Causalidad y exogeneidad en econometría*

Manuel Arellano González • Jaime García Villar
London School of Economics

> *«... el día es siempre seguido por la noche..., pero los días no producen las noches».*
>
> H. M. BLALOCK, Jr. (1964), p. 9.

Introducción

A lo largo de los setenta, paralelamente a la divulgación de los métodos de series temporales, se ha popularizado entre los economistas una noción de (no) causalidad debida a Granger, basada en la asimetría de los esquemas de correlación [1]. Sin mucha precisión, la idea de Granger (1969) consistió en que, si después de extraer toda la información posible a partir de los propios valores pasados de una variable, la adición de otra variable como regresor todavía redujese la varianza del error de predicción, entonces esta última variable sería causal. Dicha idea, en cierta forma, ya estaba presente en el trabajo de Wiener (1956). Desde luego, el principal atractivo de una definición basada en esta idea es que es susceptible de contrastación empírica. Esto es, dado un par de variables siempre podemos tratar de evaluar cuál antecede a la otra a partir de la observación de la matriz de correlaciones desfasadas correspondiente.

Este trabajo intenta concretar cuál es la utilidad del concepto de (no) causalidad de Granger en econometría con propósitos de clarificación y como medio de llamar la atención sobre ciertos abusos, tanto teóricos como aplicados, que se han hecho del concepto. Para ello intentamos compendiar, junto a la noción de causalidad en la ciencia, la literatura reciente sobre exogeneidad e invarianza estructural que proviene del importante trabajo de Engle, Hendry y Richard (1980 y 1981) (en adelante EHR).

En la primera parte presentaremos definiciones precisas de la noción de Granger y pondremos de relieve su contraste con la idea de causalidad en econometría clásica (y en realidad, común a todas las ciencias). La idea es que debe hacerse una cuidadosa distinción entre los hechos observacionales (digamos, correlaciones dinámicas y no dinámicas) y las hipótesis basadas en la teoría que se puedan proponer para explicar esos hechos. Desde luego, una

* Este trabajo ha sido realizado durante la estancia de los autores en la London School of Economics como estudiantes de Ph, D. Agradecemos especialmente el apoyo recibido del Departamento de Estadística y Econometría de la Universidad de Barcelona. Asimismo, el primer autor agradece la financiación recibida de la Caja de Ahorros y Monte de Piedad de Barcelona, y el segundo autor la recibida de la London School of Economics 1980s Fund y del International Centre for Economics and Related Disciplines en la L.S.E.

[1] Entre otros, son bien conocidos Sims (1972), Sargent (1976), Sargent (1978) y Cuddington (1980), que hacen uso de la noción de Granger de una manera u otra. Por lo que respecta a nuestro país, citemos las aplicaciones de Hoyo y Terceiro (1978), Hernández-Iglesias y Hernández-Iglesias (1981), y, recientemente, el trabajo de Angulo, Raymond y Repilado (1982).

teoría para ser aceptable debe incluir entre sus implicaciones observacionales esas correlaciones, pero el esquema de correlaciones en ningún caso puede sustituir a la estructura que lo racionaliza. Dicho en otras palabras, causalidad es equivalente a una hipótesis determinada sobre un mecanismo de transmisión, y, por tanto, es la hipótesis mantenida esencial que caracteriza la estructura que intenta explicar los hechos observados. A lo largo de la exposición, serán presentados ejemplos que intentarán hacer estas ideas más evidentes, pero por el momento insistamos citando a Hicks (1979): «todo análisis causal, como he enfatizado repetidamente, depende de la teoría...» (p. 67).

En la segunda parte se discutirá la relación entre (no) causalidad de Granger e inferencia. Varios trabajos de Sims y Geweke [Sims (1972), Sims (1977) y Geweke (1978)] han enfatizado la exogeneidad como una cuestión empírica, asimilando el concepto de *exogeneidad estricta* a capacidad predictiva, esto es, a (no) causalidad de Granger. De otra parte, sus definiciones han sido formuladas en el contexto de modelos lineales y en función de perturbaciones inobservables. Siguiendo el trabajo de EHR mostraremos que para realizar inferencia condicionada a un grupo de variables, la (no) causalidad de Granger no es necesaria ni suficiente. La definición de *exogeneidad débil* de EHR en términos de las distribuciones de las variables observables pone de relieve la noción de *parámetros de interés* y se presenta como una propiedad de ciertas variables en relación a dichos parámetros que garantiza la posibilidad de hacer inferencias condicionales con información completa. Un modelo macroeconómico sencillo con expectativas racionales será utilizado para ilustrar los conceptos presentados.

A continuación consideraremos las relaciones entre (no) causalidad de Granger e *invarianza estructural*. De nuevo aquí las definiciones de invarianza y *super-exogeneidad* de EHR parecen ser los conceptos relevantes. Mostraremos que es erróneo basar inferencias acerca de invarianza estructural en general, e inferencia acerca de la inefectividad de la política económica en particular, en tests de (no) causalidad de Granger. La situación es que los resultados de dichos tests tienen una «doble no implicación» con referencia a la neutralidad (o inefectividad) de la política económica. En este sentido, serán de interés los ejemplos presentados por Buiter (1982).

Finalmente, una vez definidos los conceptos y aclarados para qué no sirven los tests de (no) causalidad de Granger intentaremos delimitar su papel en econometría. Mientras que exogeneidad débil valida la conducción de inferencia acerca de Y condicionada a Z, la (no) causalidad de Granger valida la predicción de Z y, en consecuencia, la predicción de Y condicionada a valores futuros de Z.

En resumen, este trabajo intenta cumplir tres objetivos. Primero, resaltar los perfiles, un tanto desdibujados en buena parte del trabajo de series temporales, de la noción de causalidad como una cuestión teórica, y con ello estimular a la racionalización de los hallazgos estadísticos mediante la estimación de estructuras fundamentadas teóricamente. Extraer conclusiones a partir de parámetros de formas reducidas es muchas veces tarea huera. En conexión con ello, y segundo, poner de relieve la noción de parámetros de interés en las definiciones de los requisitos necesarios para basar inferencias, predicciones o evaluación de políticas condicionadas a determinadas variables. Tercero y último, presentar estos resultados en términos de las distribuciones de las variables observables y, por tanto, lo suficientemente generales como para no limitarse al *modelo lineal dinámico,* abarcando modelos no lineales, no gaussianos o modelos con variables dependientes limitadas entre otros casos de interés.

Causalidad de Granger y causalidad en econometría

Definiciones de (no) causalidad de Granger

Consideremos un vector $n \times 1$ de variables aleatorias observables generadas en el momento t, $\underline{x}_t$ y denotemos por X_T^1 la matriz $T \times n$ de observaciones

$$X_T^1 = \begin{pmatrix} \underline{x}'_1 \\ \vdots \\ \underline{x}'_T \end{pmatrix} \quad [1]$$

Si X_0 representa la matriz de condiciones iniciales, la información disponible en el momento t viene dada por:

$$X_{t-1} = \begin{pmatrix} X_0 \\ X_{t-1}^1 \end{pmatrix} \quad [2]$$

Representamos el proceso que genera las T observaciones mediante la siguiente función de densidad conjunta de los datos

$$D(X_T^1 | X_0, \underline{\theta})$$

donde $\underline{\theta}$ es un vector de parámetros desconocidos.

Definamos una partición del vector $\underline{x}_t$

$$\underline{x}_t = \begin{pmatrix} \underline{y}_t \\ \underline{z}_t \end{pmatrix} \qquad \underline{y}_t \in R^p \qquad \underline{z}_t \in R^q \qquad p + q = n$$

La función de densidad conjunta de los datos se puede factorizar secuencialmente:

$$D(X_T^1 | X_0, \underline{\theta}) = \prod_{t=1}^{T} D(\underline{x}_t | X_{t-1}, \underline{\theta}) \quad [3]$$

Nuestra atención se centrará en las funciones de densidad condicionales

$$D(\underline{x}_t | X_{t-1}, \underline{\theta}) \qquad t = 1, 2, \ldots, T \quad [4]$$

Establecida la notación previa podemos presentar una definición general de (no) causalidad de Granger.

Definición 1

Y_{t-1}^1 no causa en el sentido de Granger a $\underline{z}_t$ con respecto a X_{t-1} si y sólo si

$$D(\underline{z}_t | X_{t-1}, \underline{\theta}) \equiv D(\underline{z}_t | Z_{t-1}, Y_0, \underline{\theta}) \qquad t = 1, \ldots, T \quad [5]$$

esto es, si y sólo si condicionadas a Z_{t-1} e Y_0, $\underset{\sim}{z}_t$ e Y^1_{t-1} son independientes en probabilidad.

