

CAUSALIDAD Y ECONOMETRÍA *

Antonio García-Ferrer

1. INTRODUCCIÓN

Cada vez que leo o escucho la frase de que «cualquier economista sabe que la simultaneidad en el tiempo (correlación) no tiene por qué implicar causalidad», me echo a temblar. La razón de mi zozobra está relacionada con el hecho de que esta «aparente» confusión epistemológica está mucho más generalizada de lo que cabría esperar. No solamente entre legos en la materia, sino también entre los propios estudiantes y profesionales de la Economía¹. Al fin y al cabo, el concepto de causalidad en Economía es casi tan antiguo como la propia disciplina. No en vano, ya aparece en el título completo de la obra pionera de Adam Smith, *La Riqueza de las Naciones*, y en innumerables trabajos de notables filósofos y economistas como David Hume, Francis Bacon o John Stuart Mill, entre muchos otros. Sin embargo, aunque el concepto de causalidad se ha tratado de forma exhaustiva en Filosofía, y se ha utilizado igualmente a la hora de interpretar los resultados empíricos de muchas disciplinas (incluyendo la Econometría), la mayoría de los libros de texto econométricos suelen excluir los

* Este trabajo está financiado por la Universidad Autónoma de Madrid (CMUAM-09-SHD-011), y el Ministerio de Educación y Ciencia (SEJ2006-04957). Agradezco a Marcos Bujosa y a Alfonso Novales sus comentarios y sugerencias en una primera versión del trabajo, y a Juan Carlos García-Bermejo su invitación para escribirlo.

1. Aunque la mayoría de mis estudiantes entienden y sonríen al presentar ejemplos chuscos como la relación entre el precio del pan en Inglaterra y el nivel del mar en Venecia, o la relación entre el número de cigüeñas y el de nacimientos en Baviera; tuercen el gesto en señal de incompreensión cuando los ejemplos utilizados se refieren a relaciones entre oferta monetaria y tipos de interés, o sobre los efectos que una legislación sobre salarios mínimos tiene sobre la tasa de desempleo.

vocablos *causalidad* y *causa* de sus índices de materias. Estos mismos libros de texto tampoco aportan demasiada luz a la hora de esclarecer el dilema entre los conceptos de causalidad y correlación. El simple hecho de que dos variables guarden correlación no implica que una sea la causa de la otra. Una correlación significa, sencillamente, que existe una relación entre dos factores —llamémoslos X e Y—, pero no explica el sentido de dicha relación. Es posible que X cause Y; también es posible que Y cause X; y también lo es que tanto X como Y sean causadas por alguna otra variable Z².

En otros muchos textos de Economía, sin embargo, se reconoce abiertamente que el propósito de cualquier ley científica es la posibilidad de permitir una predicción fiable del fenómeno que estamos analizando, ya que dicha predicción permite un control posterior de dicho fenómeno. Sin predicción no hay control, ya que sin conocer la causa de un evento no podemos anticipar o prevenir su ocurrencia. Aunque la palabra *causa* es central en el argumento anterior, nunca encontramos una definición precisa de la misma. Algunos piensan que dicha definición es *evidente*; y otros, que su significado es *controvertido*. Cabe admitir, por tanto, la existencia de definiciones alternativas, unas contrastables y otras no. Con estos ingredientes, lo que es seguro es que la polémica está servida.

Como tendré ocasión de exponer en páginas posteriores, el concepto de causalidad (para cualquier disciplina científica) es difícil y sutil, y no viene completamente caracterizado ni por el concepto de correlación temporal ni por la idea de una reducción de la varianza predictiva con respecto a un determinado conjunto de información. También veremos cómo dicho concepto (lejos de pretender alcanzar una validez generalizada) ha de venir mediatizado por los tipos, formas, dominios y niveles de las leyes que intentamos corroborar. En la sección siguiente, presentaré una breve revisión del concepto de causalidad en Economía. En la sección 3, analizaré los problemas relacionados con los tests estadísticos de causalidad y, finalmente, en la sección 4 se resumen las conclusiones.

2. La mayoría de los libros de texto de Econometría tienden a evitar de forma explícita un lenguaje causal a la hora de describir los distintos procedimientos de estimación. Así por ejemplo, al tratar los modelos de ecuaciones simultáneas se utiliza la división entre variables endógenas o exógenas, según que dichas variables se determinen dentro o fuera del modelo. Sin embargo, nunca se indica si esa «determinación» es funcional o causal. Por otra parte, se supone que el término de error en el modelo de regresión representa perturbaciones aleatorias que captan posibles efectos de variables omitidas, errores de medida y otras influencias. De nuevo, no queda claro si estos términos conllevan un significado causal o no.

