



**TEACHING AND RESEARCH PROJECT PRESENTED
BY ANTONI ESPASA TO ACCESS TO THE CHAIR OF
ECONOMETRICS AT THE STATISTICS AND
ECONOMETRICS DEPARTMENT OF THE
UNIVERSIDAD CARLOS III DE MADRID.**

The project is written in Spanish

13, November 1992.

PROYECTO DOCENTE

Que se presenta para el concurso a la plaza del Cuerpo de Catedráticos de Universidad, área de conocimiento: Fundamentos del Análisis Económico, perfil docente: Econometría, convocada por Resolución del Presidente de la Comisión Gestora de la Universidad Carlos III de Madrid, de fecha 2 de marzo de 1992 (Boletín Oficial del Estado de 24 de marzo).

Dr. Antoni Espasa Terrades.

Madrid, 13 de Noviembre de 1992

PROYECTO DOCENTE E INVESTIGADOR

1.	INTRODUCCIÓN.....	1
I.	LA ECONOMETRÍA Y SUS FUNDAMENTOS ESTADÍSTICOS.....	9
	<i>I.1.- Definición de la Econometría.....</i>	9
	<i>I.2.- Necesidad de formalizar los fundamentos estocásticos de la Econometría.....</i>	11
	<i>I.3.- Teoría Económica y datos económicos.....</i>	13
	<i>I.4.- Formulación probabilística de las teorías económicas.....</i>	14
	<i>I.5.- La distribución de Haavelmo.....</i>	16
	<i>I.6.- El proceso de reducción de Haavelmo.....</i>	18
II.	PANORÁMICA GENERAL SOBRE LOS MODELOS ECONOMÉTRICOS.....	26
	<i>II.1.- Sistematización del proceso reductor.....</i>	26
	<i>II.1.0.- El proceso generador de datos.....</i>	27
	<i>II.1.1.- Marginalización del PGD respecto a las variables que no entran en el estudio del problema considerado.....</i>	28
	<i>II.1.2.- Condicionización según la recursividad temporal.....</i>	28
	<i>II.1.3.- Restricciones de homogeneidad.....</i>	30
	<i>II.1.4.- Restricciones distribucionales.....</i>	32
	<i>II.1.5.- Restricciones en la dependencia de los datos (memoria).....</i>	34
	<i>II.1.6.- Condicionización contemporánea.....</i>	35
	<i>II.1.7.- El proceso reductor en datos de corte transversal o de panel.....</i>	37
	<i>II.1.8.- Conclusión.....</i>	39
	II.2. Exogeneidad y determinación de los parámetros de interés.....	40
	II.3. Formulaciones dinámicas alternativas.....	44
	II.4. Esquema integrador de los modelos econométrico más usuales.....	50
	II.4.1. Esquema integrador de los modelos econométrico más usuales.....	50
	II.4.2. Modelos con la restricción de variables exógenas.....	55
	II.4.3. Modelos identificados con restricciones teóricas: MODELOS ESTRUCUTRALES.....	58
	II.4.4. Modelos recursivos.....	62
III.	EL PLANTEAMIENTO UNIECUACIONAL DE LA MODELIZACIÓN ECONOMÉTRICA.....	67

III.1. Procedimientos de información limitada y variables instrumentales.....	67
III.2. Precisiones sobre el uso de modelos econométricos.....	71
III.3. El principio “de lo general a lo particular” en la modelización econométrica.	74
III.4. El planteamiento de información limitada en sistemas cointegrados.....	77
III.5. Convergencia hacia el modelo VECM.	79
IV. DISEÑO Y EVALUACIÓN DE MODELOS ECONOMÉTRICOS	81
IV.1. Implementación el principio de adecuación estadística en la modelización econométrica.....	81
IV.2. Resumen sobre el proceso de modelización.....	92
IV.3. Conclusiones para la enseñanza de la Econometría.	93
V. BIBLIOGRAFÍA	94

1. INTRODUCCIÓN

Para explicar el proyecto docente que llevaría a cabo desde la Cátedra de Econometría, del departamento de Estadística y Econometría de la Universidad Carlos III de Madrid, en caso de acceder a ella, es conveniente empezar realizando una escueta presentación panorámica de los modelos econométricos.