La definición 1 es la definición operacional de (no) causalidad de Granger utilizada por EHR (1981) [2]. En orden a aclarar la filosofía del concepto de Granger es conveniente presentar su definición general (no operacional) recogida en Granger (1980), en donde, a la búsqueda de fundamentos últimos para su noción, Granger presenta como «axioma de causalidad» la idea de que el futuro no puede causar el pasado, junto con un supuesto de no multicolinealidad perfecta entre las variables incluidas en el conjunto de información.

Formalmente, sea Ω_{t-1} todo el conocimiento del universo disponible en el momento $t - 1$ (el conjunto de información universal). Supongamos que estamos interesados en la proposición de que la variable Y causa a la variable Z. En el momento $t - 1$, el valor $\underset{\sim}{z}_t$ será, en general, una variable aleatoria y, por tanto, puede caracterizarse mediante afirmaciones probabilísticas de la forma «Prob ($\underset{\sim}{z}_t \in A$) para un conjunto A». Esto sugiere la siguiente definición general:

Definición 2

Decimos que Y_{t-1} causa a $\underset{\sim}{z}_t$ si

$$\text{Prob}\,(\underset{\sim}{z}_t \in A|\Omega_{t-1}) \neq \text{Prob}\,(\underset{\sim}{z}_t \in A|\Omega_{t-1} - Y_{t-1}) \qquad [6]$$

para algún conjunto A y para todo t.

La definición 2 no puede ser utilizada con datos reales, puesto que el conjunto de información universal es desconocido. Por consiguiente, Granger replantea la definición general, proporcionando una definición de *(no) causalidad* con respecto a un conjunto particular de información y una definición de *causalidad* que comprende el conjunto de información universal Ω_{t-1}. La primera es equivalente a la definición 1 (en la que el conjunto de información consiste en la matriz X_{t-1}).

La definición 1 se refiere a la distribución condicional de las variables aleatorias. En algunos casos será conveniente disponer de una definición mucho menos exigente que se refiera únicamente a las medias condicionales. En vista de ello recogemos la siguiente definición adaptada a partir de la dada por Buiter (1982).

Definición 3

Y^1_{t-1} no causa una media en el sentido de Granger a $\underset{\sim}{z}_t$ con respecto a X_{t-1} si y sólo si

$$E(\underset{\sim}{z}_t|X_{t-1}, \underset{\sim}{\theta}) \equiv E(\underset{\sim}{z}_t|Z_{t-1}, Y_0, \underset{\sim}{\theta}) \qquad [7]$$

[2] Una definición similar es propuesta por Granger (1980). De otra parte, definiciones del mismo tipo se podrían presentar para los casos en que [1] no opera en ninguna de las dos direcciones «causalidad bidireccional o *feed back*» o aquellos en que la inclusión de valores contemporáneos de Y no dejase inalterada la distribución condicional $D(\underset{\sim}{z}_t|X_{t-1}, \underset{\sim}{\theta})$ («causalidad unidireccional instantánea»). [Para una introducción de estas definiciones de términos de reducción de la varianza de los errores de predicción, véase Granger (1969).]

Así, podemos referirnos a la definición 1 como «(no) causalidad de Granger en distribución» [o (no) causalidad de Granger en sentido estricto] y a la definición 3 como «(no) causalidad de Granger en media» [(no) causalidad de Granger de primer orden].

Causalidad y econometría

En este apartado contrastaremos la noción de (no) causalidad de Granger con varios de los elementos comunes de las definiciones de causación disponibles en la literatura de filosofía de la ciencia, de los que participan más o menos implícitamente las ideas tradicionales de causalidad en econometría.

Una definición filosófica de causación lo suficientemente simple y precisa para los propósitos de nuestro trabajo es la ofrecida por Feigl (1953), también adoptada por Zellner (1979) para el trabajo en econometría. De acuerdo con Feigl y Zellner, el concepto de causación es definido en términos de predecibilidad de acuerdo a una ley (o a un conjunto de leyes). A partir de esta definición es inmediato que predecibilidad sin una ley (o sin teoría) no es causación. En realidad, la necesidad del componente teórico para el trabajo econométrico es tan antigua como el trabajo econométrico mismo, y ya fue puesta de manifiesto por Frisch (1933) en el editorial del primer número de *Econométrica:*

> «La experiencia ha demostrado que cada uno de estos tres puntos de vista, el de la estadística, el de la teoría económica y el de las matemáticas, es condición necesaria, pero no por sí misma suficiente, para un entendimiento real de las relaciones cuantitativas en la vida económica moderna. Es la *unificación* de los tres lo que es poderoso. Y es esta unificación la que constituye la econometría» (página 2).

La referencia a una ley está igualmente presente en la idea de «forzamiento sistemático» de Bunge (1959) o en la noción de «producción» de Blalock (1964):

> «Si X es una causa de Y, tenemos presente que un cambio en X *produce* un cambio en Y y no meramente que un cambio en X es seguido por o asociado a un cambio en Y» (página 9).

En parecidos términos se pronuncia Jeffreys (1957), aunque utilizando predecibilidad en el sentido de alta probabilidad de que unos acontecimientos sigan a otros.

En dos puntos conviene resaltar el contraste de la definición de Feigl-Zellner con la noción de (no) causalidad de Granger. La definición de Granger comprende la idea de predecibilidad, pero no menciona leyes económicas, por lo que, en consecuencia, está en conflicto con la definición filosófica que hace referencia a la vez a predecibilidad y leyes. Adicionalmente, como señala Zellner, la definición filosófica de causalidad es adecuada para el trabajo econométrico y el de otras áreas de la ciencia

> «... una conclusión que consideramos afortunada, dada la importancia que el presente escritor asigna al principio de "unidad de la ciencia"» (página 49).

Segundo, la forma de los modelos causales que es posible considerar en el sentido de Granger es altamente restrictiva. Se consideran procesos estacionarios lineales y dentro de éstos únicamente una forma particular de modelo totalmente recursivo (lo que en econometría llamaríamos formas reducidas dinámicas).

Obsérvese que la definición de Feigl-Zellner no requiere asimetría temporal entre causa y efecto en tiempo cronológico y por ello también es aplicable a modelos en cuyas ecuaciones aparecen más de una variable endógena contemporánea.

Concluyendo, debe estar claro que (no) causalidad de Granger no es sustitutiva de la noción de causación en la literatura de filosofía de la ciencia aunque en ciertos casos puede considerarse como ingrediente de esta noción [un aspecto olvidado por la definición de Simon (1953) en la monografía número 14 de la Cowles Comission, quien enfatizando la idea de causalidad como una propiedad deductiva de los modelos no tuvo en cuenta la exigencia de capacidad predictiva]. En realidad, la definición de Granger es un ejemplo de lo que Koopmans (1947) define como *medición sin teoría,* puesto que la simple observación de regularidades (estáticas o dinámicas) en la relación entre variables no es suficiente para considerarlas como relaciones de comportamiento (causales). A este respecto destaquemos que incluso Granger y Newbold (1977) reconocen que sería más apropiado hablar de variables *temporalmente relacionadas* que de causalidad al referirse a la noción de Granger.

Granger (1980), en respuesta a la crítica de Zellner (1979), argumenta que cierta teoría podría ser útil para seleccionar la lista de variables a utilizar, pero que raramente ésta podría ir más lejos, dada la ausencia de leyes generalmente aceptadas por la mayoría de los economistas.

Desde luego, la necesidad de una lista de variables basada en alguna teoría o en el sentido común es una primera etapa en todo trabajo empírico, sin la cual se podría dar lugar a resultados tan absurdos como estadísticamente precisos, del tipo de los obtenidos en Hendry (1980). En dicho trabajo se presenta un ejercicio de exigente modelización dinámica, en el que la inflación resulta mejor explicada por la tasa acumulada de pluviosidad que por los cambios en la oferta. No obstante, disponer únicamente de esa lista de variables es insuficiente si estamos interesados en relaciones causales, puesto que es distinto encontrar *relaciones de prelación temporal* que el *mecanismo de transmisión* que las racionaliza. En otras palabras, dos variables relacionadas en la forma reducida pueden estarlo tan sólo indirectamente a través de una tercera en la forma estructural. En la práctica, la situación es probable que sea peor, debido a problemas de omisión de variables relevantes de la lista [3].