2. BREVE REVISIÓN DEL CONCEPTO DE CAUSALIDAD EN ECONOMÍA

Desde los primeros economistas, pero especialmente a partir de la segunda mitad del siglo XX, el concepto de causalidad ha generado una polémica continua cuyo final no parece cercano. En el fondo de la misma subyacen tanto su significado controvertido como la utilización que queremos dar al concepto; y en ella se han visto envueltos brillantes economistas. Desde Keynes a Friedman, pasando por una larga lista que incluiría a Simon, Hicks, Sims, Wold, Lucas, Zellner, Granger y Heckman, entre muchos otros. Con media docena de premios Nobel en medio de la polémica, no es extraño que el tema siga concitando tanto interés.

El punto de partida reciente, desde el que resulta más fácil centrar la presentación del concepto de causalidad, son los distintos trabajos pioneros de la Comisión Cowles hacia la mitad del siglo pasado. En ellos, la noción de causalidad pretendía entender el funcionamiento de un sistema complejo mediante el análisis de los mecanismos de los distintos componentes (subsistemas). Si estamos interesados en comprender cómo funciona un mercado elemental (p. e., el mercado del trigo), deberíamos considerar un determinado número de mecanismos que afectan a cantidades y precios. La complejidad aparece a la hora de determinar el número de mecanismos a incluir que, lógicamente, dependerá de dónde establezcamos los límites del sistema.

Posiblemente, la primera aportación significativa de esta aproximación metodológica sea el trabajo de H. Simon (1953) *Causal ordering and identifiability* que fue publicado como el capítulo 3 de una de las monografías más influyentes sobre metodología econométrica, *Studies in Econometric Method*. El contexto del trabajo de Simon surgió en un debate entre la Comisión Cowles (como propulsora de los modelos de ecuaciones simultáneas) y Herman Wold, que defendía la alternativa de modelos recursivos³. El tema de la causalidad estaba en el centro del debate y la intención de Simon era la de proporcionar «una base clara y rigurosa para determinar cuándo podemos establecer una ordenación causal entre dos o más variables de un modelo» (Simon, 1953, 51). La discusión, sin embargo, venía limitada en el caso de Simon a sistemas de ecuaciones lineales sin perturbaciones aleatorias. Para introducir la noción de ordenación causal, era esencial que la estructura del modelo fuese completa («autocontenida», en la terminología de Simon), entendiéndose como tal cuando el número de las variables en la estructura del modelo fuese igual al número de ecuaciones estructurales.

3. Una discusión detallada de este debate puede verse en Morgan (1991).

Para ilustrar el análisis causal de Simon revisaremos su modelo más sencillo sobre la determinación de los precios del trigo utilizando el siguiente modelo lineal de tres ecuaciones:

$$\beta_{11}y_1 = \beta_{10} \quad (1)$$

$$\beta_{21}y_1 + \beta_{22}y_2 = \beta_{20} \quad (2)$$

$$\beta_{32}y_2 + \beta_{33}y_3 = \beta_{30} \quad (3)$$

donde y_1 es un índice de tiempo (meteorológico) favorable, y_2 es el volumen de la cosecha de trigo; y_3 es el precio del trigo y los β 's son parámetros. En este modelo Simon suponía que «la climatología dependía de un parámetro β_{10} (exógena), la cosecha dependía del tiempo (ignorando una posible dependencia de la oferta sobre el precio); y el precio del trigo del volumen de la cosecha» (p. 58). Como el valor de y_1 se determina a partir de (1), es posible sustituir su valor en las ecuaciones (2) y (3) para obtener un sistema reducido con las variables y_2 e y_3 . La sustitución de (1) en (2) proporciona una relación para determinar el valor de y_2 que, una vez sustituido en (3) conjuntamente con el valor de y_1 obtenido en (1), determina el valor de y_3 . En estas condiciones, Simon interpreta que « y_1 es la causa directa de y_2 , e y_2 de y_3 » (p. 58). De forma esquemática: $y_1 \rightarrow y_2 \rightarrow y_3$ ⁴.

Quizá más que ninguna otra aportación sobre el tema, el trabajo de Simon ha sido analizado en profundidad por numerosos economistas y metodólogos⁵. En algunos casos, el esquema de Simon ha servido para revivir viejas polémicas metodológicas sobre las similitudes y diferencias en los enfoques causales de autores tan dispares como Hicks y Friedman (p. e., Held, 1988, y Hammond, 1988). En otros, ha servido para dilucidar importantes cuestiones de metodología econométrica como son las diferencias entre la identificación y las restricciones de exclusión en los modelos de ecuaciones simultáneas⁶.

4. Simon también planteó sistemas más complicados que el representado por las ecuaciones (1)-(3). Sin embargo, este sistema más sencillo revela muy bien la naturaleza de su concepto de causalidad y de ordenaciones causales. En trabajos posteriores, p. e., Simon (1988), amplía su concepto de causalidad a sistemas dinámicos.

5. Véanse, entre otros, Zellner (1979), Hicks (1979), Hoover (2001), Reiss y Cartwright (2003), y Le Roy (2006).

6. Las restricciones de exclusión juegan un papel crucial a la hora de determinar ordenaciones causales. También lo juegan a la hora de determinar si los parámetros de un modelo están o no identificados. A pesar de este rol común a la hora de determinar causalidad e identificación, se trata de conceptos distintos. La causalidad es una ordenación de las variables internas (endógenas), mientras que la identificación tiene que ver con la posibilidad de hacer inferencias sobre los parámetros de la distribución de las variables observadas.