Así, la sección I parte del supuesto de que el objetivo de la Teoría Económica, tal como se concibe en Haavelmo (1944), es explicar los hechos observados, de modo que las proposiciones teóricas sean “construcciones cognoscitivas”. A continuación se formula, en línea con el pensamiento de dicho autor, que la conexión entre Teoría y datos que la Econometría pretende desarrollar queda debidamente situada estableciendo el siguiente punto de partida: la correspondencia entre variables económicas y observaciones pasivas de la naturaleza puede verse como la correspondencia entre “universo poblacional” y “muestra” de la teoría probabilística.

Al concebir los datos observados como una realización del esquema teórico estocástico (proceso generador de datos) tenemos que los modelos econométricos sólo pueden justificarse en tanto en cuanto vengan derivados de dicho proceso, encontrando en él, en consecuencia, su fundamento estadístico.

Tras los procesos de agregación, transversal y temporal, y de marginalización respecto a las variables que no intervienen en un determinado problema económico, el proceso generador de los datos en cuestión puede verse como la función de distribución conjunta en las correspondientes variables teóricas (poblacionales). Las relaciones o funciones estocásticas entre las variables económicas que se derivan de tal función de distribución conjunta con múltiples, y el objetivo de la Teoría Económica es “realizar elecciones provechosas de tales formas”, que, a su vez, deberán que tener en cuenta las características de los datos, de modo que las funciones estocásticas resultantes sean teóricamente coherentes y compatibles con los datos. Estas funciones estocásticas, o una reformulación de las mismas, constituyen los modelos econométricos.

De lo anterior se desprende que un modelo econométrico se obtiene tras aplicar un proceso reductor a la función de distribución conjunta de los datos.

Sistematizando este proceso reductor, tal como se hace en la sección II.1, es posible, tras presentar los distintos conceptos de exogeneidad y las diferentes alternativas de las formulaciones dinámicas, construir un esquema integrador –tomado de Espasa y Cancelo (1992)- de los modelos econométricos más usuales. Este esquema, cuadro 1, ofrece una orientación clara del tipo de restricciones –basadas en la naturaleza de los datos o en las proposiciones de la teoría- que cada modelo ha incorporado en el proceso reductor aplicado para su obtención.

Como las restricciones que incorporan los distintos tipos de modelos econométricos han sido sistematizadas, el esquema anterior ofrece también una relación de mayor o menor generalidad en los modelos econométricos.

El proceso reductor se analiza teniendo en cuenta –como caso general- la dimensión temporal de las variables económicas, pero es igualmente válido para datos de corte transversal, como se discute e ilustra en el epígrafe II.1.7.

El enfoque de determinación simultánea presente en el desarrollo de los modelos econométricos estructurales del esquema anterior, siguiendo la tradición asentada por Haavelmo, fue atacado, a partir del principio lógico de encadenación casual, por Wold (1954, 1960 y 1964). Esta discusión puede sinterizarse de la siguiente forma. En los miles de millones de decisiones de los agente económicos de una determinada economía en un cierto período de tiempo, necesariamente ha de operar el principio de recursividad causal, por el que una variable no puede ser al mismo tiempo causa y efecto de otra. Pero en Economía no se analizan los datos correspondientes a los millones y millones de decisiones primarias, sino una determinada agregación de los mismos a lo largo de un período de tiempo, año, mes, día, etc., y sobre un conjunto de agentes. Este proceso de agregación es el que induce la simultaneidad causal presente en los modelos econométricos estructurales¹.

La reflexión sobre recursividad causal frente a la simultaneidad sirve para enlazar con el problema de abordar un enfoque de información limitada en la construcción y aplicación de modelos econométricos. Esto conduce a los modelos econométricos uniecuacionales, que son la base sobre la que se asienta la metodología de la London School of Economics. La estimación de estos modelos,

¹ De hecho la encadenación causal excluye incluso la correlación contemporánea entre variables con una relación de causa efecto entre ellas, cuando se aplica sobre las variables que recogen las decisiones primarias.

cuando incluyen variables explicativas endógenas, requieren la aplicación del procedimiento de variables instrumentales. La estimación mediante variables instrumentales es una técnica de aplicación más amplia, por ejemplo de modelos en los que las variables explicativas se miden con error; pero en la estimación de modelos simultáneos el procedimiento de variables instrumentales se convierte, aunque no existan errores de medición, en un tema central que necesita incorporarse, en cualquier caso, en los programas de docencia. Si todas las variables explicativas cumplen la restricción de exogeneidad se tienen el modelo de regresión dinámico, y cuando éste se aplica a problemas estáticos, por ejemplo, con los datos de corte transversal, queda reducido a un modelo de regresión clásico.