---

[3] El problema del indicador adelantado constituye un ejemplo simple de este punto. A título de ilustración, se utilizó el modelo

$$
\begin{array}{lll}
Y_t = 2. + .9X_{t-3} + \varepsilon_{1t} & & \\
Z_t = 1. + .7X_{t-1} + \varepsilon_{2t} & & \\
X_t = .7X_{t-1} + \varepsilon_{3t} & X_0 = 0 & \\
\varepsilon_{1t} \sim \mathrm{NID}\ (0, .25) & & \\
\varepsilon_{2t} \sim \mathrm{NID}(0, .16) & E(\varepsilon_{it}\varepsilon_{js}) = 0 & \forall i, j, t, s \quad i \neq j \\
\varepsilon_{3t} \sim \mathrm{NID}(0, 1.0) & &
\end{array}
$$

(en donde NID significa normal e independientemente distribuida) para generar 200 observaciones de las variables Z e Y con el fin de llevar a cabo un contraste de (no) causalidad de Granger entre estas dos variables (continúa en pág. siguiente).

Un ejemplo relativamente simple de la necesidad de restricciones teóricas en la determinación de relaciones causales son las soluciones de equilibrio a largo plazo que condicionan la especificación de los modelos de desequilibrio a corto [4].

Pese a las anteriores consideraciones, debe estar claro que el concepto de causalidad de Granger es objetivamente valioso y útil, desde un punto de vista clásico en econometría. Contrastes de la condición de Granger pueden ser capaces de generar importantes resultados empíricos que pueden entonces ser explicados por nuevas o viejas teorías económicas. Pero es importante saber exactamente qué es lo que hemos conseguido y qué no, a partir de los resultados de un contraste de (no) causalidad de Granger.

No obstante, sería ingenuo pretender que el motivo de controversia sobre el tema radica únicamente en una falta de clarificación sobre definiciones de causalidad alternativas [5]. La controversia existe y sus orígenes deben buscarse en distintos puntos de vista acerca de los tipos de inferencia posibles. Esto es, si es posible o no hacer inferencias causales. Diferentes respuestas producirán diferentes énfasis entre los conceptos estadísticos implicados. Si no creemos que las restricciones provinentes acerca de la teoría a priori pueden proporcionar estructuras identificadas, recurriremos al análisis de correlaciones para nuestro trabajo empírico [6].

No causalidad de Granger e inferencia

Cuando se trata de definir exogeneidad en relación a inferencia, el objetivo es determinar bajo qué condiciones podemos restringir nuestra atención a un modelo condicional sin perder información muestral relevante. Esto es, suponiendo que estemos interesados en la conducción de inferencia en un modelo condicional a un conjunto de variables, la exogeneidad de esas variables nos ahorraría la especificación del modelo marginal que les corresponde para aquellos propósitos.

... El test de Sims (1972) (regresión de una variable sobre los valores pasados, presentes y futuros de la otra) proporcionó los siguientes resultados:

$$\hat{Y}_t = \underset{(10.8)}{.80} + \underset{(1.2)}{.08Z_{t-2}} - \underset{(-1.4)}{.11Z_{t-1}} + \underset{(0.5)}{.04Z_t} + \underset{(0.3)}{.02Z_{t+1}} + \underset{(16.5)}{1.09Z_{t+2}} \qquad R^2 = .70$$

$$\hat{Z}_t = \underset{(-3.8)}{-.29} + \underset{(16.4)}{.58Y_{t-2}} + \underset{(2.7)}{.10Y_{t-1}} - \underset{(-0.2)}{.01Y_t} - \underset{(-2.3)}{.09Y_{t+1}} + \underset{(1.9)}{.07Y_{t+2}} \qquad R^2 = .70$$

(Entre paréntesis los valores estadísticos t.)

Tal como cabía esperar, el resultado del test muestra una relación dinámica significativa entre Z e Y, especialmente entre Y_{t-2} y Z_t, a pesar de que en el modelo que se utilizó para generar los datos no hay ninguna relación directa entre Z e Y. Es más, un cambio autónomo en Y tendría una influencia nula sobre Z.

[4] Véanse Hendry y Anderson (1977) para un modelo simultáneo. Para una utilización similar en modelos uniecuacionales, véase Davidson y col. (1978) o Dolado (1982).

[5] «Una vez he definido lo que personalmente entiendo por causación, puedo utilizar el término... Si otros quisieran referirse a mi definición pueden simplemente llamarla "causalidad de Granger" para distinguirla de definiciones alternativas» [Granger (1980), p. 333].

[6] Véanse Sims (1980) y Desai (1981) para discusiones pormenorizadas sobre este punto.

En esta sección mostraremos que en relación a tal objetivo: primero, la (no) causalidad de Granger es irrelevante y, segundo, la noción de parámetros de interés es fundamental. Para ello nos basaremos en el concepto de *exogeneidad débil* propuesto por EHR y Richard (1980). Sucintamente, podemos decir que una variable Z es débilmente exógena si para los parámetros de interés, la inferencia basada en el proceso completo (que genera X) es la misma que la inferencia basada en el proceso condicional que genera (X/Z). El punto importante es, por tanto, que las variables no son exógenas o lo dejan ser en relación a *modelos* sino en relación a *parámetros* y para determinados propósitos (inferencia, en este caso).

Consideremos el proceso generador de datos.

$$D(\underset{\sim}{x}_t | X_{t-1}, \underset{\sim}{\theta}) \qquad [8]$$

esto es, la función de densidad conjunta de los datos observables condicionada secuencialmente. En econometría se intenta modelizar [8] a partir de teorías sobre determinadas funciones de $\underset{\sim}{\theta}$, $\underset{\sim}{\psi} = \underset{\sim}{\psi}(\underset{\sim}{\theta})$, que vamos a denominar parámetros de interés.

Obsérvese que el modelo [8] se puede parametrizar en un número infinito de formas distintas, esto es, no necesariamente en términos de aquellos parámetros que se escogen para caracterizar la densidad de los datos (el vector de medias y la matriz de covarianzas para una densidad normal, por ejemplo). Por ello, en lugar de referirnos a $\underset{\sim}{\theta}$ será útil referirnos a una transformación arbitraria (biyectiva) de $\underset{\sim}{\theta}$ o reparametrización:

$$\underset{\sim}{\lambda} = \underset{\sim}{\lambda}(\underset{\sim}{\theta})$$

Es conveniente proceder de esta forma porque nos vamos a interesar en la cuestión de si los parámetros de interés son o no función únicamente de un subvector de $\underset{\sim}{\lambda}$. Es decir, suponiendo una partición de $\underset{\sim}{\lambda}$:

$$\underset{\sim}{\lambda} = \begin{pmatrix} \underset{\sim}{\lambda}_1 \\ \underset{\sim}{\lambda}_2 \end{pmatrix}$$

en la cuestión de si existe una función $\underset{\sim}{g}$ tal que

$$\underset{\sim}{\psi} = \underset{\sim}{\psi}[\underset{\sim}{\lambda}^{-1}(\underset{\sim}{\lambda})] \equiv \underset{\sim}{g}(\underset{\sim}{\lambda}_1)$$

Obsérvese también que, dada una partición de las variables $\underset{\sim}{x}'_t = (\underset{\sim}{y}'_t\ \underset{\sim}{z}'_t)$, [8] siempre se puede descomponer como la densidad de $\underset{\sim}{y}_t$ condicionada a $\underset{\sim}{z}_t$ por la densidad marginal de $\underset{\sim}{z}'_t$. En particular, si esta factorización cumple:

$$D(\underset{\sim}{x}_t | X_{t-1}, \underset{\sim}{\lambda}) = D(\underset{\sim}{y}_t | \underset{\sim}{z}_t, X_{t-1}, \underset{\sim}{\lambda}_1) D(\underset{\sim}{z}_t | X_{t-1}, \underset{\sim}{\lambda}_2) \qquad [9]$$

y la variación de $\underset{\sim}{\lambda}_1$ no está relacionada con la $\underset{\sim}{\lambda}_2$, diremos que $(\underset{\sim}{y}_t | \underset{\sim}{z}_t, \underset{\sim}{\lambda}_1)$ y $(\underset{\sim}{z}_t, \underset{\sim}{\lambda}_2)$ operan un *corte secuencial*. [9] implica que la *matriz de información*, caso de existir, *es diagonal por bloques entre* $\underset{\sim}{\lambda}_1$ *y* $\underset{\sim}{\lambda}_2$. En realidad, para muchos casos de interés, la condición de que la parametrización $\underset{\sim}{\lambda}$ opera un corte secuencial es equivalente a la condición de que la matriz de información es diagonal por bloques. Y en cualquier caso, el que la matriz de información no sea diagonal pro bloques es condición suficiente para rechazar la posibilidad de un corte secuencial. Dicho esto, estamos en condición de dar una definición precisa de exogeneidad débil.

Definición 4

z_t es *débilmente exógena* para ψ si y sólo si existe una reparametrización $\lambda' = (\lambda'_1 \lambda'_2)$ tal que

i) ψ es una función de λ_1

ii) $(y_t | z_t, \lambda_1)$, (z_t, λ_2)

operan un corte secuencial.