A modo de resumen, las principales razones por las que la caracterización de Simon no es tan utilizada hoy día, a pesar de su valor histórico, son las siguientes:

1. La hipótesis de Simon de linealidad en las variables limita la aplicabilidad de su análisis. Muchos modelos en la actualidad no son lineales y, en consecuencia, no podemos asociar las ordenaciones causales con los coeficientes.
2. Simon dejó claro que su análisis de causalidad se restringía al contexto de un modelo formal y no, tal como se aplica directamente, a la realidad o a la *realidad percibida*. Esta idea contrasta con la mayoría de las discusiones en Economía (y en otras materias) donde la causalidad se discute como una característica directa de la realidad.
3. En la definición de Simon, la causalidad no tiene nada que ver con nuestra interpretación o aplicación del modelo, sino que la ordenación causal que el modelo implica puede inferirse sin ambigüedad a partir de la estructura formal. Se trata de una propiedad formal y *deductiva* del modelo, no de una propiedad *inductiva* que suponga relación alguna con la calidad de las predicciones del modelo. De acuerdo con Zellner (1979), se trata de una definición muy restrictiva que deja fuera una de las características fundamentales del concepto filosófico de causalidad de Feigl (1953): la *calidad de las predicciones*. Sin este ingrediente, un modelo puede ser causal en el sentido de Simon aunque no proporcione predicciones útiles. Sobre estas cuestiones volveremos posteriormente, al tratar los aspectos empíricos de los tests de causalidad.

Muchas contribuciones posteriores sobre causalidad utilizadas en la práctica macroeconómica contemporánea dan la impresión de ser un refinamiento y extensión de las ideas de la Comisión Cowles, particularmente las de Simon. Entre las distintas aportaciones recientes, han sido los trabajos de Pearl (2000) y Heckman (2000) los que han recibido una mayor atención académica, especialmente relacionados con cuestiones de *política económica* del tipo ¿qué pasaría si? (*what if*), que ya aparecían como preocupación fundamental en las actas fundacionales de la Comisión Cowles. La visión de las nuevas formulaciones es que, tras el prometedor comienzo iniciado por Simon, muchos científicos sociales perdieron interés en la idea de los modelos estructurales y, consecuentemente, esto impidió un desarrollo original posterior del concepto de causalidad. Las nuevas aportaciones pretenden llenar este vacío.

El trabajo de Judea Pearl es, probablemente, la aproximación más formal, más completa y más potente de todas las disponibles en la actualidad al permitir la evaluación de causalidades complejas (*counter-*

factuals) y de sus probabilidades. Para ello desarrolla un esquema de redes Bayesianas que utiliza como una herramienta de aprendizaje de relaciones causales a partir de probabilidades condicionadas sobre la base de un conjunto de hipótesis sobre el sistema considerado. Formalmente el *modelo probabilístico causal* de Pearl se puede definir como el conjunto $\{X, Y, F, P(u)\}$, donde

- X es un conjunto de variables determinadas por un conjunto de factores externos al modelo;
- Y es un conjunto de variables determinadas por variables dentro del modelo;
- F es un conjunto de funciones $\{f_1, f_2, \dots, f_n\}$ que forman el conjunto completo de transformaciones de X a Y ;
- $P(u)$ es la función de probabilidad en el dominio de X ⁷.

Para evaluar afirmaciones del tipo ¿cuál sería el valor de Y si $X = x$?, Pearl desarrolla un teorema⁸ que demuestra cómo evaluar la probabilidad condicional $P(B_A | e)$, «si sucediera A entonces B » dada la evidencia e , utilizando tres etapas que denomina *abducción*, *acción* y *predicción* que reflejan una lógica similitud con la aplicación directa del teorema de Bayes. A lo largo del trabajo, Pearl presenta numerosos ejemplos económicos de distinta complejidad, en donde cada ecuación estructural representa una ley causal distinta para cada variable interna. A la hora de analizar una intervención de política económica, Pearl propone eliminar la ecuación que determina una variable interna y fijar el valor de dicha variable a un nivel preasignado (constante), sin alterar las restantes ecuaciones del sistema. Esta propiedad de la aproximación de Pearl, que Reiss y Cartwright (2003) denominan *modularidad*, es difícilmente satisfecha en la mayoría de los modelos económicos desde el momento en que las variaciones en las restantes variables externas pueden afectar a los valores de equilibrio de más de una variable interna. De hecho, resulta bastante difícil encontrar modelos no triviales en cualquier área científica en donde se cumpla la hipótesis de modularidad. La mayoría de los críticos del trabajo de Pearl —p. e., Fennell (2006), Reiss y Cartwright (2003), y Le Roy (2006), entre otros— cuestionan si los sistemas socio-económicos actuales se comportan de acuerdo con el planteamiento de Pearl; y si dicho esquema es útil para evaluar el tipo de políticas económicas a las que estamos acostumbrados los economistas⁹. Al margen de que algunos

7. Para una formulación más precisa, puede verse Pearl (2000, 204 ss.).

8. Teorema 7.1.7, p. 206.

9. Particularmente interesante es el ejemplo que presentan Reiss y Cartwright

autores consideran que «los modelos estructurales son un callejón sin salida a la hora de incorporar directamente el concepto de causalidad» (Le Roy, 2006, 24), la caracterización de Pearl sobre lo que constituye un modelo causal es, probablemente, demasiado estrecha. Es difícilmente compatible con sistemas donde las leyes causales pueden venir afectadas por nuestras propias políticas, tal que, si una determinada ley cambia, es probable que también lo hagan otras leyes ligadas a ella.

