313    28  918.911834           R-squared     =  0.9979
-------------+------------------------------           Adj R-squared =  0.9978
       Total |  12374912.6    30  412497.086           Root MSE      =  30.314

------------------------------------------------------------------------------
           C |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          YD |   .5164862   .1161959     4.44   0.000     .2784697    .7545028
         C_1 |   .4611183   .1232441     3.74   0.001     .2086641    .7135724
       _cons |  -38.10541   29.77948    -1.28   0.211    -99.10592    22.89509
------------------------------------------------------------------------------
Durbin-Watson d-statistic(3, 31)=  .8926652


. reg I Y r_1

      Source |       SS       df       MS              Number of obs =      31
-------------+------------------------------           F(  2,    28) =  137.32
       Model |  596368.558     2  298184.279           Prob > F      =  0.0000
    Residual |   60799.274    28  2171.40264           R-squared     =  0.9075
-------------+------------------------------           Adj R-squared =  0.9009
       Total |  657167.832    30  21905.5944           Root MSE      =  46.598

------------------------------------------------------------------------------
           I |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           Y |   .1641732   .0099203    16.55   0.000     .1438523    .1844941
         r_1 |  -5.637742   3.109188    -1.81   0.081    -12.00662    .7311408
       _cons |   32.95238   41.12491     0.80   0.430    -51.28818    117.1929
------------------------------------------------------------------------------
Durbin-Watson d-statistic(3, 31)=  .9658097

. reg r Y M

      Source |       SS       df       MS              Number of obs =      32
-------------+------------------------------           F(  2,    29) =   16.37
       Model |   125.06873     2  62.5343651           Prob > F      =  0.0000
    Residual |  110.795362    29  3.82052972           R-squared     =  0.5303
-------------+------------------------------           Adj R-squared =  0.4979
       Total |  235.864092    31   7.6085191           Root MSE      =  1.9546