Asimismo, recogemos la versión más general de las definiciones de exogeneidad estricta y variables predeterminadas tal como la de Geweke (1980).

Definición 5

Una variable en un modelo estocástico, es *estrictamente exógena* si su valor en cada período es estadísticamente independiente de los valores de todas las perturbaciones aleatorias en el modelo para todos los períodos.

Definición 6

Una variable z_t es *predeterminada* en el período t si todos sus valores presentes y pasados son independientes del vector de perturbaciones contemporáneas en el modelo, u_t, y estas perturbaciones son serialmente independientes.

Las definiciones 5 y 6, en la literatura, son referidas usualmente al modelo dinámico de ecuaciones simultáneas. Y en este contexto la idea de exogeneidad estricta se suele asimilar a la de (no) causalidad de Granger [véase, por ejemplo, Sims (1977) y Geweke (1978), en donde contrastes de (no) causalidad de Granger se denominan «test de exogeneidad]. Sin embargo, como se desprende de la definición 4, la (no) causalidad de Granger no es condición necesaria ni suficiente para realizar inferencias condicionales. Esto es, puede suceder que «y no cause en el sentido de Granger a z» y a la vez z no sea exógena débil, y viceversa.

En cualquier caso, la principal limitación de las definiciones de exogeneidad estricta es la ausencia de la noción de parámetros de interés. Esta ausencia, junto al hecho de formularse en términos de perturbaciones inobservables, hace que la exogeneidad estricta pueda conseguirse siempre por construcción para *algunos* parámetros y para *algunas* perturbaciones [en realidad, esta crítica es general: lo mismo podríamos decir con respecto a los requerimientos para la consistencia de un estimador si no se han definido cuáles son los parámetros de interés, véase, por ejemplo, Hendry (1979) o Arellano (1982)]. En consecuencia, la exogeneidad estricta de z_t no asegura que no haya pérdida de información muestral relevante al hacer inferencia condicional a z_t.

Volviendo a la definición 4, podemos ver que la exogeneidad débil proporciona una condición suficiente para la conducción de inferencia condicional a z_t, observando que si se cumple (ii) entonces la función de verosimilitud de la muestra $L(\lambda, X^1_T)$ se puede factorizar en:

$$L(\lambda, X^1_T) = L_1(\lambda_1, X^1_T) L_2(\lambda_2, X^1_T) \qquad [10]$$

donde

$$L_1(\lambda_1, X^1_T) = \prod_{t=1}^{T} D(y_t | z_t, X_{t-1}, \lambda_1)$$

$$L_2(\lambda_2, X_T^1) = \prod_{t=1}^{T} D(z_t | X_{t-1}, \lambda_2)$$

y, por tanto, las dos partes de [10] se pueden analizar independientemente una de la otra. Finalmente, si además se cumple (i), entonces toda la información muestral concerniente a los parámetros de interés ψ se puede obtener a partir de la función de verosimilitud parcial $L_1(\lambda_1, X_T^1)$ [7].

Téngase en cuenta que, como señala EHR (1980, pág. 13), el concepto de exogeneidad débil no está directamente relacionado con la validación de métodos de estimación específicos, sino que se refiere a las condiciones bajo las cuales la atención puede restringirse a submodelos condicionales sin pérdida de información muestral relevante. Una selección posterior de un estimador inapropiado puede producir ineficiencia (e inconsistencia) incluso cumpliéndose las condiciones de exogeneidad débil.

En el Apéndice desarrollamos dos de los ejemplos propuestos por EHR para ilustrar los puntos anteriores.

Por último, haremos unas breves consideraciones sobre la contrastabilidad del supuesto de exogeneidad débil. En general, la exogeneidad débil puede tener o no implicaciones contrastables. Esto es, no las tendrá en el caso en que el supuesto de exogeneidad descanse únicamente en una elección de los parámetros de interés [8]. Por el contrario, será contrastable en la medida en que implique ciertas restricciones.

Evidentemente, tests de la exogeneidad débil de z_t para ψ, requieren que el modelo conjunto $D(x_t | X_{t-1}, \lambda)$ sea especificado. Sin embargo, como señalan EHR, es algo paradójico estimar los parámetros de un (potencialmente muy complicado) modelo marginal únicamente para contrastar si se necesita especificar ese modelo. Además, errores de especificación en el modelo marginal podrían inducir falsos rechazos de la hipótesis nula de exogeneidad débil. En consecuencia, en muchas circunstancias, como cuestión práctica, la exogeneidad débil descansa en supuestos a priori, lo cual, y esto es importante, no debe entenderse como una «limitación» del concepto de exogeneidad débil sino como una limitación del trabajo econométrico. El concepto de exogeneidad débil únicamente ayuda a explicitar esta situación.

Un ejemplo macroeconómico

La utilización de los tests de causalidad propuestos por Granger y Sims ha jugado un papel importante en el debate sobre la neutralidad del dinero o, en otros términos, sobre la efectividad de las políticas estabilizadoras, especialmente la política monetaria. Dado que posteriormente se dedicará especial atención a la relación entre causalidad de Granger y política de estabilización, se ha escogido para ilustrar los conceptos expuestos en esta sección un sencillo modelo macroeconómico que puede interpretarse como un modelo típico de la llamada «Nueva Macroeconomía Clásica».

El modelo se basa en el utilizado por Sargent y Wallace (1976), si bien se ha relajado una de las hipótesis básicas en los modelos de este tipo: la ecuación de oferta no es la típica ecuación de «sorpresa». Es decir, no es sólo la parte no

---

[7] Adviértase, no obstante, que [10] no implica que sea posible la factorización de la función de densidad conjunta de los datos, que requeriría el supuesto de (no) causalidad de Granger; un punto al que nos referiremos más adelante.

[8] Véase ejemplo 3.1, en EHR (1981).

anticipada de la variable sujeta a control —en este caso la tasa de crecimiento de la oferta monetaria—, la que afecta a la variable real. Por lo demás, se mantiene la hipótesis de que los agentes forman sus expectativas racionalmente, definidas en el sentido de Muth (1961). El modelo [9] es, pues

$$y_t = \alpha m_t + \beta E(m_t | I_{t-1}) + \gamma y_{t-1} + \varepsilon_{1t} \qquad [11]$$

$$m_t = \delta y_{t-1} + \varepsilon_{2t} \qquad [12]$$

en donde y representa una variable real (tasa de desempleo o desviaciones del producto nacional bruto con respecto a su nivel normal), m es la tasa de crecimiento de la oferta monetaria, I_{t-1} es el conjunto de información de los agentes que incluye el conocimiento de los valores de cualquier variable x en $t - i$ para $i \geq 1$, así como la estructura del modelo; $E(m_t | I_{t-1})$ es el valor esperado por los individuos en $t - 1$ de m. Dado que las expectativas son racionales, ese valor esperado coincide con la esperanza matemática de m dado el conjunto de información I_{t-1}. Es decir,

$$E(m_t | I_{t-1}) = \delta y_{t-1} \qquad [13]$$

De otro lado, asumimos

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim \mathrm{NID}\left(\underset{\sim}{0}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}\right) \qquad [14]$$

(NID se utiliza como abreviación de «normal e independientemente distribuidos»). habiendo flexibilizado la especificación usual en este tipo de modelos que suelen suponer $\sigma_{12} = 0$.

Sustituyendo [13] en [11] se obtiene la *forma estructural observable*

$$y_t = \alpha m_t + \theta y_{t-1} + \varepsilon_{1t} \qquad [15]$$

$$m_t = \delta y_{t-1} + \varepsilon_{2t} \qquad [16]$$

en donde $\theta = \gamma + \delta\beta$

La correspondiente forma reducida es

$$y_t = \pi y_{t-1} + \eta_t \qquad [17]$$

$$m_t = \delta y_{t-1} + \varepsilon_{2t} \qquad [18]$$

en donde

$$\pi = \alpha\delta + \gamma + \delta\beta = \alpha\delta + \theta$$

$$\eta_t = \alpha\varepsilon_{2t} + \varepsilon_{1t}$$

y

$$\begin{pmatrix} \eta_t \\ \varepsilon_{2t} \end{pmatrix} \sim \mathrm{NID}\left(\underset{\sim}{0}, \begin{pmatrix} \alpha^2\sigma_{22} + \sigma_{11} + 2\alpha\sigma_{12} & \alpha\sigma_{22} + \sigma_{12} \\ \alpha\sigma_{22} + \sigma_{12} & \sigma_{22} \end{pmatrix}\right) = \qquad [19]$$

$$= \mathrm{NID}\left(\underset{\sim}{0}, \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & \omega_{22} \end{pmatrix}\right) = \mathrm{NID}(0, \Omega)$$

[9] Nótese que en el modelo de Sargent y Wallace $\alpha = -\beta$, de tal forma que y_t depende de la tasa de crecimiento de la oferta monetaria no anticipada.