Otra alternativa reciente es la presentada por Heckman (2000) donde se postula un modelo causal y se proporciona un conjunto de reglas con la intención de que el modelo sea capaz de proporcionar respuestas concretas a cuestiones del tipo *what if*. En particular, y dado su interés en evaluar determinados programas en mercados laborales, Heckman intenta evaluar la eficacia de un programa concreto, planteando cuestiones del tipo: ¿Qué habría pasado sin la existencia de dicho programa? O, ¿se habrían observado resultados distintos con otro programa alternativo? En este esquema causal, la distinción entre *validez interna* y *validez externa* resulta crucial. Una aseveración causal es válida internamente si es correcta dentro del sistema experimental para la que ha sido diseñada, mientras que su validez externa sólo puede garantizarse si sigue siendo firme fuera del sistema. Nos encontramos aquí, de nuevo, con los dos elementos básicos del esquema causal de Zellner (1979): explicación de los hechos (validez interna) y predicción (validez externa) sobre los que volveremos en la próxima sección.

Heckman considera el concepto de causalidad en un contexto simultáneo de oferta y demanda del tipo:

$$q_s = \beta_{sp} p_s + \beta_{sw} w \quad (4)$$

$$q_d = \beta_{dp} p_d + \beta_{di} i \quad (5)$$

$$\begin{aligned} q_s &= q_d = q \\ p_s &= p_d = p \end{aligned}$$

donde q_s es la cantidad ofertada, q_d es la cantidad demandada, q es la cantidad de equilibrio, p_s es el precio de oferta, p_d es el precio de demanda, i es la renta y w es el tiempo metereológico. Puesto que (4) y (5) no contienen variables comunes, se pueden analizar separadamente. De acuerdo con Heckman, si los precios se fijan fuera del mercado

(2003) sobre sistemas donde las expectativas de los agentes son importantes a la hora de determinar las relaciones entre la variable de control y la variable objetivo. Si las expectativas son importantes, y aparecen en determinadas leyes del sistema, los cambios reales pueden tener poca relevancia. Lo que más importa son las percepciones de los agentes; y en estas situaciones, la semántica de Pearl podría no ser aplicable.

(p. e., un programa gubernamental de precios) es posible variar p_d y p_s y obtener efectos causales en (4) y (5) como derivadas parciales o diferencias finitas de los precios, manteniendo constantes los restantes factores. Al hacerlo, Heckman interpreta que β_{sp} mide el efecto causal de p_s sobre q_s , y que β_{sw} mide el efecto causal de w sobre q_s ¹⁰.

En el ejemplo anterior Heckman trata a p_s y p_d como variables externas a pesar de que p es una variable interna, como consecuencia de las condiciones de equilibrio. Le Roy (2006), sin embargo, señala que esta hipótesis puede llevar a serios problemas de interpretación cuando la normalización se lleva a cabo sobre los precios, y no sobre las cantidades¹¹. Por otra parte, ¿qué ocurriría, por ejemplo, si i y w aparecieran de forma simultánea en (4) y (5)? Es obvio, que en el modelo resultante la ecuación de demanda no podría tratarse como una función causal y, que el efecto del precio sobre la cantidad demandada carecería de significado.

Al margen de estas consideraciones teóricas relacionadas con las hipótesis teóricas concretas, las características formales y los métodos para implementar sus distintas metodologías, ni Pearl ni Heckman (al igual que ocurría con Simon) ofrecen reglas que podamos utilizar para juzgar la validez externa de sus propuestas causales. La impresión metodológica dominante es que carecemos de dichas reglas. La mía personal es que, en determinadas circunstancias, la predicción del modelo puede y debe servirnos como criterio fundamental a la hora de evaluar propuestas causales. A ello dedicaremos la sección siguiente.

3. TESTS EMPÍRICOS DE CAUSALIDAD: LIMITACIONES Y NUEVAS PROPUESTAS

Ya veíamos en la sección anterior cómo la mera asociación entre un par de variables económicas es insuficiente para determinar la causalidad. Hoy día, se admite sin discrepancias que la causalidad implica una relación asimétrica fundamentada en cuestiones relacionadas con la Teoría Económica, la información externa y la prioridad temporal.