------------------------------------------------------------------------------
           r |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           Y |   .0082649   .0014477     5.71   0.000     .0053039    .0112259
           M |  -.0239728   .0042814    -5.60   0.000    -.0327292   -.0152164
       _cons |  -12.03228   3.462509    -3.48   0.002     -19.1139   -4.950652
------------------------------------------------------------------------------
Durbin-Watson d-statistic(3, 32)=  .6534176
```

Cuando se comparan los resultados producidos por MCO y MC2E se observa: **Primero**, no hay mucha diferencia entre ambos conjuntos de estimados. Pero si MCO es sesgado,

¿cómo ocurre esto? Cuando el ajuste de las ecuaciones reducidas en la primera etapa es excelente, entonces Y y Ŷ son virtualmente idénticas y, por lo tanto, la segunda etapa es bastante similar que los estimados de MCO

Segundo, se esperaría un sesgo positivo en la estimación de MCO y sesgo negativo pequeño en la estimación de MC2E, pero las diferencias entre MCO y MC2E parecen estar en la dirección esperada sólo la mitad de las veces. Esto puede estar causado por casos de multicolinealidad extrema en las estimaciones de MC2E y por los excelentes ajustes de las ecuaciones de formas reducidas

El problema de identificación

El problema de identificación se refiere a la obtención de estimaciones numéricas de los parámetros de las ecuaciones estructurales a partir de los coeficientes estimados de las ecuaciones de forma reducida

Si se pueden obtener tales estimaciones se dice que la ecuación está identificada; en caso contrario se dice que no está identificada o está subidentificada

Cuando una ecuación está identificada puede estar exactamente identificada o sobreidentificada. Exactamente identificada significa que se pueden obtener valores numéricos únicos de los parámetros estructurales. Mientras que cuando

está sobreidentificada se pueden obtener más de un valor numérico para los parámetros de las ecuaciones estructurales

Este problema surge debido a que diferentes conjuntos de coeficientes estructurales pueden ser compatibles con la misma información. Es decir, una ecuación de forma reducida dada puede ser compatible con diferentes ecuaciones estructurales

En el ejemplo de la oferta y demanda, dada la información sobre precios y cantidades y ningún dato adicional, no hay manera de garantizar que la estimación sea de la función de oferta o demanda; porque P_t y Q_t representan sólo puntos de intersección de las curvas de demanda y oferta debido a la condición de equilibrio. Es decir,

$$Q_{Dt} = \alpha_0 + \alpha_1 P_t + \varepsilon_{1t}$$

$$Q_{St} = \beta_0 + \beta_1 P_t + \varepsilon_{2t}$$

A pesar de haber etiquetado una ecuación como demanda y la otra como oferta, la computadora no será capaz de identificarlas a partir de los datos disponibles porque las variables del lado derecho y del izquierdo son las mismas en ambas ecuaciones. Sin la presencia de alguna variable predeterminada para distinguir entre ambas ecuaciones, seria imposible distinguir entre oferta y demanda

Si se añade una variable predeterminada a la ecuación de la oferta, por ejemplo, la ecuación vuelve:

$$Q_{St} = \beta_0 + \beta_1 P_t + \beta_2 Z_t + \varepsilon_{2t}$$

Por lo tanto, cada vez que Z (condiciones climáticas u otros factores externos) cambia, la curva de oferta se desplazará en el tiempo, pero la demanda permanece relativamente estable, por lo que los puntos dispersos trazan una curva de demanda, entonces se dice que la curva de demanda ha sido identificada

Una interpretación similar se produce para la demanda, se requiere información adicional sobre la naturaleza de dicha curva, por ejemplo, la demanda se desplaza en el tiempo debido a cambios en el ingreso, gustos, preferencias, etc., pero la oferta permanece constante ante estos cambios. En esta situación se dice que se ha identificado la curva de oferta

La manera de identificar ambas curvas es tener al menos una variable predeterminada en cada ecuación pero que no está en la otra, como por ejemplo:

$$Q_{Dt} = \alpha_0 + \alpha_1 P_t + \alpha_2 X_t + \varepsilon_{1t}$$

$$Q_{St} = \beta_0 + \beta_1 P_t + \beta_2 Z_t + \varepsilon_{2t}$$

Ahora, cuando Z cambia, la curva de oferta se desplaza y se puede identificar la curva de demanda de los datos sobre el equilibrio de precios y cantidades. Pero cuando X cambia, la curva de demanda se desplaza y se puede identificar la curva de oferta a partir de los datos. Si X y Z están altamente correlacionadas, tendremos problemas de estimación

La identificación es una precondición para la aplicación de MC2E a ecuaciones simultáneas. Es decir, el método de MC2E no puede aplicarse a una ecuación de forma reducida a menos que sea **identificada**

Por consiguiente, antes de estimar cualquier ecuación en un sistema simultáneo, el econometrista debe tratar el problema de identificación. Una vez que una ecuación se encuentra que es identificada, entonces puede ser estimada con MC2E pero si una ecuación no es identificada (subidentificada), entonces no puede ser utilizada MC2E ni importa cuán grande sea la muestra

Es importante señalar que una ecuación identificada (y que puede ser estimada por MC2E) no asegura que los estimados MC2E resultantes sean buenos.

Reglas de identificación

Como se vio anteriormente, siempre es posible utilizar las ecuaciones de forma reducida para determinar la identificación de una ecuación en un sistema de ecuaciones simultáneas, a pesar que puede se un trabajo laborioso y que toma mucho tiempo

Para evitar este trabajo se llevan a cabo las pruebas de **condiciones de orden y rango de identificación**.

Condición de orden

La condición de orden es un método sistemático para determinar si una ecuación particular en un sistema simultáneo tiene el potencial de ser identificado. La condición de orden es una condición necesaria pero no suficiente

La condición necesaria para que una ecuación esté identificada es que el número de variables predeterminadas en el sistema sea mayor o igual que el número de coeficientes de pendiente en la ecuación de interés. Es decir,

Número predeterminadas = Número coeficientes de pendiente

(en el sistema simultáneo) (en la ecuación)

De manera equivalente se puede escribir, en un modelo de M ecuaciones simultáneas, una ecuación está identificada si el número de variables predeterminadas excluidas en esta ecuación no debe ser menor que el número de variables endógenas incluidas menos 1, es decir,

$$K - k \geq m - 1$$

Pero si K – k = m – 1, entonces la ecuación está exactamente identificada, y si K – k > m – 1 la ecuación estará sobreidentificada.

Donde:

M= número de variables endógenas en el modelo

m= número de variables endógenas en una ecuación dada

K= número de variables predeterminadas en el modelo

k= número de variables predeterminada en una ecuación dada

Las variables endógenas son aquellas que están conjuntamente determinadas en el sistema en el periodo actual. Las variables predeterminadas son las exógenas más las endógenas rezagadas que podrían estar en el modelo

Ejemplos

Sea el modelo de oferta y demanda de las bebidas gaseosas:

$$Q_{Dt} = \alpha_0 + \alpha_1 P_t + \alpha_2 X_{1t} + \alpha_3 X_{2t} + \varepsilon_{Dt}$$

$$Q_{St} = \beta_0 + \beta_1 P_t + \beta_2 X_{3t} + \varepsilon_{St}$$

El modelo tiene dos variables endógenas (Q y P) y tres variables exógenas (X_1, X_2 y X_3)

Para la primera ecuación se tiene que M=2, K=3, m=2, k=2, por lo tanto:

$$3 - 2 >? 2 - 1 \text{ ó } 1 = 1$$

Por lo tanto, la primera ecuación está **exactamente identificada**. También se puede ver que puesto el número de variables predeterminadas (X_1, X_2 y X_3) es igual al número de coeficientes de pendiente de la primera ecuación (α_1, α_2, α_3), en consecuencia, la ecuación está exactamente identificada

Para la segunda ecuación se tiene que M=2, K=3, m=2, k=1, por lo tanto:

$$3 - 1 >? 2 - 1 \text{ ó } 2 > 1$$

Por lo tanto, la segunda ecuación está **sobreidentificada**. También se puede ver que puesto el número de variables predeterminadas (X_1, X_2 y X_3) es mayor al número de coeficientes de pendiente de la ecuación (β_1, β_2), en consecuencia, la ecuación está sobreidentificada

En consecuencia, el método de MC2E puede aplicarse al sistema porque ambas ecuaciones están identificadas (exactamente identificadas y/o sobreidentificadas)

Un sistema más complicado es el modelo macroeconómico simple:

$$Y_t = C_t + I_t + G_t + XN_t$$

$$C_t = \beta_0 + \beta_1 YD_t + \beta_2 C_{t-1} + \varepsilon_{1t}$$

$$I_t = \beta_3 + \beta_4 Y_t + \beta_5 r_{t-1} + \varepsilon_{2t}$$

$$YD_t = Y_t - T_t$$

El modelo tiene 5 variables predeterminadas (G_t, XN_t, T_t, C_{t-1}, r_{t-1}) y 4 endógenas (Y_t, C_t, I_t, YD_t)

La ecuación de consumo tiene dos coeficientes de pendiente a ser estimados (β_1, β_2), de modo que esta ecuación está sobreestimada ($5 > 2$), así que cumple la condición de orden de identificación

La ecuación de la inversión también tiene dos coeficientes de pendiente a ser estimados (β_4, β_5), de modo que esta ecuación está también sobreestimada ($5 > 2$), de modo que cumple la condición de orden de identificación