La factorización de la densidad conjunta de los datos proporciona

$$D(y_t, m_t | Y_{t-1}, M_{t-1}, \underset{\sim}{\lambda}) = D_1(y_t | m_t, Y_{t-1}, M_{t-1}, \underset{\sim}{\lambda}_1) \cdot D_2(m_t | Y_{t-1}, M_{t-1}, \underset{\sim}{\lambda}_2) \quad [20]$$

Para definir las densidades D_1 y D_2 utilizamos la forma reducida definida en [17].

De esta forma, si llamamos f_N a la función de densidad de una distribución normal univariante, podemos escribir

$$D_1(y_t | m_t, Y_{t-1}, M_{t-1}, \underset{\sim}{\lambda}_1) = f_N(y_t | bm_t + cy_{t-1}, \omega_{11.2}) \quad [21]$$

$$D_2(m_t | Y_{t-1}, M_{t-1}, \underset{\sim}{\lambda}_2) = f_N(m_t | \delta y_{t-1}, \sigma_{22}) \quad [22]$$

en donde $\underset{\sim}{\lambda}'_1 = (b, c, \omega_{11.2})$ y $\underset{\sim}{\lambda}'_2 = (\delta, \sigma_{22})$ y siendo los parámetros que definen la distribución condicional [10].

$$b = \alpha + \frac{\sigma_{12}}{\sigma_{22}} \qquad c = \gamma + \delta\left(\beta - \frac{\sigma_{12}}{\sigma_{22}}\right) \quad [23]$$

$$\omega_{11.2} = \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}} \quad [24]$$

Con respecto a la densidad marginal D_2, es inmediato derivar sus parámetros a partir de [18].

Dado que para discutir si z_t es estrictamente exógena y/o predeterminada, debemos conocer cov $(\varepsilon_{1t-i}, m_t)$ $\forall i$, es fácilmente demostrable que

$$\begin{aligned} \text{cov}\,(\varepsilon_{1t}, m_t) &= \sigma_{12} \\ \text{cov}\,(\varepsilon_{1t-i}, m_t) &= \delta\sigma_{11} \\ \text{cov}\,(\varepsilon_{1t-i}, m_t) &= 0 \qquad \forall i < 0 \;\; \text{y} \;\; \forall i > 1 \end{aligned} \quad [25]$$

---

[10] Para derivar los parámetros de la densidad condicional a partir de los de la conjunta se ha hecho uso del siguiente resultado.

Si

$$\begin{pmatrix} y \\ x \end{pmatrix} \sim \text{NID}\left(\begin{pmatrix} \mu_y \\ \mu_x \end{pmatrix}, \begin{pmatrix} v_{11} & v_{12} \\ v_{12} & v_{22} \end{pmatrix}\right)$$

entonces

$$y | x \sim \text{NID}(\mu_{y \cdot x}, v_{11 \cdot 2})$$

en donde

$$\mu_{y \cdot x} = \mu_y + \frac{v_{12}}{v_{22}}(x - \mu_x)$$

$$v_{11 \cdot 2} = v_{11} - \frac{v_{12}^2}{v_{22}}$$

Caso 1: Modelo de Sargent y Wallace ($\alpha = -\beta$, $\sigma_{12} = 0$)

La función de verosimilitud logarítmica para este modelo es

$$\log L = -\frac{T}{2}\log(\sigma_{11}\sigma_{22}) - \frac{1}{2\sigma_{11}\sigma_{22}}\left[\sigma_{22}\sum_{1}^{T}\eta_t^2 - \right. \qquad [26]$$

$$\left. - 2\alpha\sigma_{22}\sum_{1}^{T}\eta_t\varepsilon_{2t} + (\alpha^2\sigma_{22} + \sigma_{11})\sum_{1}^{T}\varepsilon_{2t}^2\right]$$

Supongamos que los parámetros de interés son α, γ y σ_{11}. Así, consideraremos log L como una función del vector de parámetros

$$\underset{\sim}{\phi}' = (\alpha\gamma\sigma_{11} \vdots \delta v_{22}) = (\underset{\sim}{\phi}_1' \vdots \underset{\sim}{\phi}_2')$$

La comprobación de que la matriz

$$-\,E\left(\frac{\partial^2 \log L}{\partial\underset{\sim}{\phi}_1 \partial\underset{\sim}{\phi}_2'}\right)$$

es distinta de cero es condición suficiente para rechazar la exogeneidad débil de m_t con respecto a α, γ y σ_{11}.

Nótese simplemente que

$$\frac{\partial^2 \log L}{\partial\gamma\partial\delta} = \frac{\alpha}{\sigma_{11}}\sum_{t=1}^{T} y_{t-1}^2$$

cuya esperanza no se anula.

Por lo demás, particularizando [25] se puede ver que m_t es predeterminada, pero no estrictamente exógena. Además, dado que $\delta \neq 0$, y causa en el sentido de Granger a m.

Caso 2: $\delta = 0$

En este caso, la parte determinista de la regla especifica un nivel constante de la cantidad de dinero. Sólo variaciones puramente aleatorias afectarán a la misma. Por tanto, y no causa en el sentido de Granger a m.

Si $\sigma_{12} \neq 0$, m_t no es débilmente exógena para α, β y γ ni tampoco predeterminada ni estrictamente exógena. Esto último, junto al hecho de que y no causa a m, ejemplifica la matización apuntada por Sargent (1979) en el sentido de que no causalidad de Granger de y a m es condición necesaria, pero no suficiente, para que m sea estrictamente exógena. Lo cual, en conexión con la idea de «exogeneidad débil basada en exogeneidad estricta», apuntada por Geweke (1980), acerca este último concepto al esquema de EHR al aceptar implícitamente la necesidad de referir exogeneidad a determinados parámetros.

Por el contrario, si $\sigma_{12} = 0$, m_t será débilmente exógena para α y γ.

Caso 3: Modelo sin restricciones

El modelo más general integra las distintas complicaciones que han aparecido en los casos anteriores: y causa en el sentido de Granger a m, m no es débilmente exógena para α, β y γ y tampoco es predeterminada ni estrictamente exógena. Obsérvese que m tampoco es exógena débil para π.

El ejemplo anterior pone de manifiesto el hecho de que (no) causalidad de Granger, exogeneidad estricta y variables predeterminadas no son conceptos comparables al de exogeneidad débil y, por tanto, no son requisitos necesarios ni suficientes para llevar a cabo inferencias condicionales de información completa.

Causalidad de Granger e invarianza estructural

Uno de los objetivos de la especificación y estimación de modelos econométricos es la evaluación de políticas alternativas. Lucas (1976) critica la práctica econométrica consistente en el uso de formas reducidas con parámetros fijos para tales fines; dado que, siempre que los agentes tengan en cuenta las reglas de control al formar sus expectativas, los parámetros de esas formas reducidas no serán invariantes a los cambios de dichas reglas. Es más, incluso los parámetros estructurales podrían verse afectados como consecuencia de reacciones de los agentes a cambios en las políticas.

Las consideraciones anteriores ponen de manifiesto la insuficiencia del concepto de exogeneidad débil como requisito para el análisis condicional de política entre regímenes, ya que no incluye ningún supuesto de invarianza estructural. De otro lado, mostraremos que es erróneo asociar (no) causalidad de Granger con ineficacia, en términos de estabilización, de las políticas económicas.

El elemento básico es, pues, la idea de invarianza estructural ante ciertos cambios en determinados parámetros *(intervenciones)*, definida por EHR en la siguiente forma.

Definición 7

Un parámetro en una ecuación de comportamiento es *estructuralmente invariante* para una clase de intervenciones si y sólo si permanece constante bajo dichas intervenciones.

La idea de invarianza estructural se extiende a modelos condicionales cuando lo son todos y cada uno de los parámetros.

En relación a ello, EHR definen el concepto de *super-exogeneidad* como requisito necesario para validar inferencias condicionales en procesos sujetos a intervenciones.

Definición 8

$\underline{z}_t$ es *super-exógena* para $\underline{\psi}$ si $\underline{z}_t$ es exógena débil para $\underline{\psi}$ y el modelo condicional $D(\underline{y}_t | \underline{z}_t, X_{t-1}, \underline{\lambda}_1)$ es estructuralmente invariante.

Obsérvese que afirmaciones de invarianza estructural y, por tanto, de *super-exogeneidad* son contrastables en relación a cambios ocurridos en el pasado en los procesos exógenos [tests clásicos de este tipo son los propuestos por Chow (1960)].