10. La interpretación de los coeficientes de la ecuación de demanda es similar. Nótese aquí el contraste con el concepto de Simon, en donde la causalidad se analiza para el sistema completo. Para Simon, el parámetro β_{sp} no tiene causalidad asociada, puesto que p no causa q . Ambas vienen determinadas de forma simultánea.

11. Bajo esta normalización alternativa, ¿Podríamos interpretar el parámetro $\frac{1}{\beta_{sp}}$ como el coeficiente que mide el efecto de q_s sobre p_s ? ¿Sería esta nueva interpretación compatible con la propuesta anterior de Heckman o tendríamos que elegir entre ambas?

La cuestión lógica posterior tiene que venir, necesariamente, relacionada con la posibilidad de llevar a cabo tests estadísticos a partir de los datos disponibles. Para ello necesitamos una definición alternativa y un procedimiento operativo que puedan distinguir entre *causalidad temporal* y *causalidad sincrónica*.

Supongamos un conjunto observado de hogares y un conjunto de variables económicas para cada uno de ellos asociadas, por ejemplo, con el consumo de electricidad en cada hogar. En esta situación, pueden aparecer dos tipos de cuestiones. La primera estaría relacionada con averiguar por qué el consumo de dos hogares A y B es distinto en un momento determinado del tiempo. Se trataría aquí de una cuestión relacionada con la causalidad de sección cruzada o sincrónica. También es posible observar los momentos de la distribución del consumo —como, por ejemplo, la media y la varianza— y observar si estos momentos cambian en el tiempo. Si es así, la pregunta asociada es: ¿por qué la demanda de electricidad es mayor este año que el pasado? En este caso, estaríamos hablando de una causalidad temporal. Puesto que sus definiciones e interpretaciones pueden ser distintas, analizaremos cada concepto separadamente.

3.1. *Causalidad temporal*

Existen numerosos tests de *causalidad temporal*, la mayoría de ellos derivados del concepto de causalidad de Granger (1969). Esta aproximación —que denominamos CG— se basa en dos axiomas: que la causa ocurrirá antes que el efecto (prioridad temporal estricta) y que la causa contiene información única sobre el efecto. Si definimos I_t como toda la información universal acumulada hasta el momento $t-1$ y llamamos $I_t - Y_t$ a toda esa información con excepción de la contenida en Y_t , podemos definir la CG y afirmar que Y_t causa X_t , ($Y_t \Rightarrow X_t$) si se cumple que: $\sigma^2(X | I) < \sigma^2(X | I - Y)$; es decir, que Y_t causa X_t si la predicción de X_t obtenida utilizando el conjunto de información I_t es mejor que la obtenida si ese conjunto de información no incluye a Y_t ¹².

La CG ha recibido numerosas críticas y cualificaciones, incluyendo las del propio autor. La más significativa, sin duda, es el trabajo de Zellner (1979), cuya definición de causalidad basada en Feigl (1953) suponía *capacidad predictiva de acuerdo con una ley o conjunto de le-*

12. Granger (1969, 428-429) también definió otros conceptos relacionados como son los de causalidad instantánea, retardada, y simultánea que no son relevantes para la intención de este trabajo.

yes. De acuerdo con Zellner, la ausencia de teoría económica relevante convierte al criterio CG en una herramienta puramente estadística a la hora de definir la causalidad y representa un caso evidente del conocido eslogan de «medición sin teoría». Por otro lado, la crítica de Zellner se centra en otros aspectos que están directamente relacionados con la aplicación empírica de los tests: la relevancia del conjunto de información I_t , el supuesto de linealidad y el uso de predictores lineales, el criterio cuadrático de mejora predictiva, la exclusión de variables no estocásticas y la utilización de «filtros» para transformar las series. Como corresponde a cualquier polémica entre dos gigantes de la metodología econométrica, la respuesta de Granger (1980) reconocía que su definición contenía algunas ventajas, pero también algunos problemas. En teoría, los tests CG —admitía Granger— no se ven alterados por la aplicación de filtros o transformaciones a los datos, si bien, determinados ajustes estacionales o errores de medida pueden provocar resultados absurdos.

Críticas al margen, la aplicación de los tests CG provocó un aluvión de resultados empíricos en cientos, si no miles, de publicaciones durante los años setenta y parte de los ochenta¹³. Aunque no resulta fácil llevar a cabo una valoración exhaustiva de los resultados, es notable el nivel de indefinición y contradicción de muchos de los trabajos más citados. Así, por ejemplo, mientras que Sims (1972) aceptaba una causalidad unidireccional del dinero a la renta (y rechazaba la causalidad inversa) para datos trimestrales norteamericanos, Feige y Pearce (1976) —utilizando datos similares— no encontraban relación alguna entre oferta monetaria y PIB. Es probable que, en este caso como en muchos otros resultados contradictorios, la crítica de Zellner asociada a problemas de la práctica econométrica elegida —potencia de los tests, imprecisión de las estimaciones, ajustes estacionales y efectos del filtrado de variables— pueda explicar esa miríada de resultados que tan poco ayudan en la búsqueda de un procedimiento válido para la evaluación de políticas económicas. ¿Sigue siendo posible dicha evaluación utilizando modelos econométricos? En el caso de una respuesta afirmativa, aparece un segundo interrogante no menos complejo. ¿Por qué camino? ¿Con qué método? No soy capaz de proporcionar una respuesta contundente sin caer en el grave riesgo del dogmatismo científico. Por ello, sólo intentaré reflejar algunas puntualizaciones que son el fruto de mi propia experiencia y de lo aprendido de las experiencias de otros.

