

A. ESPASA



## *Modelos macroeconómicos: Simultaneidad y recursividad, estimación y observaciones escasas\**

---

El criterio que guía el enfoque del presente artículo es básicamente expositivo. En él se analizan algunas ideas acerca de la naturaleza de los modelos simultáneos y de su estimación; el nivel de la exposición presupone un conocimiento básico de Econometría (equiparable, por ejemplo, al del libro «Métodos en Econometría» de J. Johnston (1)) y en la misma, por obvias razones de espacio, no se persigue la demostración de cada uno de los puntos considerados. El objetivo del trabajo es familiarizar al economista interesado en la estimación de un modelo macroeconómico con algunos de los problemas con que puede encontrarse, e introducirlo tanto a las soluciones más recientemente propuestas como a una serie de criterios para elegir entre una gama diversa de estimadores. Para aquellos lectores que no intenten realizar una labor empírica, el artículo puede ser de interés como ayuda a la comprensión de los problemas inherentes en los trabajos de aplicación práctica.

El artículo consta de tres secciones. En la primera se introduce una consideración previa sobre la aplicabilidad a la Economía Española de los métodos de estimación simultánea. La segunda se dedica a presentar los problemas de la simultaneidad y los modelos de tipo recursivo propuestos por Wold y Fisher. La tercera sección se dedica a la estimación de modelos macroeconómicos introduciendo primero los estimadores de clase  $k$ , señalando también una serie de estudios sobre pequeñas muestras, que se enfrentan con la cuestión de cuál es el valor de  $k$  más conveniente. A continuación se exponen los estimadores por variables instrumentales. Por último, sobre el tema de observaciones escasas se presentan dos tipos de soluciones: Una que utiliza la forma reducida no restringida, pero opera con sólo algunos componentes de ella (Klok-Mennes y Fisher), y otra que utiliza la forma reducida restringida.

\* Este trabajo se basa en un curso que impartí en el Departamento de Teoría Económica de la Universidad de Barcelona en diciembre de 1972.

I. UNA CONSIDERACIÓN PREVIA: LA APLICABILIDAD A ESPAÑA  
DE MÉTODOS DE ESTIMACIÓN SIMULTÁNEA

Es interesante, antes de entrar en el tema central del artículo, considerar hasta qué punto las cuestiones aquí tratadas están o no alejadas de las posibilidades de investigación econométrica en nuestro país. La objeción que de forma más inmediata puede plantearse se refiere a la limitación de datos con que el investigador ha de enfrentarse: *las series estadísticas disponibles son tan cortas en observaciones que impiden en principio toda consideración acerca de modelos simultáneos*. La objeción tiene ciertamente su fundamento. Sin embargo, conviene aclarar de entrada que ello no cierra las posibilidades de trabajo econométrico en nuestro país.

El argumento contenido en esta objeción puede ilustrarse con base en un trabajo pionero dentro del campo econométrico español: el modelo contenido en uno de los anexos del III Plan de Desarrollo (2); concretamente la versión definitiva del modelo a largo plazo. Supongamos que se quiere estimar este modelo para el período 1959-1969, considerando que los cambios estructurales en la Economía Española en 1959 fueron tan profundos que no es aconsejable utilizar observaciones anteriores a tal fecha y al mismo tiempo las observaciones posteriores a 1969 se quieren utilizar para los tests de predicción. Supongamos además que se incluye en el modelo variables *dummies* para la intersección de los hiperplanos de regresión. Si bajo tales supuestos los problemas de estimación que la simultaneidad del modelo presenta, los abordáramos con la aplicación de los mínimos cuadrados bietápicos (MCB), nos encontraríamos con que los resultados obtenidos coincidirían con los correspondientes a la aplicación de los estimadores mínimo cuadrático ordinarios (MCO) o directos. La simultaneidad consecuentemente, en dichas circunstancias no se podría tener en cuenta en la estimación. El motivo por el que en dichas circunstancias los MCB coincidirían con los MCO, obedece a que el número de variables predeterminadas que entran en el sistema ( $m$ ) son once; es decir, igual al número de observaciones disponibles ( $T$ ).<sup>1</sup> Por tal razón, la matriz de variables predeterminadas ( $Z$ ) sería cuadrada ( $T \times T$ ) y, por tanto, la condición de los grados de libertad en la estimación mínimo cuadrático bietápica — $m < T$ — no se cumple.<sup>2</sup> Esta circunstancia hace que  $Z^{-1}$  exista<sup>3</sup> y así, en la primera etapa de la estimación, los instrumentos para la matriz

1. El número de observaciones son once, suponiendo que las variables retardadas que entran en el sistema no hacen perder una observación, sino que se toma el valor de 1958 como válido. Si se pierde una observación, la forma reducida empleada en la primera etapa de los MCB es inestimable.

2. Este es el problema que se trata en la sección III, pero aquí es más grande, pues no sólo es un inconveniente en la primera etapa sino también en la segunda.

3. Si no existe no es posible estimar la forma reducida.

de variables endógenas, que aparecen como regresores ( $Y_1$ ) en la ecuación bajo estudio y que vienen dados por

$$[1] \quad \hat{Y}_1 = Z(Z' Z)^{-1} Z' Y_1$$

se reducirán a

$$[2] \quad \hat{Y}_1 = Y_1$$

es decir, los instrumentos empleados en el procedimiento MCB ( $Y_1$ ) serían exactamente las variables endógenas que se pretende suplantar, por lo que volvemos a estar en el caso de MCO. El problema radica en que cuando  $Z$  es ( $T \times T$ ) estamos estimando los coeficientes de la forma reducida ( $F R$ ) no restringida con cero grados de libertad, es decir, no dejamos ninguno para la estimación del componente aleatorio que se quiere sustraer de  $Y_1$  y que evidentemente es inestimable en el ejemplo propuesto, y el hiperplano de regresión contiene todos los puntos de la muestra. Por ello, los MCB no se utilizan en la práctica a no ser que

$$T - m > q > 0$$

siendo  $q$  un número que el investigador considere adecuado. En una intuición ampliamente aceptada en Econometría, cuando el número de observaciones y el grado de sobreidentificación son pequeños, los MCO producen mejores resultados que los MCB. Esta convicción ha sido probada por Kadane (3) para el caso de ecuaciones estáticas.

El ejemplo que hemos expuesto está sometido a ciertos supuestos, pero en la medida en que los modelos macroeconómicos para la economía española estén muy próximos a ellos, la objeción señalada al iniciar este epígrafe se confirmará en los trabajos empíricos. En mi opinión, no obstante, el presente artículo posee —incluso en estas circunstancias— una razón de ser. La simultaneidad en Economía no es un fenómeno al que se le puede buscar una explicación alternativa, sino un fenómeno que necesariamente hay que tener en cuenta. Si la escasez de datos impide la estimación adecuada de un modelo simultáneo, no por ello debe ser ignorada la cuestión, sino que el problema debe ser atacado por la base estadística, procurando, si es posible, la trimestralización de las series y poniendo medios para que en lo sucesivo éstas se calculen con periodicidad trimestral. Por otra parte, los modelos trimestrales son los verdaderamente importantes desde el punto de vista de la planificación y política económicas. En especial, esta última necesita tomar medidas para modificar la marcha de las magnitudes económicas dentro del año y el político actual no se conforma con el dictamen de expertos que se refiera solamente al comportamiento anual de las variables económicas. Asimismo, los modelos trimestrales permiten una mejor especificación de las interrelacio-

nes retardadas entre las variables. Por último, incluso los modelos de crecimiento o a muy largo plazo se construyen ya con base trimestral. Es más, en algunos sectores como el monetario, los modelos trimestrales están siendo sustituidos por los mensuales.

Además de lo dicho, el investigador econométrico puede colaborar a remediar el problema dedicándose al estudio sobre la utilización de modelos con errores en las observaciones, para poder hacer pleno uso de toda la información existente introduciendo en el programa series de fiabilidad escasa. También, y esto es más importante, modelos con observaciones perdidas permitirían combinar series trimestrales con series anuales en un modelo trimestral.