Con ello, desde la panorámica más general posible, la distribución conjunta de los datos, se llega a la derivación de los modelos más simples en Econometría la regresión y la regresión dinámica. Estos modelos, siendo sencillos, son básicos en la formación econométrica, pues en ellos aparecen ya la mayor parte de las cuestiones de estimación e inferencia en Econometría. En consecuencia con ello, el programa de licenciatura comienza en la regresión, pero presentándola dentro de un marco con grandes restricciones sobre el proceso generador de los datos. Esto tiene la ventaja de que desde el principio el alumno es instruido dentro de un contexto estadístico adecuado y conectado con la Teoría Económica. Así, generando la regresión a partir de una determinada hipótesis sobre la función de distribución conjunta de los datos económicos, es muy simple introducir en las primeras lecciones el concepto de endogeneidad y exogeneidad, que lleva a que el alumno asimile desde el principio la necesidad de que la Teoría Económica oriente la construcción de los modelos econométricos.

En el proceso reductor tiene especial relevancia el conjunto informativo \mathbb{D}_t sobre el que se condiciona la función de distribución de las variables de interés, Y_t , donde t , puede referirse al tiempo o a agentes económicos según sea el tipo de datos utilizados. A partir de los diferentes conceptos de exogeneidad, la condicionalización queda estructurada de forma que no hay pérdida de información en formular los modelos econométricos a partir de la función de distribución condicional $D(Y_t/\mathbb{D}_t; \lambda)$, donde λ es un vector de parámetros, que es función de características de la función de distribución de y_t/\mathbb{D}_t ignorando la distribución marginal de \mathbb{D}_t .

Con ello denominando $E(Y_t/\mathbb{D}_t; \lambda) = \mu_t$ a la esperanza matemática condicional de Y_t , se tiene que

$$Y_t = E(Y_t/\mathbb{D}_t; \lambda) + e_t = \mu_t + e_t \quad (1)$$

donde e_t es un residuo ortogonal a $E(\cdot)$.

Suponiendo que dicha esperanza es lineal se tiene que μ_t , la parte sistemática del modelo, es una función lineal de las variables contenidas en \mathbb{D}_t . En ambas expresiones e_t constituye una innovación respecto al conjunto informativo \mathbb{D}_t , es decir, es independiente en media respecto a él.

En la formulación de modelos de tipo (1) la Teoría Económica interviene fundamentalmente señalando; sobre qué variables se desea marginalizar el proceso generador de datos para concluir con el vector Y_t como el vector de variables de interés, y \mathbb{D}_t como el conjunto de variables con respecto a las cuales se quiere explicar Y_t .

Con ello, la derivación de (1) se realiza directamente de $D(Y_t/\mathbb{D}_t; \lambda)$, de modo que la forma funcional entre las variables económicas recogida en la parte sistemática del modelo $-E(Y_t/\mathbb{D}_t)-$ es compatible con la Teoría Económica, pero no ha sido restringida por ella. Por eso, a tales modelos se les puede denominar modelos estadísticos, que no son más que modelos de forma reducida.

La Teoría Económica no siempre formula las relaciones entre variables de acuerdo con (1) sino que puede formular una relación de tipo;

$$Y_t = F(Y_t/\mathbb{D}_t; \varphi) + a_t \quad (2)$$

en donde a_t es una innovación, pero no es ortogonal a $F(Y_t/\mathbb{D}_t)$ pues ésta no es una esperanza matemática condicional. La ecuación (2) constituye un modelo estructural y φ es un vector de parámetros estructurales.

Conceptualmente, se puede reformular (2) como un modelo de forma reducida del tipo

$$Y_t = G(\mathbb{D}_t, \varphi, a_t) \quad (3)$$

Si F es lineal, (3) se puede asociar con una especificación de la media condicional como ocurre en (1), en cuyo caso la forma reducida se podría adoptar como base para la estimación. Sin embargo, en general esto no es posible y la estimación se tiene que basar directamente en la ortogonalidad entre a_t y variables instrumentales, esto es, las variables en que condicionamos o funciones de las mismas. En definitiva, los modelos de variables instrumentales especifican características de la distribución $D(Y_t/\mathbb{D}_t)$ consistentes en la formulación de la media condicional de funciones de Y_t y \mathbb{D}_t , por lo que en cierta forma generalizan los modelos de regresión en los cuales invariablemente el interés se centra en la media condicional de Y_t .