13. Un resumen de los resultados de los tests CG puede verse en Pierce y Haugh (1977), y su evaluación en Nelson y Schwert (1982).

Cualquier respuesta contundente a la pregunta anterior ha de venir mediatizada por las *características de los datos*. En particular, hay dos aspectos claves que considerar: *los efectos a corto y largo plazo*, y el *intervalo muestral* elegido. Es sabido que la agregación temporal puede destruir muchas de las propiedades de las series económicas, llegando incluso a producir resultados inadmisibles desde un punto de vista teórico (p. e., García-Ferrer *et al.*, 2006). En estas situaciones, no es extraño encontrar relaciones unidireccionales de causalidad en series trimestrales que se convierten en simultáneas al pasar a datos anuales. En este sentido, y para el caso de series temporales españolas, también nos hemos encontrado con casos de efectos temporales muy rápidos, especialmente en variables microeconómicas¹⁴. Por último, la inestabilidad de los parámetros estimados —especialmente si utilizamos un periodo muestral amplio— puede echar por tierra los dos pilares básicos del concepto de causalidad: la interpretación teórica de los resultados y la obtención de predicciones eficientes.

3.2. *Causalidad sincrónica*

La validación del concepto de *causalidad sincrónica* presenta otro tipo de peculiaridades relacionadas, fundamentalmente, con el tipo de inferencia causal y con la posibilidad o no de llevar a cabo experimentos controlados. Para un gran número de economistas y metodólogos, la Economía es una ciencia abstracta y deductiva donde sus principios básicos no reflejan situaciones reales sino situaciones «ideales del mundo económico». Esta visión se remonta a los trabajos pioneros de John Stuart Mill (1948/1830) para quien el método inductivo (de evidencias particulares al conocimiento general) era una alternativa inviable en Economía. Sus argumentos, continuamente repetidos en la actualidad, se basaban en que: 1) el mundo económico es demasiado complejo para llevar a cabo una inferencia inductiva y, 2) el control experimental es difícilmente operativo en las ciencias sociales. Por el contrario, en multitud de fenómenos naturales (a pesar de su complejidad) es posible aislar mecanismos causales de forma experimental y llevar a cabo tests de hipótesis sobre la base de un experimento con-

14. Con referencia a este último punto y en relación al efecto del precio sobre la demanda de gasolina en España, García-Ferrer y Del Hoyo (1987) encontraron una respuesta sustancial y significativa dentro de un periodo mensual. Resultados similares encontrados por otros autores en distintos estudios energéticos ponen seriamente en duda la vieja idea de que la demanda energética sólo reacciona a variaciones de precio a largo plazo.

creto. En consecuencia, la Economía sólo puede tratarse como una ciencia deductiva¹⁵.

Aun aceptando las dos premisas básicas de Mills —complejidad e imposibilidad de experimentación— los métodos inductivos han avanzado considerablemente hasta el punto de que, actualmente, es posible llevar a cabo *quasi*-experimentos controlados capaces de permitir una inferencia inductiva causal, incluso en situaciones complejas donde el control experimental no es posible. De esta forma, las nuevas propuestas de los llamados *experimentos naturales* pueden tratarse como un caso particular del método de variables instrumentales (VI) en Econometría, y se han llevado a cabo bajo diversas variaciones en función de las características de los datos¹⁶. En algunos casos, el experimento «natural» contiene dos poblaciones idénticas con respecto a la distribución de las causas posibles, pero una de esas poblaciones recibe el «tratamiento» mientras que la otra no (p. e., Card y Krueger, 1994). En otros, como por ejemplo en Angrist (1990), es posible construir un experimento «aleatorio», ya que tanto la pertenencia al grupo de tratamiento como al grupo de control vienen determinadas por un mecanismo aleatorio.

Veamos dos ejemplos planteados, recientemente, en un famoso y controvertido libro de Levitt y Dubner (2006)¹⁷. El primero de ellos está relacionado con las causas que explican la caída reciente de las tasas de criminalidad en las ciudades norteamericanas. Entre las hipótesis manejadas, el *aumento del número de policías* aparecía como una causa obvia, aunque demostrarlo no resultaba fácil¹⁸. Para demostrar

15. La traslación directa de esta visión a los tiempos actuales es que, en un mundo ideal, un economista podría llevar a cabo un experimento controlado al igual que un físico o un biólogo: preparando dos muestras, manipulando una de ellas al azar, y midiendo el efecto. Pero un economista rara vez puede permitirse el lujo de un experimento puro semejante. Por regla general, sólo contamos con un conjunto de datos con multitud de variables, ninguna de ellas generadas aleatoriamente, unas relacionadas y otras no.