La literatura sobre este último tema es escasa (véase (39) y (40)) y en ella la investigación se ha concentrado en el problema de observaciones perdidas al principio (o al final) del período muestral para un subconjunto de variables endógenas. Esta especialización refleja la necesidad surgida, en países como el Reino Unido, al estudiar ciertos sectores económicos como salarios y empleo, comercio internacional, etc. Estos estudios, además de requerir series procedentes de la Contabilidad Nacional, usan otras publicadas independientemente y que aparecieron con periodicidad trimestral con anterioridad a la Contabilidad Nacional. Así pues, en un modelo para estos sectores, variables como PNB tienen un bloque de observaciones perdidas al principio del período.

Sargan y Drettakis estudian el problema dentro de un modelo simultáneo autorregresivo y el procedimiento empleado consiste en estimar por el método de la máxima verosimilitud los parámetros del sistema para el subperíodo en el que todas las series están completas. Con dichos resultados y la información estadística para el subperíodo inicial se estiman los datos perdidos que se usan posteriormente para reestimar el modelo para todo el período. Con estos nuevos estimadores se vuelven a calcular las observaciones perdidas y así sucesivamente hasta alcanzar el *maxima maximorum* de la función de verosimilitud.

En Sargan y Drettakis (39) y Drettakis (40) no se usa la información anual sobre las series con observaciones perdidas, pues esto complica considerablemente el problema aunque no necesariamente lo convierte en insoluble. Pero si este tipo de solución se aplicase a la economía española sería importante, quizás imperativo, investigar cómo utilizar dicha información anual. En cualquier caso este tipo de solución exige una experiencia, por corta que sea, de observaciones trimestrales. Si esto no es así y en el modelo entra una serie con observaciones sólo anuales a lo largo de todo el período, habría que buscar si existe otra variable, suficientemente correlacionada con ella, con observaciones trimestrales, que se pudiese utilizar para interpolar la primera. No obstante, quedará todavía el problema de la variación estacional que se puede atacar operando con series ajustadas.

En realidad, el tratamiento de la base estadística y el estudio de modelos apropiados a nuestra situación es un paso necesario, sin el cual los estudios

econométricos empíricos sobre nuestra economía tendrán que ser necesariamente deficientes y por tanto poco fiables para basar una predicción y simulación económica en ellos.

## II. SIMULTANEIDAD Y RECURSIVIDAD

Hay que advertir que a lo largo de estas notas consideraremos simplemente sistemas lineales. Ello no es debido a que tales sistemas supongan una representación adecuada de la realidad económica —al contrario, un «buen» modelo económico esencialmente tiene que contener elementos no lineales—, sino porque con la extracomplejidad de la no linealidad, el tratamiento de los problemas que nos proponemos quedaría algo oscurecido y no permitiría captar fácilmente otros aspectos fundamentales de los programas económicos. No obstante, para la estimación por el método MCO o por variables instrumentales, que no requieran operar con la forma reducida, es suficiente que el modelo sea lineal en los parámetros, pudiendo no serlo en las variables. Para la estimación en este caso por MCB se remite al lector a Kelejian (7) y en general para los problemas sobre la no linealidad en Econometría una referencia muy útil en Goldfeld y Quandt (8).

Una de las principales características de los modelos económicos que los diferencia de los físicos es la determinación conjunta de distintos outputs en el modelo; es decir, los outputs son al mismo tiempo inputs. Este fenómeno es conocido en Econometría con el nombre de simultaneidad y constituye un problema que ha preocupado a los investigadores desde los comienzos de la Econometría moderna. La simultaneidad fue tratada inicialmente por Haavelmo (4). Con posterioridad, en la Cowles Commission se enfocaron los problemas de la identificación y estimación de dichos modelos y sus resultados se publicaron en las monografías diez a catorce (5, 6). Desde entonces la investigación sobre los referidos programas ha ocupado la mayor parte de la literatura econométrica.

Es bien conocido (véase Goldberger (9) capítulo siete) que si se pretende estimar un sistema simultáneo por MCO obtenemos estimadores inconsistentes e ineficientes de los parámetros; tal circunstancia ha llevado a muchos estudiosos a proponer otros estimadores que, al menos asintóticamente, tuvieran las deseadas propiedades de consistencia y eficiencia. Dichas propiedades las tenían los estimadores de la Máxima Verosimilitud, de la Cowles Commission, aunque los mismos ofrecían el inconveniente de que las ecuaciones normales en las que se basan, son altamente no lineales y los algoritmos conocidos en aquella época para su resolución exigían mucho tiempo en un ordenador electrónico (por otra parte no tan perfeccionados como los de nuestros días). Ello indujo a que Theil y Bassman, separadamente, dedujeran los MCB (10, 11, 12) y posteriormente Theil y Zellner, los estimadores mínimo cuadráticos trietápicos (MCT) (13). Ambos estimadores, bajo ciertas hipótesis bastante

generales, tienen respectivamente las mismas propiedades asintóticas que los estimadores de la máxima verosimilitud de información limitada (MVIL) y de información completa (MVIC) de la Cowles Commission y a su vez son más fáciles de obtener. Finalmente, debido a Theil y Nagar, dentro de la clase de estimadores de información limitada, se han propuesto los estimadores de clase  $k$ , clase  $b$  y doble clase  $k$  (11 y 15).

Otro enfoque en la investigación de programas econométricos, debido a Herman Wold (14), es el de especificar el modelo de forma recursiva, de modo que los MCO se puedan aplicar consistentemente. La idea básica reside en que, después de normalizar el sistema, adjudicando a cada ecuación una variable endógena como variable dependiente, se pueden ordenar las ecuaciones de modo que en la  $i$ -ésima ecuación, explicando la  $i$ -ésima variable, las variables endógenas que entran en ella como regresores son de orden menor a  $i$ , es decir, han sido explicadas previamente. Sin embargo, a pesar de la lógica de los modelos recursivos, el problema radica en que en Economía no podemos observar las variables que entran en el sistema (o al menos no se dispone de estadísticas) a intervalos de tiempo tan pequeños como sea necesario para hacer el programa recursivo. Así, los modelos estándar actuales son anuales o trimestrales y en ellos la hipótesis de recursividad expuesta arriba no es una hipótesis realista. Conviene señalar también que, incluso en un sistema estático (es decir, sin *retardos endógenos* que afecten al output), para que los estimadores MCO sean consistentes no basta con que el modelo sea recursivo respecto a las variables endógenas, sino que además las perturbaciones que entran en un momento dado en el sistema sean independientes unas de otras, condición que ciertamente no se cumple en un modelo económico. Ésta puede ser también muchas veces una característica que diferencia y complica la estimación de un modelo económico respecto a uno físico; los modelos físicos de control, por ejemplo, asumen generalmente independencia en las diversas perturbaciones aleatorias que afectan el sistema en un momento determinado de tiempo ( $t$ ).

De lo dicho podemos concluir que en la práctica no es posible obtener un *modelo económico recursivo* al que la aplicación de los MCO sea adecuada. Ahora bien, si tal tipo de recursividad no es una hipótesis aceptable, la idea de modelos recursivos en bloque de F. M. Fisher (16) puede serlo más fácilmente. Estos sistemas propuestos por Fisher son una generalización de los de Wold y la idea consiste en considerar al modelo subdividido en bloques correspondientes a diferentes conjuntos de variables endógenas contemporáneas con sus respectivas ecuaciones en vez de considerar variables endógenas individualmente. La ventaja de tal descomposición del sistema es que entonces la hipótesis de que en el  $i$ -ésimo bloque no entran variables endógenas que se explican en bloques de orden mayor a  $i$  es aceptable y la hipótesis de que las perturbaciones de distintos bloques sean independientes es mucho menos restringente que la correspondiente a la propuesta de Wold. Con ello se ha dividido el sistema en bloques recursivos, que a su vez son, aisladamente consi-

derados, simultáneos respecto a las variables endógenas contemporáneas que en ellos se definen. Asumiendo que no hay correlación serial en las perturbaciones, lo expuesto es suficiente para conseguir que las variables endógenas de cualquier bloque no estén correlacionadas con las perturbaciones aleatorias de bloques de mayor orden. Desgraciadamente, en Economía la ausencia de correlación serial (en contra de lo que frecuentemente se presupone), no es una hipótesis aceptable. No obstante, introduciendo la correlación serial en el programa, bajo ciertas hipótesis sobre la matriz de retardos endógenos y la de covariancias, un modelo del tipo del de Fisher todavía tiene la mencionada propiedad. Como veremos posteriormente, esta generalización de la idea de recursividad es todavía una hipótesis incorrecta para un modelo económico (como el mismo Fisher demuestra en el artículo referido); no obstante, es mucho más general que la idea de Wold y puede ser aceptada con más facilidad como hipótesis de trabajo en una investigación empírica.