Dadas las restricciones impuestas en (2) por F , la identificación de los parámetros φ dependerá del número de instrumentos válidos, de acuerdo con la Teoría Económica, que estén disponibles en los datos. Si los parámetros no están identificados, el modelo teórico no puede estimarse, al menos con la información disponible.

En la medida en que el modelo (3) contenga restricciones de sobreidentificación –dictadas por la teoría- éstas pueden ser contrastadas.

En conclusión, los modelos econométricos, estructurales y de forma reducida, se pueden formular a partir del proceso generador de los datos, mediante un esquema reductor que especifique el conjunto informativo condicional relevante, teniendo para ello en cuenta las características de los datos y los resultados de la Teoría.

Con esta racionalización, los modelos estructurales constituyen un paso ulterior en los programas de Econometría, pero en perfecta conexión, con un mismo esquema estadístico, de modo que no hay ruptura de continuidad en la línea de aprendizaje econométrico establecida.

En la Econometría de series temporales existen diversas metodologías, pero podemos fijarnos en tres. La basada en modelos simultáneos (SEM) –o su correspondiente forma reducida restringida-, siguiendo la tradición de Haavelmo perpetuada por la Cowles Commission, la basada en modelos de forma reducida, y particularmente en modelos VAR, que no imponen a los datos ninguna orientación teórica, (Sargent y Sims (1977), Sims (1980)) y la de la London School of Economics, basada en modelos uniecuacionales dinámicos generales. El desarrollo de la teoría de cointegración está produciendo un acercamiento de todas esas metodologías, que convergen hacia los modelos VAR con mecanismos de corrección del error o modelos VECM (vector error corrección models). Esta convergencia se explica bien en el contexto estadístico empelado para derivar los modelos econométricos y en la sección III.5 se comenta tal proceso convergente que sirve para cerrar los programas de formación básica en licenciatura y doctorado. Dicho proceso se está produciendo bastante rápidamente pues los usuarios del enfoque VAR están cada vez más interesados en hacer análisis estructural, lo que conlleva la imposición de restricciones, que de acuerdo con Sims (1980) se considerarían a posteriori, una vez estimado el modelo VAR.

No obstante, con la cointegración puede ser importante imponer también algunas restricciones al principio, lo cual puede verse como operar con la forma reducida restringida correspondiente a un modelo SEM. Por último, la formulación de modelos uniecuacionales en presencia de variables cointegradas puede ser muy compleja, por lo que será recomendable trabajar con sistema de ecuaciones.

La formulación de los modelos econométricos requiere, como ya se ha dicho, la consideración de una serie de restricciones, que suponen unas hipótesis sobre el comportamiento de los datos, por lo que son contrastables. En consecuencia, el esquema reductor descrito sirve también para basar en él la estrategia de diseño y evaluación de los modelos econométricos, que se discute en la sección IV.

Así pues, en las distintas etapas del aprendizaje econométrico los alumnos van a ser expuestos, no sólo a las técnicas de estimación que nos proporcionan

hipótesis concretas sobre los valores de sus parámetros, sino, principalmente, al contraste de hipótesis sobre los mismos, que conduzcan a obtener una experiencia útil en la especificación y validación de modelos econométricos.

El plan docente que se propone en esta memoria está dentro de la tradición econométrica de la LSE, pero existen varios aspectos del mismo que quedan algo al margen de ella. Señalamos dos, ilustrando la discusión en el ámbito usual de la metodología empleada, el modelo uniecuacional. En el análisis económico aplicado de corto plazo, tanto a nivel macro como micro, el problema de predicción es especialmente importante. Para abordarlo, no siempre es factible, con los datos disponibles, la especificación y estimación de modelos uniecuacionales estructurales (más precisamente, los modelos de forma reducida restringida que de ellos se derivan). Por ello, el economista encuentra que en el desarrollo de su profesión se ve necesitado, según los casos, tanto de la disponibilidad de modelos de estructurales como de forma reducida general. En bastantes ocasiones ni tan siquiera se puede disponer de observaciones de las variables explicativas al mismo nivel de desagregación que se dispone para la variable endógena, en cuyo caso sólo la modelización univariante es posible. En problemas sobre datos temporales, los modelos de forma reducida se especifican haciendo un menor uso de la Teoría Económica, que muchas veces se limita a especificar las variables que se pueden marginalizar y aquéllas que pueden ser o son candidatas a consideraciones exógenas. Estos modelos son claramente de naturaleza econométrica tal como se desprende de la discusión incluida en la sección II, si bien su desarrollo está muy ligado a la literatura estadística de Series Temporales. Obviamente su naturaleza econométrica y su necesidad en el campo aplicado, lleva a que estos modelos, incluidos en el esquema integrador de la sección II.4, tengan un tratamiento adecuado en los programas de licenciatura y de doctorado.