16. Véase, Reiss (2003). El método de VI se utiliza cuando una relación que pretende explicar Y utilizando la información en X, se encuentra potencialmente distorsionada por la influencia de otras variables. El instrumento viene definido por una tercera variable Z con dos características: Z está correlacionada con X, pero no lo está con el término de error de Y.

17. Lo que hace de *Freakonomics* una lectura apasionante no sólo son sus resultados controvertidos y sus retos a la sabiduría convencional, sino la base estadística utilizada y el nivel de las publicaciones científicas en que sus resultados originales fueron publicados.

18. El motivo es que cuando la criminalidad aumenta, la gente clama por protección e, invariablemente, se destina más dinero a la policía. Cuando se observan las correlaciones brutas entre policía y crimen nos encontramos con que cuanto más policía hay, tiende a haber más crímenes.

la causalidad necesitamos un escenario «experimental» en el que se contrate un mayor número de policías por razones que no guarden relación alguna con el aumento del crimen. ¿Cómo podría introducirse esa aleatoriedad necesaria para verificar nuestra hipótesis de causalidad? Levitt y Dubner (2006, 132) lo describen perfectamente en la generación de los propios datos, al afirmar:

Pues resulta que los políticos ávidos de votos crean a menudo el escenario perfecto. En los meses anteriores a las elecciones, los alcaldes tratan de ganar votos contratando más policías, aun cuando el índice de criminalidad permanezca estable. Así que, si comparamos un grupo de ciudades que han tenido elecciones recientemente (y que, por consiguiente, contrataron más efectivos policiales) con otro grupo de ciudades que no las han tenido (y, por lo tanto, no aumentaron sus fuerzas policiales), es posible deducir el efecto adicional en el crimen. La respuesta: en efecto, el aumento del número de policías reduce sustancialmente el índice de criminalidad.

El segundo ejemplo, que tiene que ver con la *evaluación de las políticas educativas* utilizando los datos del Estudio Longitudinal de la Primera Infancia (ECLS según sus siglas en inglés), pretendía evaluar el progreso académico de más de veinte mil niños norteamericanos desde la guardería hasta el quinto curso. Los datos incluían toda la información típica acerca de cada niño: raza, sexo, estructura familiar, posición socioeconómica, nivel de educación de sus padres, etc. También incluían entrevistas con padres y profesores, en donde aparecían todo tipo de preguntas personales que no son comunes en la típica entrevista gubernamental. Supongamos que queremos plantear una pregunta fundamental del tipo: ¿tener multitud de libros en casa implica que su hijo rendirá satisfactoriamente en la escuela? Un modelo de regresión no puede responder satisfactoriamente a esta pregunta, pero puede contestar a una sutilmente distinta: ¿tiende un niño con multitud de libros en su casa a rendir más que un niño sin libros? La diferencia entre ambas preguntas es la diferencia entre causalidad (pregunta 1) y correlación (pregunta 2). Lo que Fryer y Levitt (2004) idearon para resolver este dilema fue alinear a los niños que compartían numerosas características y precisar, posteriormente, la única característica que no compartían. Es decir, lo que realmente hicieron fue comparar a dos niños muy similares en todos los aspectos excepto en uno —en nuestro ejemplo, el número de libros que hay en su casa— y comprobar si ese único factor implica una diferencia en sus resultados escolares. Una base de datos tan amplia como la del ECLS permite este tipo de análisis para la mayoría de las evidencias causales que queramos contrastar.

En estas condiciones, también es posible verificar la capacidad predictiva del modelo llevando a cabo una partición aleatoria muestral, estimando con parte de la muestra y comparando las predicciones con los datos reales observados.

4. CONCLUSIONES

Es posible que los objetivos iniciales asignados a la Econometría —estimación y contrastación de teorías, evaluación de políticas económicas y predicción—, en las actas fundacionales de la Sociedad de Econometría y en los trabajos posteriores de la Comisión Cowles, fuesen demasiado ambiciosos para ser alcanzados de forma simultánea. De hecho, la experiencia de los últimos años nos ha enseñado que es muy difícil elaborar modelos que puedan abarcar de forma satisfactoria estos tres objetivos conjuntamente. Hay muchos economistas —entre los que me incluyo— que creen que deberíamos ser más humildes en el planteamiento de nuestros objetivos, y que sería suficiente con que fuéramos capaces de encontrar modelos útiles.