Es conveniente expresar analíticamente lo hasta ahora dicho. Supongamos que el modelo en cuestión en tiempo  $t$  puede ser representado de la siguiente forma:

$$[3] \quad y_t = Ay_t + By_{t-1} + Cz_t + u_t$$

donde  $y_t$  e  $y_{t-1}$  son respectivamente dos vectores ( $n \times 1$ ) de variables endógenas contemporáneas y con un retardo;  $z_t$  es un vector ( $m \times 1$ ) de variables exógenas y  $u_t$  es un vector ( $n \times 1$ ) de perturbaciones aleatorias. También suponemos que [3] cumple las hipótesis de costumbre sobre  $z$  y  $u_t$ .  $A$ ,  $B$ ,  $C$ , son matrices ( $n \times n$ ), ( $n \times n$ ) y ( $n \times m$ ) de parámetros que tienen que ser estimados.

$A \neq \emptyset$  donde  $\emptyset$  es la matriz en que todos sus elementos son ceros y, por tanto, el sistema es simultáneo de forma directa. Además suponemos que el modelo ha sido normalizado de forma que todos los elementos en la diagonal principal de  $A$  son ceros; es decir, en cada ecuación la variable endógena que aparece a la izquierda del signo igual no se repite en el otro lado de la expresión. Adviértase también que para que el conjunto de ecuaciones propuestas en [3] sea aceptable, es necesario que  $(I - A)$  —donde  $I$  es la matriz unidad— sea una matriz no singular, pues de lo contrario el modelo no representaría una teoría de formación conjunta de las variables dependientes; en cierta manera, una garantía de que  $(I - A)^{-1}$  existe nos viene dada por el mismo contenido de la Teoría Económica acerca del carácter interdependiente con que las variables se relacionan entre sí.

El modelo es dinámico pues  $B \neq \emptyset$ . Obsérvese que si en el modelo entran variables endógenas con más de un retardo, la expresión en [3] es válida tras una redefinición de las variables endógenas. En efecto, supongamos que en el sistema intervienen hasta  $r$  retardos en la determinación de  $y_t$  de forma que la representación en tiempo  $t$  es

$$[4] \quad y_t^* = A^*y_t^* + B_1^*y_{t-1}^* + B_2^*y_{t-2}^* + \dots + B_r^*y_{t-r}^* + Cz_t^* + u_t$$

fácilmente se comprueba que definiendo:

$$y_t = \begin{bmatrix} Y_t^* \\ Y_{t-1}^* \\ \vdots \\ Y_{t-r+1}^* \end{bmatrix}$$

$$A = \begin{bmatrix} A^* & 0 & 0 & \dots & 0 \\ 0 & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} B_1^* & B_2^* & \dots & \dots & B_r^* \\ I & 0 & 0 & \dots & 0 \\ 0 & I & 0 & & \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & 0 & I & 0 \end{bmatrix}$$

[3] es una válida representación de [4]. En lo que sigue tomaremos [3] como una expresión bastante generalizada y supondremos, sin entrar aquí en su discusión, que el sistema es estable.

En un modelo estático, es decir,  $B = 0$ , las condiciones de recursividad de Wold bajo las cuales los MCO son consistentes son:

- W-1.  $A$  es triangular con ceros arriba de la diagonal principal y  
 W-2. la matriz de variancias de  $u_t$  ( $V_0$ ) es diagonal.<sup>4</sup>

Dado que estamos bajo el supuesto de un modelo estático las perturbaciones pueden estar sometidas a correlación en el tiempo y los estimadores MCO ser todavía consistentes, aunque no eficientes. Pero en la realidad  $B$  nunca es 0 y, por tanto, se necesita una tercera condición para la consistencia de MCO. Para su formulación definamos antes la matriz  $V(\theta)$  como la matriz de covariancias entre los vecto-

4. Esta forma de exposición es la que se encuentra en Fisher (16) en cuyo artículo está basada principalmente la presente discusión de recursividad y recursividad en bloque.



res  $u_t$  y  $u_{t-s}$ . Como se ha dejado implícito arriba al establecer las propiedades de [3], estamos suponiendo que cuando  $V(\theta)$  no es cero, es independiente de  $t$ ; es decir, el proceso aleatorio  $u$  es estacionario; sin la hipótesis de estacionariedad no es posible estimar un sistema disponiendo de una sola observación por cada unidad de tiempo. Con esta aclaración la condición se puede expresar;

- W-3.  $B$  es triangular como  $A$ ,  
 $V(\theta)$  es triangular para cualquier  $\theta \geq 0$  y  $B$  o todas las matrices  $V(\theta)$  ( $\theta > 0$ ) tienen ceros en todos los elementos de la diagonal principal.

Es muy probable que ninguna de las tres condiciones W se cumpla para un modelo macroeconómico. Con todo, el problema no radica tanto en su exacto cumplimiento como en el hecho de que las mismas se aproximen en cierto modo a las circunstancias que la realidad presenta. Cabe pues preguntarse, ¿constituyen por lo menos una aproximación aceptable? Las contribuciones de Fisher, y la práctica que ya hoy en día existe sobre este tipo de modelos, nos llevan a aceptar la opinión de que W-1, W-2 y W-3 no son aceptables ni tan siquiera como algo aproximado.

Las condiciones de recursividad de Fisher se pueden expresar:

- F-1.  $A$  es triangular por bloques  
 F-2.  $V(\theta)$  es diagonal por bloques  
 F-3.  $B$  es triangular por bloques,  
 $V(\theta)$  es triangular por bloques para cualquier  $\theta > 0$  y o bien todos los bloques en la diagonal de  $B$  son cero, o lo son todos los bloques en las diagonales de todas las matrices  $V(\theta)$ , para  $\theta > 0$ .

F-1 es una aproximación aceptable, tanto más cuanto más se desagregue en el tiempo. Sin embargo, F-2 y F-3 no lo son. Las perturbaciones aleatorias que entran en el modelo se componen principalmente de los efectos de las variables omitidas y algunas de ellas pueden influir en todos los bloques aunque ciertamente con intensidades distintas; en todo caso, es muy difícil pretender que tales efectos sean independientes. Con base al mismo razonamiento acerca de errores en las ecuaciones compuestas por la acción de las variables omitidas, cuanto más se desagregue en el tiempo el sistema —para que F-1 sea aproximadamente cierta— más parecidos o relacionados estarán los vectores  $u_t$  y  $u_{t-1}$  y, por tanto, más fuerte será la correlación serial. Por supuesto, al desagregar en el tiempo los parámetros del proceso  $u$  cambiarán considerablemente (véase por ejemplo Thomas y Wallis) (18). Un proceso autorregresivo de primer orden en un modelo anual será sustituido fácilmente por un proceso autorregresivo de quinto grado con parámetros distintos de cero para los retardos primero, cuarto y quinto, en un modelo trimestral. La alternativa a la hipótesis de ausencia de correlación serial es F-3, pero en ella se asume que  $B$  es triangular y, por tanto, también lo es  $(I - A)^{-1}B$ ; es decir,

el sistema de ecuaciones en diferencias finitas [3] se puede descomponer de forma que «lo que ocurre en sectores de orden superior nunca influye en lo que ocurre en los sectores de orden inferior»<sup>5</sup> y esto, como el lector puede suponer, es irreal. Así pues, F-2 y F-3 no se dan en la práctica ni siquiera de forma aproximada; en consecuencia, el uso de variables endógenas contemporáneas como instrumentos en sectores de orden superior puede ocasionar sesgos considerables, es decir, dichas variables endógenas necesitan ser reemplazadas por otros instrumentos en la estimación. Fisher demuestra que bajo ciertas condiciones bastante generales y aceptables en las perturbaciones aleatorias, la hipótesis de que los retardos endógenos no están correlacionados con las perturbaciones aleatorias contemporáneas de los bloques de mayor orden se cumple con bastante aproximación, si bien tales retardos pueden estar correlacionados con las perturbaciones correspondientes a bloques de menor o igual orden. En consecuencia, dichas variables se pueden usar como instrumento en aquellos sectores, mientras que para usarlos en los últimos con igual nivel de consistencia Fisher propone retardarlos un período más. Este resultado, que puede parecer muy débil, es muy importante para el investigador empírico en este campo.