Asimismo, tal como se discute en la sección II.3, la tradición econométrica de la LSE se centra, fundamentalmente, en utilizar la hipótesis de truncamiento temporal en los supuestos sobre “memoria”, que se realizan en el proceso reductor seguido para formular modelos econométricos sobre series temporales. Una hipótesis más general es la formada por las condiciones de ergodicidad e invertibilidad, que lleva a modelos con formulaciones dinámicas racionales. Es decir, modelos en los que las dependencias dinámicas no se agotan hasta el infinito, pero son convergentes, con lo que admiten una aproximación mediante polinomios racionales. En la práctica la elección entre ambas alternativas se reduce a un problema de parquedad en la

parametrización, pero no es un problema sobre el conjunto informativo sobre el que se condiciona, como algún lector poco atento pudiera deducir de ciertas publicaciones.

En las secciones I a III se discuten los fundamentos y la conexión lógica existente entre los distintos elementos del plan docente. Tales fundamentos y desarrollo se basan en un trabajo previo recogido en los capítulos II y III de Espasa y Cancelo (1992). Sin embargo, en esta memoria se va más allá de lo expresado en la obra citada, ofreciendo un esquema plenamente integrador de los modelos estructurales y de los modelos que hemos denominado estadísticos, al tiempo que – siguiendo la tradición de la London School of Economics- se utiliza, véase sección IV, el proceso reductor subyacente en la formulación de los modelos econométricos, para sistematizar, dentro de la enseñanza de la Econometría, los problemas de diseño y evaluación de los modelos.

Dada la conexión de esta memoria con la obra citada, en la confección de la primera me he visto muy beneficiado por las discusiones, que en la realización de la segunda, tuve con distintas personas. De entre ellas debo destacar a Manuel Arellano, Carlos Ballabriga, Olymplia Bover, Julia Campos, Miguel Delgado, Juan José Dolado, Álvaro Escribano, Juan Romo, Miguel Sebastián, Mark Steel y Rubén Zamar, y el coautor del libro mencionado, José Ramón Cancelo.

En la redacción del proyecto docente de esta memoria me han servido de especial apoyo, además del volumen Espasa y Cancelo (1992), diferentes trabajos, muchos de ellos dentro de la tradición econométrica de la London School of Economics. De entre tales trabajos debo destacar, Engle et al. (1983), Hendry y Richard (1982), Ericsson et al. (1990), Ericsson (1991), Ericsson y Hendry (1985), Spanos (1986, 1988, 1989), Hendry et al. (1989), Hendry y Ericsson (1991), Philips (1988, 1991), Granger (1990), Gilbert (1986) y Morgan (1990).

La última parte de la memoria se dedica a describir mi proyecto investigador.

I. LA ECONOMETRÍA Y SUS FUNDAMENTOS ESTADÍSTICOS

I.1.- Definición de la Econometría

La Econometría moderna está ligada a la fundación de la “Econometric Society” en 1931 y a su revista *Econometrica*, que empezó a publicarse en 1933. Como reconoce Frisch (1936), el término “Econométrica” (Oekonometrie) fue usado por primera vez de Ciompa (1910), asignándole un contenido descriptivo y es en Frisch (1923) en donde tal término aparece en su significado moderno, recogándose así en la editorial del primer número de *Econometrica*, Frisch (1933). Pesaran (1987) perfila bien dicho concepto partiendo de la definición que él extrae de las referencias mencionadas de Frisch, “la aplicación de las matemáticas y de los métodos estadísticos al análisis de los datos económicos” y precisándola con las palabras de Samuelson, Koopmans y Stone (1954), “como el análisis cuantitativo de los fenómenos apropiados de inferencia”.

Con la definición precedente, los Métodos Econométricos se convierten en uno de los principales instrumentos para el progreso de la Economía, ya que la conjunción de teoría y observación realizada a través de los procedimientos racionales que proporciona la inferencia estadística, desarrolla un proceso progresivo en el conocimiento económico. En efecto, una teoría o generalización económica inicial es necesaria para organizar los datos reales disponibles y sobre ellos contrastar estadísticamente mediante modelos econométricos dicha teoría. Con tal contraste se obtendrá que ciertos aspectos de la Teoría parecen adecuados para explicar el mundo real, pero otros deben modificarse, dando lugar a una nueva formulación teórica que será también confrontada con los datos para que genere otra formulación más precisa, y así sucesivamente. Frisch (1993) resume esta idea con las siguientes palabras: “... estudios estadísticos nuevos y otros estudios empíricos (“factual”) deben ser el sano elemento de perturbación que constantemente amenace e inquiete al teórico y que le impida descansar sobre un conjunto heredado y obsoleto de hipótesis”.