Como puede verse, y esto es la cuestión importante, la noción de (no) causalidad de Granger no es un ingrediente de ninguna de las definiciones previas. En consecuencia, es erróneo asimilar la (no) causalidad (causalidad) de Granger a inefectividad (efectividad) de las políticas económicas.

En lo que queda utilizaremos dos ejemplos para ilustrar esta doble no implicación: Primero, que causalidad de Granger e inefectividad pueden ir asociadas y, segundo, que lo mismo puede ocurrir con relación a (no) causalidad y efectividad.

Como ejemplo de la primera utilizaremos el modelo propuesto por Sargent (1976)

$$u_t = \lambda u_{t-1} + \beta_0[m_t - E(m_t | I_{t-1})] + \beta_1[m_{t-1} - E(m_{t-1} | I_{t-2})] + \varepsilon_{1t} \quad [27]$$

$$m_t = \sum_{i=1}^{n} \delta_i m_{t-i} + \varepsilon_{2t} \quad [28]$$

en donde u es la tasa de desempleo, m es el logaritmo de la cantidad de dinero en términos nominales, I_{t-1} es el conjunto de información en $t - 1$ y ε_1 y ε_2 son los términos de perturbación distribuidos normalmente y no autocorrelacionados. Para que m no cause a u en el sentido de Granger se debe satisfacer

$$E(u_t | U_{t-1}) = E(u_t | U_{t-1}, M_{t-1}) \quad [29]$$

en donde U_{t-1}, M_{t-1} se han definido como en [2].

Teniendo en cuenta que

$$E(u_{t-1} | U_{t-1}) = u_{t-1}$$

y que

$$E[E(u_{t-1} | U_{t-2}) | U_{t-1}] = E(u_{t-1} | U_{t-1})$$

podemos afirmar

$$E(u_t | U_{t-1}, M_{t-1}) = \lambda u_{t-1} + \beta_1\left(m_{t-1} - \sum_{i=1}^{n} \delta_i m_{t-1-i}\right)$$

es decir que m causa en el sentido de Granger a u. De otro lado, es inmediato que la parte determinista de la regla que define la política monetaria no tiene influencia alguna en la función de densidad de u, es decir, la política monetaria es neutral. Así pues, (no) causalidad en el sentido de Granger y neutralidad pueden actuar en direcciones distintas.

Buiter (1981) describe varios ejemplos en los que, pese a que los modelos son de características similares a los que preconizan la neutralidad del dinero, existe un margen de actuación estabilizadora para la política monetaria. Buiter (1982) muestra que el caso anterior es consistente con el hecho de que la cantidad de dinero no cause en el sentido de Granger a la variable real. Como ilustración de este punto presentamos el modelo con contratos laborales a largo plazo basado en el propuesto por S. Fischer (1977) y elaborado por Buiter. En dichos modelos dado que la duración de los contratos es mayor que los períodos entre intervenciones sucesivas sobre las variables objeto de control, los agentes no pueden responder a cambios en la política monetaria con lo cual se dan efectos sobre las variables reales [11].

El modelo es

$$m_t - p_t = y_t + u_t^m \quad [31]$$

$$y_t = \beta_1[p_t - E(p_t | I_{t-1})] + \beta_1'[p_t - E(p_t | I_{t-2})] + \beta_2 y_{t-1} + u_t^y \quad [32]$$

Las expectativas de la segunda ecuación, formadas en distintos períodos, muestran el hecho de que algunos salarios nominales fueron fijados en el período $t - 1$, mientras que otros lo fueron en $t - 2$. La regla que define la política monetaria es

$$m_t = \mu_1 u_{t-1}^y + \mu_2 u_{t-1}^m \quad [33]$$

Resolviendo el modelo para y_t se obtiene

$$\begin{aligned} y_t = \beta_2 y_{t-1} - \frac{\beta_1 + \beta_2'}{\theta} u_t^m + \frac{1}{\theta} u_t^y + \frac{\beta_1'}{1 + \beta_1'}\left[\mu_1 - \frac{\beta_2}{\theta}\right] u_{t-1}^y + \\ + \frac{\beta_1'}{1 + \beta_1'}\left[\mu_2 + \frac{\beta_2(\beta_2 + \beta_1')}{\theta}\right] u_{t-1}^m \end{aligned} \quad [34]$$

en donde $\theta = 1 + \beta_1 + \beta_1'$

La simple observación de la ecuación anterior muestra que se puede reducir la varianza de la función de densidad y_t fijando

$$\mu_1 = \frac{\beta_2}{\theta} \qquad \mu_2 = \frac{\beta_2(\beta_1 + \beta_1')}{\theta}$$

lo que proporcionaría

$$y_t = \beta_2 y_{t-1} - \frac{(\beta_1 + \beta_1')}{\theta} u_t^m + \frac{1}{\theta} y_t^y \quad [35]$$

Según [35] m no causa ni en media ni en varianza a y en el sentido de Granger, ya que

$$E(y_t | Y_{t-1}) = E(y_t | Y_{t-1}, M_{t-1}) = \beta_2 y_{t-1}$$

$$E[(y_t - E(y_t | Y_{t-1}))^2 | Y_{t-1}] = E[[y_t - E(y_t | Y_{t-1}, M_{t-1}))^2 | Y_{t-1}, M_{t-1}] =$$

$$= \left(\frac{\beta_1 + \beta_1'}{\theta}\right)^2 \sigma_m^2 + \frac{1}{\theta^2}\sigma_y^2$$

en donde

$$\text{Var}(u_t^y) = \sigma_y^2$$

y

$$\text{Var}(u_t^m) = \sigma_m^2$$

[11] Sólo una indicación de los contratos que reprodujera cuál hubiese sido el comportamiento de los individuos en caso de poder responder a la nueva información, mantendría el carácter neutral de la política monetaria.

Es decir, una política estabilizadora es consistente con el hecho de que la cantidad de dinero no causa en el sentido de Granger a la variable real (PNB).

Los ejemplos presentados en esta sección descalifican los intentos de justificar la eficacia o ineficacia de determinadas políticas en base a tests de (no) causalidad de Granger.

No causalidad de Granger y predicción condicional

A partir de la discusión previa, situar la utilidad del concepto de (no) causalidad de Granger en econometría es inmediato. Para ello, el punto de partida es la consideración de los requisitos necesarios para la factorización de la función de densidad conjunta de los datos como opuesta al problema de factorización de la función de verosimilitud. En la segunda parte, observamos que la exogeneidad débil era condición suficiente para la segunda (y de ahí que se definiera como el requisito necesario para validar inferencia condicional) pero no para la primera. Precisamente es la *adición* del supuesto de (no) causalidad de Granger lo que permite la factorización de la función de densidad conjunta de los datos. Esto es, si $\underset{\sim}{y}$ opera un corte secuencial y además $\underset{\sim}{y}$ no causa en el sentido de Granger a $\underset{\sim}{z}$, $D(X_T^1 | X_0, \underset{\sim}{\lambda})$ se puede factorizar de la siguiente forma:

$$D(X_T^1 | X_0, \underset{\sim}{\lambda}) = D(Y_T^1 | Z_T^1, X_0, \underset{\sim}{\lambda}_1) D(Z_T^1 | X_0, \underset{\sim}{\lambda}_2) \qquad [36]$$

en donde

$$D(Y_T^1 | Z_T^1, X_0, \underset{\sim}{\lambda}_1) = \prod_{t=1}^{T} D(\underset{\sim}{y}_t | \underset{\sim}{z}_t, X_{t-1}, \underset{\sim}{\lambda}_1) \qquad [37]$$

$$D(Z_T^1 | X_0, \underset{\sim}{\lambda}_2) = \prod_{t=1}^{T} D(\underset{\sim}{z}_t | Z_{t-1}, Y_0, \underset{\sim}{\lambda}_2) \qquad [38]$$

(Véase, EHR, 1981, pág. 10.)

Insistimos en la idea de adición, puesto que la definición de (no) causalidad de Granger no incluye supuestos acerca de parámetros y por ello debe ser completada asumiendo la exogeneidad débil.

La trascendencia de la factorización anterior radica en que nos permite tratar a $\underset{\sim}{z}_t$ como si fuese fijo en muestras repetidas. En otras palabras, nos asegura que los valores tomados en el pasado por $\underset{\sim}{y}$ no afectarán a valores presentes o futuros de $\underset{\sim}{z}$, y, por tanto, podemos hacer predicciones de $\underset{\sim}{y}$ condicionadas a $\underset{\sim}{z}$.

A partir de la consideración conjunta de exogeneidad débil y (no) causalidad de Granger, EHR definen *exogeneidad fuerte* en tanto que valedora de la predicción condicional [12]. De ello se desprende que los tests de (no) causalidad de Granger son tests legítimos para contrastar la exogeneidad fuerte.