El trabajo de Fisher se basa en un estudio de los estimadores principalmente desde el punto de vista de su primer momento y obtiene unas recomendaciones para la estimación por el procedimiento de variables instrumentales, que es el más usado en macromodelos amplios. Tal método es de información limitada, es decir, no incorpora los conocimientos *a priori* sobre las demás ecuaciones<sup>6</sup> ni tiene en cuenta la correlación entre las perturbaciones. Al no introducir estos aspectos, fácilmente se intuye que dichos estimadores serán menos eficientes que aquellos que hagan uso de ella, es decir, los de información completa. Todo esto es válido bajo el supuesto de que la información extra es correcta y sirve para mejorar y no empeorar los resultados (punto éste muy debatido en el que no entramos). Pero sí que se desprende de lo dicho que un estudio basado en los segundos momentos de los estimadores se tendría que referir a los métodos de información completa. Ahora bien, cuando en la práctica se realiza la estimación simultánea de un modelo (véase (19)), las hipótesis F-2 y «no correlación serial» se toman necesariamente como ciertas debido a la alta escasez de observaciones con que se trabaja. D. Hendry (20) ha desarrollado y aplicado el estimador propuesto por Sargan (21) MVIC con un proceso autorregresivo en los errores. Para modelos pequeños éste es el procedimiento recomendable siempre que los datos lo permitan. Para modelos amplios el número de observaciones requerido impedirá por completo su uso. Una posible solución sería descomponer el modelo en

5. FISHER (16), p. 609.

6. El procedimiento de variables instrumentales es el que generalmente no tiene en cuenta la información contenida en las otras ecuaciones, pero el procedimiento que Fisher propone en la segunda parte de su artículo y que se comenta en la sección III de estas notas, sí que la tiene presente.

pequeños bloques recursivos y aplicar en cada uno de ellos un procedimiento similar al realizado por Hendry. Otro modo de prescindir de la hipótesis de ausencia de correlación serial es aplicar el método de Hannan (36) (capítulo VII), véase también (37), a un modelo simultáneo (labor que no ha realizado todavía nadie hasta la fecha) y estimarlo utilizando la información espectral en vez de los segundos momentos. Éste es actualmente el tema de investigación del que escribe estas líneas.

Poco más añadiremos sobre estimación simultánea en el resto de estas notas que se dedicarán, por razones mencionadas, a las técnicas de información limitada. Veremos que incluso en ellas la escasez de observaciones es todavía problema y discutiremos algunas de las soluciones propuestas.

### III. LA ESTIMACIÓN DE MODELOS MACROECONOMÉTRICOS

Como se ha señalado en la Introducción, para la estimación por MCB y en general por cualquier método del tipo clase  $k$ , se necesita estimar en una primera etapa la forma reducida no restringida. Si el modelo es suficientemente amplio surge el problema de que el número de variables predeterminadas ( $m$ ) es superior al de observaciones y, por tanto, dicha FR no puede ser estimada de forma única y no se pueden obtener sin modificación estimadores estructurales únicos de clase  $k$ .

#### 1. Estimadores de clase $k$

Consideremos una ecuación estructural.

$$[5] \quad y_1 = y_1 b_1 + Z_1 c_1 + u_1 \quad ? = X_1 a_1 + u_1$$

Los estimadores de clase  $k$  consisten en estimar [5] por variables instrumentales (véase fórmula [11]) utilizando como instrumentos para  $Y_1$ ,  $\hat{Y}_1 - KV_1$  donde  $\hat{Y}_1$  y  $\hat{V}_1$  son los valores estimados a través de la forma reducida de  $Y_1$  y de los errores. La fórmula para tales estimadores es:

$$[6] \quad \begin{pmatrix} \hat{b}_1 \\ \hat{c}_1 \end{pmatrix} = \begin{pmatrix} Y_1' Y_1 - K \hat{V}_1' \hat{V}_1 & Y_1' Z_1 \\ Z_1' Y_1 & Z_1' Z_1 \end{pmatrix}^{-1} \begin{pmatrix} Y_1' - K \hat{V}_1' \\ Z_1' \end{pmatrix} y_1 = \hat{\delta}(K)$$

donde  $\hat{V}_1$  es la matriz estimada de los errores de la FR correspondiente a  $Y_1$ .

7. En esta formulación la matriz  $Z$  puede contener variables endógenas retardadas además de variables exógenas. El sufijo 1 sirve para indicar las variables que entran en la ecuación, mientras pueden haber otras en el sistema no incluidos en ella.

Fácilmente se ve que si  $k = 0$  tenemos las fórmulas para los MCO y si  $k = 1$  las correspondientes a los MCB. En tal caso, [6] se puede expresar de la siguiente manera

$$[7] \quad \begin{pmatrix} \hat{b}_1 \\ \hat{c}_1 \end{pmatrix} = \begin{pmatrix} Y_1'Z(Z'Z)^{-1}Z'Y_1 & Y_1'Z(Z'Z)^{-1}Z'Z_1 \\ Z_1'Z(Z'Z)^{-1}Z'Y_1 & Z_1'Z(Z'Z)^{-1}Z'Z_1 \end{pmatrix}^{-1} \begin{pmatrix} Y_1'Z(Z'Z)^{-1}Z'y_1 \\ Z_1'Z(Z'Z)^{-1}Z'y_1 \end{pmatrix}$$

o utilizando  $Z = (Z_1Z_2)$ ,  $Z_1 = Z \begin{pmatrix} I \\ 0 \end{pmatrix}$  Y

$$Z(Z'Z)^{-1}Z'Z_1 = Z(Z'Z)^{-1}Z'Z \begin{pmatrix} I \\ 0 \end{pmatrix} = Z_1$$

$$[8] \quad \begin{pmatrix} \hat{b}_1 \\ \hat{c}_1 \end{pmatrix} = \begin{pmatrix} Y_1'Z(Z'Z)^{-1}Z'Y_1 & Y_1'Z \\ Z_1'Y_1 & Z_1'Z_1 \end{pmatrix}^{-1} \begin{pmatrix} Y_1'Z(Z'Z)^{-1}Z'y_1 \\ Z_1'y_1 \end{pmatrix} = \delta \quad (1)$$

o bien

$$[8a] \quad \delta \quad (1) = (X_1'Z(Z'Z)^{-1}Z'X_1)^{-1} (X_1'Z(Z'Z)^{-1}Z'y_1)$$

fácilmente se ve en [8] y también en [6] que si

$$m > T$$

$(Z'Z)$  que es una matriz  $(m \times m)$ , tiene a lo más rango  $T$  y, por tanto, es singular y  $(Z'Z)^{-1}$  no existe y los estimadores expresados en [6] no se pueden obtener.