¿Útiles para qué? ¿Para estimar teorías causales o para producir predicciones eficientes? La definición más general de causalidad supone alcanzar ambos objetivos (teoría y predicción), y —estrictamente hablando— no es posible utilizar otras acepciones del término *causa* sin que éste pierda su significado filosófico. Sin embargo, hemos visto cómo los ejemplos empíricos de inferencia causal más exitosos son aquellos donde la causalidad ha venido mediatizada por los tipos, formas, dominios y niveles de las leyes económicas que intentamos corroborar. Por regla general, cuanto más completo (grande) y mayor sea su nivel de agregación —modelos estructurales con un número considerable de ecuaciones— mayor será la probabilidad de encontrarnos con los problemas que hemos mencionado en las secciones previas de este trabajo. Nunca he encontrado demasiado atractiva la vieja idea de que «el mundo económico es complicado y, en consecuencia, necesitamos modelos complicados». Por el contrario, modelos sencillos que utilizan datos microeconómicos desagregados, referidos a determinados periodos muestrales, y a países o regiones concretas, pueden ser muy útiles, tanto en términos de estimación de teorías como en la verificación de políticas económicas y predicción. Ello no implica, por supuesto, que sus resultados sean directamente extrapolables a otros países y situaciones históricas distintas a las del estudio específico.

BIBLIOGRAFÍA

- Angrist, J. (1990), «Lifetime earnings and the Vietnam era draft lottery»: *American Economic Review*, 80/3, 313-336.
- Card, D. y Krueger, A. (1994), «Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania»: *American Economic Review*, 84/4, 772-793.
- Feige, E. L. y Pearce, D. K. (1976), «Economically rational expectations: Are innovations in the rate of inflation independent of innovations in measures of monetary and fiscal policy?»: *Journal of Political Economy*, 84, 499-552.
- Feigl, H. (1953), «Notes on Causality», en H. Feigl y M. Brodbeck (eds.), *Readings in the philosophy of Science*, Appleton-Century-Crofts, New York, 408-418.
- Fennell, J. (2006), *Causality mechanisms and modularity: Structural models in econometrics*, Mimeo, London School of Economics.
- Fryer, R. G. y Levitt, S. D. (2004), «Understanding the black-white test score gap in the first two years of school»: *The Review of Economics and Statistics*, 86/2, 447-464.
- García-Ferrer, A. y Del Hoyo, J. (1987), «Analysis of the car accident indexes in Spain: A multiple time series approach»: *Journal of Business and Economic Statistics*, 5/1, 27-38.
- García-Ferrer A., De Juan, A. y Poncela, P. (2006), «Forecasting traffic accidents using disaggregated data»: *International Journal of Forecasting*, 22/2, 203-222.
- Granger, C. W. J. (1969), «Investigating causal relations by econometric models and cross-spectral methods»: *Econometrica*, 37, 424-438.
- Granger, C. W. J. (1980), «Tests for causation – a personal viewpoint»: *Journal of Economic Dynamics and Control*, 2, 329-352.
- Hammond, J. D. (1988), «How different are Hicks and Friedman on method?»: *Oxford Economic Papers*, 40, 392-394.
- Heckman, J. (2000), «Causal parameters and policy analysis in economics: A twentieth century retrospective»: *Quarterly Journal of Economics*, 115/1, 45-97.
- Helm, D. (1988), «Reply to Hammond»: *Oxford Economic Papers*, 40, 395-396.
- Hicks, J. (1979), *Causality in Economics*, Basil Blackwell, London.
- Hoover, K. D. (2001), *Causality in Macroeconomics*, Cambridge University Press, Cambridge.
- Le Roy, S. F. (2006), *Causality in Economics*, Mimeo, MS U.C. Santa Barbara.
- Levitt, S. D. y Dubner, S. J. (2006), *Freakonomics*, Ediciones B, Barcelona.
- Mill, J. S. (1948/1830), *Essays on Some Unsettled Questions of Political Economy*, Parker, London.
- Morgan, M. (1991), «The stamping out of process analysis in econometrics», en N. de Marchi y M. Blaug (eds.), *Appraising Economic Theories: Studies in Methodology of Research Programs*, Edward Elgar, Aldershot, 237-265.

- Nelson, C. R. y Schwert, G. W. (1982), «Tests for predictive relationships between time series variables»: *Journal of the American Statistical Association*, 77, 11-18.
- Pearl, J. (2000), *Causation: Models, Reasoning and Inference*, Cambridge University Press, Cambridge.
- Pierce, D. A. y Haugh, L. D. (1977), «Causality in temporal systems»: *Journal of Econometrics*, 5, 265-293.
- Reiss, J. (2003), *Practice ahead of theory: Instrumental variables, natural experiments and inductivism in econometrics*, London School of Economics, CPNSS. Technical Report 12/03.
- Reiss, J. y Cartwright, N. (2003), *Uncertainty in econometrics: Evaluating policy counterfactuals*, London School of Economics, CPNSS. Technical Report 11/03.
- Simon, H. (1953), «Causal ordering and identificability», en W. C. Hood y T. C. Koopmans (eds.), *Studies in Econometric Method*, John Wiley, New York, 49-74.
- Simon, H. (1988), «Causal ordering, comparative statics and near decomposability»: *Journal of Econometrics*, 39, 149-173.
- Sims, C. A. (1972), «Money, income and causality»: *American Economic Review*, 62, 540-552.
- Zellner, A. (1979), «Causality and Econometrics», en K. Brunner y A. H. Meltzer (eds.), *Carnegie-Rochester Conference on Public Policy*, vol. 10, 9-54.