Por último, es necesario resaltar la importancia de disponer de los tres conceptos diferentes de exogeneidad definidos a lo largo del trabajo (débil, fuerte y

[12] Véase el ejemplo 2 del Apéndice para una ilustración del concepto de exogeneidad fuerte en contraste con el de exogeneidad débil.

super-exogeneidad). El objeto de una afirmación de exogeneidad es permitir el análisis de un conjunto de variables sin tener que especificar con exactitud cómo se determina un segundo conjunto de variables relacionado con aquéllas. Tal análisis puede incluir conjuntamente o por separado inferencia, predicción o evaluación de políticas. Puesto que la «exogeneidad» en cada caso requiere diferentes condiciones, los tres conceptos de exogeneidad son necesarios.

Comentario final

Esperamos que la discusión y el material presentados en este trabajo contribuyan a que se ponga más énfasis en la búsqueda de relaciones empíricas fundamentadas en mecanismos lógicos. Si queremos determinar la existencia de relaciones estables entre magnitudes económicas, que sean útiles al político y al estudioso, debemos ser tan cuidadosos en la coherencia teórica del trabajo realizado como en su rigor estadístico.

Apéndice

Ejemplo 1

En este ejemplo se considera un caso en que la variable explicativa en la ecuación de interés no es causada en el sentido de Granger por y_t, aunque, sin embargo, no es exógena débil. El modelo es:

$$y_t = \beta z_{t-1} + u_t \qquad [A1]$$

$$u_t = \varepsilon_{1t} + \phi \varepsilon_{2t-1} \qquad [A2]$$

$$z_t = \gamma z_{t-1} + \varepsilon_{2t} \qquad [A3]$$

donde

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim IN\left(\underset{\sim}{0}, \begin{pmatrix} \omega_{11} & 0 \\ 0 & \omega^{22} \end{pmatrix}\right) \qquad [A4]$$

Este modelo describe una situación razonable: z_t es generada por un proceso autorregresivo de primer orden y los agentes, cuyo comportamiento viene dado por [A1] y [A2], ajustan con un período de retraso tanto el componente sistemático como el inesperado de z_t. Asumimos, por tanto, que los parámetros de interés son β y ϕ.

La forma reducida viene dada por:

$$y_t = \pi_1 z_{t-1} + \pi_2 z_{t-2} + \varepsilon_{1t} \qquad [A5]$$

$$z_t = \gamma z_{t-1} + \varepsilon_{2t} \qquad [A3]$$

donde

$$\pi_1 = \beta + \phi \quad \text{y} \quad \pi_2 = -\gamma\phi \qquad [A6]$$

La función de verosimilitud logarítmica viene dada por

$$\log L = -\frac{T}{2}(\log \omega_{11} + \log \omega_{22}) + \frac{1}{2}\left(\frac{\sum_{t=1}^{T} \varepsilon_{1t}^2}{\omega_{11}} + \frac{\sum_{t=1}^{T} \varepsilon_{2t}^2}{\omega_{22}}\right) \quad [A7]$$

que podemos considerar parametrizada en términos del vector

$$\underset{\sim}{\delta}' = (\beta\phi\omega_{11} \mid \gamma\omega_{22}) = (\underset{\sim}{\delta}'_1 \underset{\sim}{\delta}'_2)$$

si los parámetros de interés son β, ϕ y ω_{11}, o

$$\underset{\sim}{\delta}^{*\prime} = (\pi_1 \pi_2 \omega_{11} \mid \gamma\omega_{22}) = (\underset{\sim}{\delta}_1^{*\prime} \underset{\sim}{\delta}_2^{*\prime})$$

en el caso de que estemos interesados en π_1, π_2 y ω_{11}.

Obsérvese que la descomposición [3] se particulariza en este caso a:

$$\log L_1 = -\frac{T}{2}\log \omega_{11} - \frac{\sum_{t=1}^{T} \varepsilon_{1t}^2}{2\omega_{11}} \quad [A8]$$

$$\log L_2 = -\frac{T}{2}\log \omega_{22} - \frac{\sum_{t=1}^{T} \varepsilon_{2t}^2}{2\omega_{22}} \quad [A9]$$

Es condición necesaria para la exogenediad débil de z_t respecto a β, ϕ y ω_{11} que la matriz de información sea diagonal por bloques entre $\underset{\sim}{\delta}_1$ y $\underset{\sim}{\delta}_2$. Es decir, la matriz

$$-E\begin{pmatrix} \dfrac{\partial^2 L}{\partial\beta\,\partial\gamma} & \dfrac{\partial^2 L}{\partial\beta\,\partial\omega_{22}} \\ \dfrac{\partial^2 L}{\partial\phi\,\partial\gamma} & \dfrac{\partial^2 L}{\partial\phi\,\partial\omega_{22}} \\ \dfrac{\partial^2 L}{\partial\omega_{11}\partial\gamma} & \dfrac{\partial^2 L}{\partial\omega_{11}\partial\omega_{22}} \end{pmatrix}$$

debería anularse. Sin embargo, para el caso que nos ocupa se púede comprobar que toma el valor

$$-E\begin{pmatrix} \dfrac{\phi}{\omega_{11}}\sum_{t=1}^{T} z_{t-1}z_{t-2} & 0 \\ \dfrac{1}{\omega_{11}}\left(\phi\sum_{t=1}^{T} z_{t-1}z_{t-2} - \sum_{t=1}^{T} z_{t-2}\varepsilon_{1t} - \gamma\phi\sum_{t=1}^{T} z_{t-2}^2\right) & 0 \\ \dfrac{\phi}{\omega_{11}^2}\sum_{t=1}^{T} z_{t-2}\varepsilon_{1t} & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{T\gamma\phi}{1-\gamma^2} \cdot \frac{\omega_{22}}{\omega_{11}} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

que obviamente es distinto de una matriz de ceros. En consecuencia, z_t no es débilmente exógena para β, ϕ y ω_{11}, pese a que y no causa a z en el sentido de Granger.

Por el contrario, es fácilmente comprobable que

$$-E \begin{pmatrix} \frac{\partial^2 L}{\partial\pi_1 \partial\gamma} & \frac{\partial^2 L}{\partial\pi_1 \partial\omega_{22}} \\ \frac{\partial^2 L}{\partial\pi_2 \partial\gamma} & \frac{\partial^2 L}{\partial\pi_2 \partial\omega_{22}} \\ \frac{\partial^2 L}{\partial\omega_{11} \partial\gamma} & \frac{\partial^2 L}{\partial\omega_{11} \partial\omega_{22}} \end{pmatrix} = 0$$

Lo cual, para el modelo que estamos considerando, es equivalente a que $(y_t | z_t, \delta_1^*)$ y (z_t, δ_2^*) operan un corte secuencial. Por tanto, z_t es débilmente exógena para π_1, π_2 y ω_{11}.

Ejemplo 2

Consideremos el modelo

$$y_t = \beta z_t + u_t \qquad [A10]$$

$$u_t = \rho u_{t-1} + \varepsilon_{1t} \qquad [A11]$$

$$z_t = \gamma y_{t-1} + \varepsilon_{2t} \qquad [A12]$$

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim IN \left(\underset{\sim}{0}, \begin{pmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{pmatrix} \right) \qquad [A13]$$

cuyo esquema también responde a una situación usual en economía: los agentes reaccionan al control z_t, que a su vez se basa en el comportamiento del sistema en el pasado (y_{t-1}). Supongamos que β y ρ son los parámetros de interés.

La forma reducida consiste en

$$y_t = \pi_1 y_{t-1} + \pi_2 z_{t-2} + v_t \qquad [A14]$$

$$z_t = \gamma y_{t-1} + \varepsilon_{2t} \qquad [A12]$$

con

$$\pi_1 = \rho + \beta\gamma \quad y \quad \pi_2 = -\beta\rho \qquad [A15]$$

$$\begin{pmatrix} v_t \\ \varepsilon_{2t} \end{pmatrix} \sim IN \left(\underset{\sim}{0}, \begin{pmatrix} \sigma_{11} + \beta^2\sigma_{22} & \beta\sigma_{22} \\ \beta\sigma_{22} & \sigma_{22} \end{pmatrix} \right) = IN(0, \Omega) \qquad [A16]$$

La función de verosimilitud logarítmica de este modelo es

$$\log L = -\frac{T}{2} \log (\det \Omega) - \frac{1}{2} \sum_{t=1}^{T} (v_t \varepsilon_{2t}) \Omega^{-1} (v_t \varepsilon_{2t})' \qquad [A17]$$

que puede descomponerse en

$$\log L = \log L_1 + \log L_2 \qquad [A18]$$

donde

$$\log L_1 = -\frac{T}{2}\log\sigma_{11} - \frac{1}{2\sigma_{11}}\sum_{t=1}^{T}(y_t - \beta z_t - \rho y_{t-1} + \beta\rho z_{t-1})^2$$

y

$$\log L_2 = -\frac{T}{2}\log\sigma_{22} - \frac{1}{2\sigma_{22}}\sum_{t=1}^{T}(z_t - \gamma y_{t-1})^2$$

se puede comprobar que z_t es exógena débil para β y ρ. Sin embargo, y causa en el sentido de Granger a z (como es evidente a partir de [A12]). En consecuencia, z no es fuertemente exógena. El que z_t sea débilmente exógena para β y ρ permite llevar a cabo inferencias sobre β y ρ condicional a z_t, cuyo modelo no precisa ser especificado (nótese que β y ρ deben estimarse conjuntamente, por ejemplo, utilizando un método de máxima verosimilitud. El uso de un procedimiento tipo Cochrane-Orcutt produciría estimadores inconsistentes para β y ρ). Sin embargo, no podemos predecir y a partir de z sin tener en cuenta cómo, a su vez, y influencia los valores posteriores de z.