## 2. Estimadores MVIL

El mismo problema ocurre con MVIL, pues éstos son un caso particular de [6] en los que  $k$  es estocástico y viene dado por el valor de la raíz característica más pequeña ( $l_1$ ) de la ecuación determinantal

$$[9] \quad \det (W^{**} - 1W^*) = 0$$

donde las matrices  $W$  consisten en sumas de cuadrados y productos transversales de los residuos en la estimación mínimo cuadrática de la FR de las variables endógenas ( $y_1 Y_1$ ). Se incluye  $y_1$  en [9] pues MVIL es independiente de la variable que se normalice ya que la estandarización empleada en ellos es

$$b^{*'} W^* b^* = 1$$

donde  $b^*$  es el vector de coeficientes de la ecuación en cuestión sin normalizar. En [9] la matriz  $W^*$  proviene de la estimación hecha con todas las variables predeterminadas y la matriz  $W^{**}$  con sólo las predeterminadas que aparecen en la ecuación estructural.

Otra forma de ver los estimadores MVIL es como el vector característico normalizado asociado con  $l_1$ .

Antes de considerar el caso  $m > T$  puede ser interesante comentar qué valor de  $k$  en [6] es conveniente elegir; desgraciadamente, veremos que no es muy definitivo lo que se puede decir sobre ello.

### 3. Elección de $k$ cuando $(Z'Z)^{-1}$ existe

Comencemos por observar que para valores de  $k$  correspondientes a las raíces características de

$$[10] \quad \det \begin{pmatrix} Y_1'Y_1 - KV_1'\hat{V}_1 & Y_1'Z_1 \\ Z_1'Y_1 & Z_1'Z_1 \end{pmatrix}$$

los estimadores propuestos en [6] obviamente son explosivos. De ahí que el valor (la raíz característica menor de (10)) sea, como han señalado Savin y Farenbrother (22), un estadístico de interés.

Los MCO ( $k = 0$ ) tienen la propiedad de minimizar la variancia de los residuos, pero el inconveniente de ser inconsistentes. Para la consistencia de los estimadores definidos en [6] se requiere que <sup>8</sup>

$$[11] \quad \text{Plim } k = 1 \text{ o } \lim k = 1.$$

Según sea  $k$  estocástica o no.

MCB es, pues, consistente y también MVIL, ya que

$$\text{Plim } l_1 = 1$$

El problema de elegir el valor de  $k$  no se presenta asintóticamente, puesto que si, como es deseable,  $k$  cumple [11] en el límite, cualquier estimador consistente coincidirá con MCB. El problema de este epígrafe es un problema de pequeñas muestras, que por otra parte son la realidad cotidiana. Esta cuestión se podría resolver derivando la distribución exacta de  $\delta$  para pequeñas muestras y elegir aquel valor de  $k$  que minimizase la suma de los cuadrados de los errores de los estimadores respecto el verdadero valor de  $\delta$ . Desgraciadamente hasta la fecha tal distribución no ha sido derivada por ningún autor,

8. Ver H. THEIL (43), pp. 504-506.

pues envuelve una complicación extraordinaria, aunque existen trabajos para el caso de MCB y MVIL con sólo un regresor endógeno en la ecuación [véase (24), (25) y (26)], en un sistema estático. Para  $\delta(k)$  el máximo resultado obtenido es el debido a Kadiyala (23) que demuestra que la suma de los cuadrados de los residuos es una función monótona creciente de  $k$  para  $0 \leq k \leq \lambda$  por lo que él apunta que MCB son preferibles a MVIL. Este criterio, aunque muy ilustrativo, no es el único o definitivo: Kadane (3) ha demostrado que para grados de sobreidentificación pequeños ( $\leq 6$ ) la esperanza matemática de los cuadrados de los errores (respecto al verdadero valor) de cualquier combinación lineal de los parámetros estimados por MCB es menor que la correspondiente a MVIL. No obstante, para sistemas amplios, lo contrario es posiblemente cierto; no es claro pues en qué ocasiones los MCB son mejores que MVIL, pues, como advierte Kadane, los estudios de Monte Carlo que concluyen prefiriendo MCB a MVIL se basan en pequeños modelos y los resultados pueden ser muy distintos para modelos más amplios. La sugerencia de Kadiyala es tomar como estimador una media ponderada de MCO, MCB y MVIL, pero no es intuitivo qué ponderación señalar a cada uno de ellos. Además Farebrother y Savin han demostrado que si está lejos de  $I_1$  el gráfico de  $\delta(K)$  en función de  $k$  es bastante plano y por tanto no se gana nada por la ponderación de Kadiyala. Las primeras sugerencias en este campo fueron debidas a Nagar (27) quien propuso un valor

$$K = 1 + \frac{\alpha}{T}$$

donde  $\alpha$  es un número real independiente de  $T$ ; tal estimador es consistente y él deduce el valor de  $\alpha$  que minimiza la variancia. Tal valor en general resulta menor que uno aunque aproxime a uno; para casos en que  $m$  es muy grande  $\alpha$  posiblemente es superior a uno.

El inconveniente de todos los trabajos sobre las propiedades de los estimadores en pequeñas muestras es que se basan en modelos que no contienen retardos endógenos y éste no es el tipo de programa con el que se opera en la realidad. No obstante, por lo trabajos mencionados y por muchos otros en este campo, puede decirse que bajo «ciertas condiciones» y en «cierto sentido» los MCB son tan buenos o quizá mejores que los MVIL y más fáciles de obtener. Por ello, la elección del valor uno para  $k$  es la más popular, pero frecuentemente todavía existe el problema de qué instrumentos emplear en la primera etapa y a ese problema volveremos nuestra atención en el epígrafe 4.

Quizás en el resto de este artículo será muy útil considerar a los MCB como un tipo de estimadores por variables instrumentales más que como un caso de [6].

Estimadores por variables instrumentales <sup>9</sup>

Fijémosnos en la ecuación [5] y supongamos que tenemos exactamente  $n_1 + m_1$  instrumentos para sustituir a  $X_1$  y llamemos  $P$  a la matriz de observaciones de esos instrumentos: premultiplicando [5] por  $P'/T$  tenemos

$$[11a] \quad \frac{P'y_1}{T} = \frac{P'X_1}{T} a_1 + \frac{P'u_1}{T}$$

y como, por construcción,

$$p \lim \frac{P'u_1}{T} = 0$$

podemos sin perder consistencia prescindir de tal información a medida que  $T$  tiende a infinito en [11a] y estimar  $a_1$  por la fórmula

$$[11b] \quad \hat{Q}_1 = (P'X_1)^{-1}P'y_1$$

En general, el número de variables instrumentales (VI) es mayor que  $n_1 + m_1$  y se nos presenta el problema de qué instrumentos elegir. La solución es buscar la combinación lineal óptima de las VI a nuestra disposición. Es decir, buscar una matriz  $H$  tal que, aplicada a  $P_1$  nos dé  $n_1 + m_1$  instrumentos óptimos,  $HP_1$ . Pero  $H$  no se puede determinar de forma única, por lo que hay que imponerle alguna restricción y usando

$$H'K = I$$

donde  $K$  es

$$K = p \lim \left( \frac{P'X_1}{T} \right)$$

el estimador óptimo por VI es

$$[11c] \quad \hat{a}_1 = (X_1'P(P'P)^{-1}P'X_1)^{-1} (X_1'P(P'P)^{-1}P'y_1)$$

Comparando [11c] con [8a] vemos que los MCB son un caso particular

9. Esta sección ha sido escrita basándose en el artículo del profesor J. D. SARGAN (38) y en su curso de «Teoría de la Econometría» en la London School of Economics, 1970-71.

de la clase de estimadores por VI en el que  $P$  es la matriz de todas las variables predeterminadas del sistema ( $Z$ ). O comparando [11b] y [8a] vemos que

$$P' = \begin{pmatrix} Y'_1 \\ Z'_1 \end{pmatrix} (Z'Z)^{-1} Z' = \begin{pmatrix} \hat{\pi} Z' \\ Z'_1 \end{pmatrix} = \begin{pmatrix} \hat{Y}'_1 \\ Z'_1 \end{pmatrix}$$

es decir, MCB equivale a tomar las variables exógenas que entran en [5] como instrumentos de sí mismas y a los valores estimados a través de la forma reducida para las variables endógenas incluidas en [5] como instrumentos de  $Y_1$ . Obsérvese que empleamos la palabra instrumentos en dos circunstancias distintas. Para referirnos a los instrumentos usados en la primera etapa,  $Z$ , o para referirnos a los utilizados en la segunda ( $\hat{Y}_1 Z_1$ ). Este uso simultáneo se continúa haciendo a lo largo del artículo sin especificar a veces si nos referimos a la primera o a la segunda etapa, pero por el contexto queda claro cuál de ellas es la envuelta.

Se puede demostrar que si todos los instrumentos posibles entran en el sistema, MCB es eficiente dentro de la clase de VI, no ganando nada considerando retardos de las variables ya incluidas en  $Z$ . Si existen instrumentos adecuados fuera del modelo es indicio de que éste contiene errores en la especificación, y antes de pasar a estimarlo debemos mejorar aquélla.

#### 4. *El uso de componentes principales en la primera etapa*<sup>10</sup> de MCB

Volviendo al caso, en [8] no es aplicable, pues  $m > T$ . Este problema no es en absoluto nuevo y una de las primeras soluciones propuestas fue la debida a Kloek y Mennes (28) que consiste en usar como instrumentos para los regresores endógenos su valor estimado en una regresión sobre componentes principales de las variables predeterminadas y no sobre todo el conjunto. Con ello, utilizando un número de componentes principales compatible con las observaciones disponibles, los MCB son factibles. Como es bien sabido (véase (29) secciones 2.2 y 5.6) la técnica de los componentes principales consiste en que dado un conjunto de variables se debe encontrar un número de combinaciones lineales de ellas mutuamente independientes y que en cierto sentido tengan máxima variabilidad. Si  $Z$  es la matriz de observaciones de variables predeterminadas, la correspondiente matriz de  $r$  componentes principales viene dada por

$$[12] \quad C_r = Z A_r$$

10. La terminología de la primera y de la segunda etapa es ambigua, pues para obtener MCB no hay que operar en la FR en una primera etapa y con dichos resultados pasar a la segunda, sino que [8] permite el cálculo directo e igualmente [6] se puede expresar de forma donde el cálculo es directo. Si aquí se mantienen a veces los términos como primera etapa es con la esperanza de que de esa forma la expresión sea más clara.



donde

$$[13] \quad A_r = (a_1, \dots, a_r)$$

donde  $a_i$  es el vector característico normalizado correspondiente a la  $i$ ésima raíz característica de  $Z'Z/T$ .

Teniendo en cuenta que

$$[14] \quad \left( \frac{Z'Z}{T} \right)^{-1} = A \Lambda^{-1} A' \quad 11$$

donde

$$[15] \quad \Lambda_m = \begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \lambda_m \end{bmatrix}$$

siendo  $\lambda_i$  la  $i$ ésima raíz característica de  $Z'Z/T$ , [14] puede ser aproximado mediante

$$A_r \Lambda_r^{-1} A_r'$$

y sustituyendo en [8] tenemos:

$$[16] \quad \begin{pmatrix} b_1^c \\ C_1^c \end{pmatrix} = \begin{pmatrix} Y_1' Z A_r \Lambda_r^{-1} & A_r' Z' Y_1 & Y_1' Z A_r \Lambda_r^{-1} & A_r' Z_1 \\ Z_1' Z A_r \Lambda_r^{-1} & A_r' A_r & Z_1' Z_1 & Z_1' Z' \end{pmatrix}^{-1} \begin{pmatrix} Y_1' Z A_r \Lambda_r^{-1} & A_r' Z' y \\ Z_1' y_1 \end{pmatrix}$$

que es la fórmula <sup>12</sup> que sugiere Dhrymes para operar.

Estos estimadores tienen la misma eficiencia que los MCB si a medida que aumenta  $T$  también se aumenta  $r$ . Pero como señala Dhrymes (41) en este tipo de estimadores el tomar las propiedades asintóticas como una aproximación para pequeñas muestras a partir de cierto tamaño es más peligroso, pues al tender  $T$  a infinito no es sólo la ley de los grandes números la que opera

11. La matriz  $A$  en [14] es como la definida en [13], pero con todos los componentes principales en vez de sólo los  $r$  primeros.

12. En [16] las variables entran medidas como desviaciones sobre la media.

sino que la naturaleza del estimador mismo cambia, de una que omite cierta información muestral a otra que finalmente la utiliza.

El método tiene además ciertos inconvenientes. Primero los componentes principales no son independientes de las unidades con que se miden los elementos de  $Z$ . El número de grados de libertad

$$[17] \quad q = T - r$$

con que se va a operar, es una elección arbitraria del investigador. Un procedimiento bastante usado para determinar  $q$  (véase Klein (19)), que tiene cierta justificación intuitiva, es ir escogiendo componentes principales (de mayor a menor orden) hasta obtener cierto porcentaje fijado de antemano de la variancia generalizada de  $Z$ .

$$[18] \quad \text{Var. 6. } (Z) = \left| \frac{Z'Z}{T} \right|$$

explicado por los componentes principales escogidos. En realidad se ha trasladado el problema de elegir  $q$  a  $p$ , pero quizá la arbitrariedad ha quedado reducida al introducir un nuevo criterio que tiene presente a la variancia generalizada.

Tercero, como observa Fisher, el procedimiento no envuelve principios económicos. Por último, como han advertido varios autores y Maddala (30) es particularmente claro en ello, no tiene por qué ser cierto que los componentes principales de mayor orden sean los más correlacionados con la variable endógena considerada, ya que es la matriz  $Z'Z/T$  la que se considera para calcular los componentes principales. Para asegurar esto último<sup>13</sup> habría que calcular todos los componentes principales y seleccionar entre ellos por el procedimiento de regresión por etapas, pero tal proceso se puede realizar directamente sobre  $Z$  sin recurrir a  $C$  a no ser que se esté interesado en obtener regresores ortogonales.

Señalemos por último que el método de Kloek y Mennes (en lo sucesivo K-M) sirve también para resolver el problema de la multicolinealidad en la primera etapa.

13. Esta descripción está tomada de (30).

5. *El método de variables instrumentales estructuralmente ordenadas de Fisher (VIEO)*

Al final de 3 el problema ha quedado planteado en cómo escoger los instrumentos en la primera etapa y discutiendo la solución de K-M se ha señalado que en ella no se tienen en cuenta los principios económicos que relacionan a las variables. Lo que se propone Fisher es hacer uso de tal información.

En la elección de instrumentos hay que considerar tres aspectos: *a*) que los instrumentos no estén correlacionados con las perturbaciones aleatorias y obtener así estimadores consistentes; *b*) que los instrumentos estén causalmente conexiónados con las variables que reemplazan para lograr estimadores eficientes, y *c*) que los instrumentos no estén correlacionados entre sí en la primera etapa de la estimación.

La multicolinealidad en la segunda etapa, como señala Fisher, es inevitable, pues estamos escogiendo los instrumentos de entre las variables que entran en el sistema, que por otra parte es el único modo racional de hacerlo. Este problema es, pues, común a todas las soluciones propuestas por distintos autores.

El conjunto sobre el que tenemos que elegir los instrumentos está formado por las variables exógenas y las variables endógenas retardadas. Los dos primeros subconjuntos por definición no están correlacionados con las perturbaciones pero en general no ofrecen candidatos suficientes que cumplan con cierta aproximación las condiciones *b*) y *c*) y por ello hay que recurrir también a los retardos endógenos, sobre los que ya hemos discutido en qué condiciones cumplen aproximadamente *a*). Se trata pues ahora, de hacer la elección final mirando al principio de causalidad. Conviene aclarar que los criterios *a*) y *b*) son incompatibles en el sentido de que cuanta más causalidad, más correlación con los errores y cuanta menos correlación, menos causalidad con las variables que se reemplazan. En efecto, desde el punto de vista de la causalidad, el mejor candidato es la variable en cuestión, que precisamente se quiere sustituir por estar correlacionada con las perturbaciones y un candidato no correlacionado con los errores es uno totalmente ajeno al sistema y en consecuencia sin ningún vínculo causal, la solución ha de ser necesariamente un compromiso entre ambos criterios.