Bibliografía

ANGULO, J.; RAYMOND, J. L., y REPILADO, A.: *Relaciones de causalidad en Economía y criterios estadísticos para detectar su existencia*. Instituto de Estudios Fiscales. Monografía núm. 16, 1982.

ARELLANO, M.: *Errores de especificación en modelos dinámicos: nota sobre el análisis de inconsistencias*. Facultad de Ciencias Económicas de Barcelona, 1982.

BLALOCK, Jr., H. M.: *Causal Inference in Nonexperimental Research*. The University of North Carolina Press. Chapel Hill, 1964.

BUITER, W. H.: «The Superiority of Contingent over Fixed Rules in Models with Rational Expectations». *The Economic Journal*, vol. 91, núm. 363, 647-670, 1981.

BUITER, W. H.: «Granger-Causality and Stabilization Policy». *Centre for Labour Economics. London School of Economics. Discussion Paper*, núm. 128, 1982.

BUNGE, M.: *Causality*. Cambridge: Harvard University Press, 1959.

CHOW, G. C.: «Tests for Equality Between Sets of Coefficients in Two Linear Regressions». *Econometrica*, vol. 28, núm. 3, 591-605.

CUDDINGTON, J. T.: «Simultaneous-equations Tests of the Natural Rate and Other Classical Hypotheses». *Journal of Political Economy*, vol. 88, núm. 3, 539-549, 1980.

DAVIDSON, J.; HENDRY, D. F.; SRBA, F., y YEO, S.: «Econometric Modelling of the Aggregate Time-Series Relationship Between Consumers' Expenditure and Income in the United Kingdom». *The Economic Journal*, vol. 88, núm. 352, 661-692, 1978.

DESAI, M.: *Testing Monetarism*. Frances Pinter Publishers Ltd. Londres, 1981.

DOLADO, J. J.: «Procedimientos de búsqueda de especificación dinámica: el caso de la demanda de M3 en España». *Banco de España. Servicio de Estudios. Estudios Económicos*, núm. 27, 1982.

ENGLE, R. F.; HENDRY, D. F., y RICHARD, J. F.: «Exogeneity, Causality and Structural Invariance in Econometric Modelling». *C. O. R. E. Discussion paper*, 80-38, 1980.

ENGLE, R. F.; HENDRY, D. F., y RICHARD, J. F.: «Exogeneity». *Econometrica*, vol. 51, núm. 2, 277-305, 1983.

FEIGL, H.: «Notes on Causality», en H. Feigl y M. Brodbeck (ed.), *Readings in the Philosophy of Science*. New York: Appleton-Century-Crofts, 1953.
FISCHER, S.: «Long-term Contracts, Rational Expectations and the Optimal Money Supply Rule». *Journal of Political Economy*, vol. 85, núm. 1, 191-205, 1977.
FRISCH, R.: Editorial. *Econometrica*, vol. 1, núm. 1, 1-4, 1933.
GEWEKE, J.: «Testing the Exogeneity Specification in the Complete Dynamic Simultaneous Equations Model». *Journal of Econometrics;* vol. 7, núm. 2, 163-185, 1978.
GEWEKE, J.: «Causality, Exogeneity and Inference». IV World Congress of the Econometric Society. Aix-en-Provence, 1980.
GRANGER, C. W. J.: «Investigating Causal Relations by Econometric Models and Cross-spectral Methods». *Econometrica*, vol. 37, núm. 3, 424-435, 1969.
GRANGER, C. W. J.: «Testing for Causality: a Personal Viewpoint». *Journal of Dynamics and Control*, vol. 2, núm. 4, 329-352, 1980.
GRANGER, C. W. J., y NEWBOLD, P.: *Forecasting Economic Time Series*. Academic Press, 1977.
HENDRY, D. F.: «The Behaviour of Inconsistent Instrumental Variables Estimators in Dynamic Systems with Autocorrelated Errors». *Journal of Econometrics*, vol. 9, núm. 3, 295-314, 1979.
HENDRY, D. F.: «Econometrics: Alchemy or Science?». *Economica*, vol. 47, núm. 188, 387-406, 1980.
HENDRY, D. F., y ANDERSON, G. J.: «Testing Dynamic Specification in Small Simultaneous Systems: an Application to a Model of Building Society Behavior in the United Kingdom», en M. D. Intriligator (ed.). *Frontiers in Quantitative Economics*, vol. III. A. North-Holland. Amsterdam, 1977.
HICKS, J.: *Causality in Economics*. Brasil Blackwell. Oxford, 1979.
HERNANDEZ-IGLESIAS, C., y HERNANDEZ-IGLESIAS, F.: «Causality and the Independence Phenomenon: the Case of the Demand for Money». *Journal of Econometrics*, vol. 15, núm. 2, 247-263, 1981.
HOYO, J. del, y TERCEIRO, J.: «Causalidad en series temporales. Alguna evidencia empírica para la Economía española». *Cuadernos de Economía*, vol. 6, núm. 17, 431-459, 1978.
JEFFREYS, H.: *Scientific Inference* (2.ª ed.). Cambridge, 1957.
KOOPMANS, T. C.: «Measurement Without Theory». *Review of Economics and Statistics*, vol. 29, núm. 3, 161-172, 1947.
LUCAS, R. E., Jr.: «Econometric Policy Evaluation: a Critique», en R. Brunner y A. H. Meltzer (ed.). *The Phillips Curve and Labor Markets*. North-Holland. Amsterdam, 1976.
MUTH, J. F.: «Rational Expectations and the Theory of Price Movements». *Econometrica*, vol. 29, núm. 3, 315-335, 1961.
RICHARD, J. F.: «Models with Several Regimes and Changes in Exogeneity». *The Review of Economic Studies*, vol. 47, núm. 146, 1-20, 1980.
SARGENT, T. J.: «A Classical Macroeconometric Model for the United States». *Journal of Political Economy*, vol. 84, núm. 2, 631-640, 1976.
SARGENT, T. J.: «Estimation of Dynamic Labor Demand Schedules under Rational Expectations». *Journal of Political Economy*, vol. 86, núm. 6, 1009-1044, 1978.
SARGENT, T. J.: «Causality, Exogeneity and Natural Rate Models: Reply to C. R. Nelson and B. T. McCallum». *Journal of Political Economy*, vol. 87, núm. 2, 403-409, 1979.
SARGENT, T. J., y WALLACE, N.: «Rational Expectations and the Theory of Economic Policy». *Journal of Monetary Economics*, vol. 2, núm. 2, 169-183, 1976.
SIMON, H. A.: «Causal Ordering and Identifiability», en W. C. Hood y T. C. Koopmans (ed.). *Studies in Econometric Method Cowles Comission Monograph*, núm. 14. John Wiley & Sons. New York, 1953.
SIMS, C. A.: «Money, Income and Causality». *American Economic Review*, vol. 62, núm. 4, 540-552, 1972.
SIMS, C. A.: «Exogeneity and Causal Ordering in Macroeconomic Models», en C. A. Sims (ed.). *New Methods in Business Cycle Research*. Documentos presentados a una Conferencia en noviembre de 1975, financiada por Federal Reserve Bank of Minneapol, 1977.
SIMS, C. A.: «Macroeconomics and Reality». *Econometrica*, vol. 48, núm. 1, 1-48, 1980.
WIENER, N.: «The Theory of Prediction», en E. F. Beckenback (ed.). *Modern Mathematics for Engineers*. Series 1, Cap. 8, 1956.
ZELLNER, A.: «Causality and Econometrics», en K. Brunner y A. H. Meltzer (ed.). *Three Aspects of Policy and Policy Making*. North-Holland. Amsterdam, 1979.