La sugerencia de Fisher es como sigue: Considérese una variable endógena cualquiera que aparezca como explicativa en la ecuación que tiene que ser estimada y asígnesele un orden causal cero. Búsqese la ecuación que determina tal variable endógena y asígnese un orden causal de uno a las variables que aparezcan como explicativas en ella. Tómense las variables endógenas de orden causal uno, búsqense las ecuaciones que las determinan y asígnese un orden causal de dos a las variables explicativas. Continúese el proceso asignando un orden causal ( $i + 1$ ) a las variables explicativas de las variables en-

dógenas de orden causal  $i$ . Con ello hemos utilizado la información económica contenida en el sistema para subdividir las variables predeterminadas según el criterio de la causalidad. Una completa clasificación de estas variables respecto a una endógena de orden cero se puede obtener de la siguiente forma, aunque no es la única. Llámese  $p$  al mayor número de órdenes causales diferentes a los que pertenezca una variable predeterminada cualquiera. A cada instrumento se le habrán asignado varios órdenes causales, tómense sólo aquellos que sean diferentes y ordénense de forma creciente en un vector de  $p$  componentes en el que a los elementos vacíos se les da el valor infinito. Este vector asignado a cada variable predeterminada sirve para establecer un orden de preferencia entre ellas, dando un valor mayor a uno de dos instrumentos, si o bien el primer elemento de su correspondiente vector es mayor que el primer elemento del otro, o si al menos un elemento del vector es mayor que el correspondiente en el otro vector, cuando todos los elementos previos son iguales. Esta prelación por el contrario causal puede modificarse para tener en cuenta el criterio de la inconsistencia, por ejemplo prefiriendo siempre a una variable exógena o sus retardos con orden causal finito sobre los retardos endógenos. El siguiente paso es combinar la información *a priori* contenida en este ordenamiento con la información *a posteriori* de la muestra. Para ello, Fisher sugiere regresar la variable de orden cero con los  $T-2$  primeros instrumentos por medio de una regresión por etapas hacia atrás partiendo del instrumento menos preferido y el valor estimado de la variable de orden cero es el que se usa como instrumento en la segunda etapa.

El hecho de que a cada variable de orden cero corresponderán generalmente instrumentos distintos en la primera etapa, hace que los peligros de inconsistencia en la segunda son mayores. Fisher (16) ya señala este punto, que es asimismo el mayor inconveniente que Griliches (31) encuentra al método propuesto. Además (véase Mitchell y Fisher (32)), en tal caso, no es claro cómo se puede obtener la matriz asintótica de variancia. Para resolver esta cuestión, Mitchell y Fisher proponen que, una vez obtenidos los instrumentos (llámeseles  $\hat{Y}_1$ ), para sustituir  $Y_1$  en la segunda etapa, se regrese cada componente de  $Y_1$  con  $\hat{Y}_1$  y  $Z_1$  y utilizar el valor estimado para  $Y_1$  procedente de tal regresión como el instrumento en la segunda etapa en vez de  $\hat{Y}_1$ . Mitchell y Fisher también resuelven otros problemas de VIEO, como el que VIEO no son independientes de la normalización empleada en las identidades, así como el problema de que varias variables predeterminadas tengan el mismo lugar en la ordenación mencionada. Desgraciadamente, todo ello complica más aún el método que por tal motivo, con excepción de Mitchell (33) no ha sido aplicado nunca.

Mitchell (33) ha cogido el reto lanzado por Griliches (31) y ha comparado los estimadores de K-M con los VIEO, aplicándolos a una versión de setenta y siete ecuaciones del modelo Brookings con ochenta variables exógenas, usando 54 observaciones. La conclusión alcanzada es que los estimadores VIEO son similares a los K-M con 15 componentes principales. En los VIEO

se usan menos de 15 variables instrumentales pero requieren mucho más tiempo en un ordenador electrónico.

No obstante, el método VIEO contiene también cierta arbitrariedad al ordenar los instrumentos, sobre todo si se quiere cierto compromiso entre el criterio de causalidad y el de consistencia. Además, está basado en que  $F - 1$  sea aproximadamente cierto, con lo que los retardos de variables endógenas determinadas en bloques de igual o mayor orden quedan descartados como posibles instrumentos. Si  $A$  no es triangular, o bien se prueban todos los retardos endógenos como posibles instrumentos, alargando más el proceso, o se decide arbitrariamente cuáles de ellos se van a considerar como candidatos y cuáles no.

6. *Estimadores por variables instrumentales basados en la forma reducida restringida*

El problema que se presenta cuando  $m > T$  en la estimación por métodos de información limitada puede ser atacado de dos maneras. Una, utilizando en la primera etapa una forma reducida no restringida con un número de variables predeterminadas compatibles con el número de observaciones, los métodos de K-M y VIEO siguen dicho camino. Otra utiliza en la primera etapa una forma reducida restringida. Este último tipo de solución, al que dedicamos esta sección, ha sido alcanzado independientemente por Dhrymes (41) y Brundy y Jorgenson (34) y coincide esencialmente con las propuestas de Lyttkens (35) y Maddala (30). En este epígrafe expondremos el método basándonos en (41) y (30) principalmente y aportaremos como nuevo una comparación de ambos procedimientos indicando en qué sentido la sugerencia de Maddala puede ser más atractiva que la de Dhrymes.

El método de Dhrymes, denominado por él mismo estimadores de información limitada iterativos por variables instrumentales (ILIVI) se basa en estimar  $a_1$  en [5] por medio de VI según la fórmula [11b] y luego resolver el sistema para dichos valores obteniendo así la FR restringida, que se usa para calcular la parte sistemática de  $Y_1$  ( $\hat{Y}_1$ ). Por último, se estima definitivamente [5] por medio de [11b] usando como matriz  $P$  a:

$$P = (\hat{Y}_1 \ Z_1)$$

Para la existencia de ILIVI se necesita que

$$T > \max_1 (n_i + m_i)$$

que en general lo cumple todo modelo. Estos estimadores son asintóticamente equivalentes a MCB.

Dhrymes propone también un procedimiento similar con el que se obtienen unos estimadores asintóticamente equivalentes a MCT, pero para su existencia el modelo ha de cumplir:

$$T > n$$

condición que no satisfacen los sistemas amplios.

En la obtención de ILIVI hay que decidir arbitrariamente qué instrumentos emplear en la primera regresión de la forma estructural. De hecho existen  $\binom{m - m_1}{ni}$  posibles decisiones. Asintóticamente independiente de

qué elección inicial se haga, obtenemos estimadores equivalentes a MCB, pero en pequeñas muestras las diferencias pueden ser importantes.

Maddala (30) estudia varios métodos de estimación y el que ofrece mejor comportamiento es el que él denomina, estimación por VI basados en la FR restringida iterando mediante la resolución de la FR (VIFRR). El procedimiento es el mismo que el propuesto por Dhrymes pero continuándolo de forma iterativa hasta que converja. Es decir, con los estimadores definitivos de Dhrymes se vuelve a resolver el sistema, obteniendo la forma reducida y estimando con ella de nuevo la parte sistemática de  $Y_1$  que se usa como instrumento en la estimación de la forma estructural y así sucesivamente. Maddala demuestra que los estimadores VIFRR tienen la misma distribución asintótica que los MCB, es decir, con la iteración no se ha ganado en eficiencia respecto al método ILIVI.

Obsérvese que los métodos que usan la FR restringida utilizan toda la información sistemática de la FE y no sólo la contenida en las restricciones cero de la ecuación en cuestión, con todas las ventajas e inconvenientes que de ello se puedan derivar. Maddala demuestra que el uso de extrainformación no aumenta la eficiencia, lo cual, de paso, indica que la mayor eficiencia de MCT y MVIC sobre MCB y MVIL es por vía de la información adicional sobre los procesos aleatorios y, por tanto, no sólo la parte sistemática del modelo, sino también las perturbaciones aleatorias, deben ser introducidas con precisión en el programa siempre que sea posible.

No obstante lo mencionado arriba, los estimadores VIFRR tienen sobre los ILIVI la ventaja —según los resultados detenidos por Maddala— de que, independientemente de los instrumentos iniciales, VIFRR convergen al mismo punto o si en varios puntos sólo uno es estable. Vemos, pues, que el iterar sirve para eliminar la arbitrariedad señalada al método de Dhrymes, logrando una solución única independiente del punto de partida.

La última indicación de Maddala en su artículo muestra que los estimadores VIFRR no son asintóticamente equivalentes a los MVIC cuando la variancia de los errores es diagonal (MVICD), pero una pequeña modificación en las ecuaciones normales de VIFRR nos dan las ecuaciones correspondientes a MVICV que son más eficientes que los primeros.

Para concluir digamos que cuando  $m > T$  no fallan los métodos de

estimación estructural de información limitada, sino sólo aquellos que operan a través de la FR no restringida. En tal caso, además de los métodos propuestos por Kloek-Mennes y Fisher, otra solución, quizá más prometedora, es operar con la FR restringida. Con qué instrumentos iniciar el proceso de los estimadores que siguen esta línea es una decisión arbitraria, pero los VIFR obtienen soluciones únicas independientes del punto inicial, lo que quizá sirve al mismo tiempo para demostrar que tener en cuenta las restricciones de todas las ecuaciones favorece para encontrar una solución única, aunque para ello es necesario un proceso iterativo y no basta el pararse en la primera vuelta. No obstante, los MVICD requieren poco más trabajo y son más eficientes, por lo que es el tipo de estimador recomendable, cuando MVIC u otros más sofisticados como los propuestos por Sargan o aquellos que utilizan la información espectral no son posibles. Por último, la experiencia con MVICD en el modelo alemán (42) es el que produce estimadores más estables.

## BIBLIOGRAFÍA

1. JOHNSTON, J.: *Métodos de Econometría* (Ed. Vicens Vives, Barcelona, 1967).
2. Comisaría del Plan de Desarrollo Económico y Social: «Elaboración y especificación del modelo Econométrico», Madrid, 1972.
3. KADANE, J. B.: «Comparison of *K*-Class Estimators when the disturbances are small», *Econometrica*, septiembre, 1971.
4. HAAVELMO, T.: «The Statistical implication of a System of simultaneous equations», *Econometrica*, enero, 1943.
5. KOOPMANS, T. C. (editor): *Statistical Inference in Dynamic Economic Models* (Wiley, Nueva York, 1950).
6. KOOPMANS, T. C., and HOOD (editores): *Studies in Econometric Methods* (Wiley, Nueva York, 1953).
7. KELEJIAN, H. H.: «Two-stage Least Squares and Econometric Systems linear in Parameters but nonlinear in the endogenous variables», *JASA*, junio, 1971.
8. GOLDFELD, S. M., and QUANDT, R. E.: *Nonlinear Methods in Econometrics* (North-Holland, 1972).
9. GOLDBERGER, A. S.: *Teoría Econométrica* (Tecnos, Madrid, 1970).
10. THEIL, H.: *Estimation Simultaneous correlation complete equation System* (La Haya, Centraad Plaubureau, 1953).
11. THEIL, H.: *Economic Forecasts and Policy* (North-Holland, 1961).
12. BASMANN: «A generalised classical method of linear estimation of coefficients in a Structural equation», *Econometrica*, enero, 1957.
13. ZELLNER y THEIL: «Three-Stage Least Squares: simultaneous estimation of simultaneous equations», *Econometrica*, enero, 1962.
14. WOLD, H.: «Causality and Econometrics», *Econometrica*, abril, 1957.
15. NAGAR, A. L.: «Double *K*-class estimators of parameters in simultaneous equations and their small sample properties», *International Economic Review*, 1962, páginas 168-188.
16. FISHER, F. M.: *Dynamic Structure and Estimation in Economy-wide Econometric Models*, capítulo 15 de (17).
17. DUESENBERY, FROMM, KLEIN y KUHN: *The Brookings Quarterly Econometric Model of the U.S.A.* (North-Holland, 1965).
18. THOMAS, J. J., y WALLIS, K. F.: «Seasonal Variation in Regression Analysis», *JASA*, 1971, pp. 57-72.

19. KLEIN, L. R.: «Estimation of Interdependent System in macro-econometrics», *Econometrica*, abril, 1969.
20. HENDRY, D.: «Maximum Likelihood Estimation of Systems of Simultaneous Regression Equations with errors Generated by a vector autoregressive process», *International Economic Review*, junio, 1971.
21. DENIS SARGAN, J.: «The Maximum Likelihood estimation of Economic relationships with autoregressive residuals», *Econometrica*, julio, 1961.
22. FAREBROTHER y SAVIN: «Notes on the Graph of the *K*-class estimator: on algebraic and statistical interpretation», Notas ciclostiladas, Universidad de Manchester, noviembre, 1972.
23. KADIYALA, K. R.: «An exact Small Sample property of the *K*-class estimators», *Econometrica*, noviembre, 1970.
24. RICHARDSON: «The exact distribution of a structural coefficient estimator», *Journal American Statistical Association, JASA*, 1968.
25. SAWA: «The exact sampling distribution of O.L.S. and 2S.L.S. estimators», *JASA*, 1969, pp. 923-937.
26. MARIANO y SAWA: «The exact finite-sample distribution of L.I.M.L. estimation in the case of two included endogenous Variables», *JASA*, 1972.
27. NAGAR: «The bias and the moment matrix of the general *K*-class estimators of the parameters in simultaneous equations», *Econometrica*, octubre, 1959.
28. KLOK y MENNES: «Simultaneous equation estimation based on Principal Components of Predetermined Variables», *Econometrica*, enero, 1960.
29. DHRYMES: *Econometrics, statistical foundations and applications* (Harper, 1970).
30. MADDALA: «Simultaneous Estimation methods for large and medium-size Econometric Models», *Review of Economic Studies*, 1971.
31. GRILICHES: «The Brookings Model Volume: A review article», *Review of Economic and Statistics*, mayo, 1968.
32. MITCHELL y FISHER, F. M.: «The choice of instrumental variables in the estimation of Economic-wide Econometric Models: some further Thoughts», *International Economic Review, I.E.R.*, junio, 1970.
33. MITCHELL: «Estimation of large Econometric models by Principal Components and Instrumental Variables methods», *Review of Economics and Statistics*, 1971.
34. BRUNDY y DALE JORGENSON: «Efficient estimation of simultaneous equations by Instrumental Variables», *Review of Economics and Statistics*, agosto, 1971.
35. LYTTKENS: «Non-interactive estimation of Special Interdependent systems by G.E.I.D. Specification», *Econometric Society, Blaricum Meeting*, enero, 1967.
36. HANNAN: *Multiple Time Series*, Wiley, 1971.
37. HANNAN y TERREL: «Time Series regressions with linear constraints», *I.E.R.*, junio, 1972.
38. SARGAN, J. D.: «The estimation of Economic relationships using instrumental Variables», *Econometrica*, 1958, pp. 393-415.
39. SARGAN, J. D., y DRETTAKIS, E. G.: «Missing chain in an autoregressive model», Notas ciclostiladas, London School of Economics y University of Leeds, 1971.
40. DRETTAKIS, E. G.: «Missing data in Econometric estimation», tesis doctoral, no publicada, L.S.E., 1971.
41. DHRYMES: «A simplified structural estimation for large scale econometric models», *The Australian Journal of Statistics*, noviembre, 1971.
42. KRELLE, MARTIENSEN, SONDERMANN y SCHLOENBACH: «Effects of Government decisions on Economic Development. Experiments with a medium Term Forecasting Model for the West German Economy», *European Econometric Meeting*, Budapest, 1972.
43. THEIL, H.: *Principles of Econometrics* (North-Holland Publishing Co., Amsterdam, 1971).