En este proceso los avances de la Teoría también condicionan a la Econometría, exigiendo de ésta el desarrollo de modelos probabilísticos, que cada vez más constituyan a priori un mejor marco para inferencia: piénsese, por ejemplo, en modelos probabilísticos no lineales con dependencia temporal.

1.2.- Necesidad de formalizar los fundamentos estocásticos de la Econometría

Un punto básico de la definición anterior reside en el término de “métodos apropiados de inferencia”, lo que conduce al tema de los fundamentos estadísticos de la Econometría. Un planteamiento a fondo del problema se encuentra en el trabajo de Haavelmo (1944), “El enfoque probabilístico en Econometría”², sobre el que Hendry et al. (1989) dicen que “mercidamente es considerado como el trabajo que constituyó a la Econometría moderna como una disciplina separada”³

Dicha publicación es una monografía de 115 páginas precedidas de un prólogo, que apareció como un suplemento de *Econometrica*. Haavelmo comienza el prólogo señalando que su estudio pretende ser una aportación a la Econometría y representa un intento de ofrecer unos fundamentos teóricos para el análisis de las interrelaciones entre variables económicas, que se basa en la teoría moderna de la probabilidad y la inferencia estadística⁴. Esta teoría estadística moderna era la que se venía consolidando desde los años treinta y que con Kolmogorov (1933) había recibido los fundamentos axiomáticos y su firme inserción como una parte de las Matemáticas.

Haavelmo justifica su trabajo diciendo, en línea con su maestro Frisch, que el objetivo de la Econometría es la conjunción de la Teoría Económica con los datos, usando la inferencia estadística como puente, pero ese puente nunca ha sido explicado. La resistencia de los economistas a adoptar modelos probabilísticos y la incongruencia que ello implica, junto con una defensa del enfoque que él propone, se discute en el resto del prólogo, que vale la pena resumir aquí.

Haavelmo señala que de hecho la adopción de modelos probabilísticos ha sido considerada como un crimen en la investigación económica, una violación de la auténtica naturaleza de los datos. Sin embargo, instrumentos de inferencia estadística, como desviaciones estándar y coeficientes de correlación múltiple, se han venido utilizando para apoyar determinadas opiniones. Ahora bien, utilizar la Teoría

² Sobre este trabajo de Haavelmo véase Morgan (1990) y Hendry et al. (1989). Los comentarios que siguen recogen varias interpretaciones y consideraciones expresadas en dichos artículos.

³ Este trabajo, como señala el autor en el prólogo, se escribió en Harvard en 1941. Así, fue conocido en los círculos “econométricos” antes de su publicación en 1944.

⁴ Entre las citas que realiza Haavelmo se encuentran, Neyman (1937), Levy (1937), Wilks (1937), etc. De hecho, la conversación de Haavelmo, discípulo de Frisch, al enfoque probabilístico se debió a sus discusiones con Neyman (véase Morgan (1990), pág. 242, nota 10).

Estadística sin referirla a un esquema estocástico es un modo inconsistente de proceder.

La resistencia entre los economistas de la época a aceptar modelos probabilísticos como una base para la investigación económica se debía –argumenta Haavelmo- “a que se tiene un concepto muy estrecho de probabilidad y variables aleatorias”. A la creencia de que los modelos probabilísticos requieren observaciones independientes, característica que ciertamente no se da en las series temporales. Pero tal requisito no es necesario. “Es suficiente suponer que el conjunto n observaciones puede ser considerado como una observación de n variables que siguen una ley probabilística conjunta de n dimensiones, cuya existencia puede ser meramente hipotética. Con ello se pueden contrastar hipótesis sobre la ley probabilística conjunta e inferir sobre su posible forma, mediante un punto muestral en n dimensiones. La Teoría Estadística moderna ha hecho considerables progresos en resolver tales problemas de inferencia”.

La investigación económica actual –continúa diciendo Haavelmo- se basa en vagas nociones de probabilidad. Así, términos como “prácticamente imposible” o “caso seguro” se pueden formular como “la probabilidad es casi cero” o “la probabilidad es casi uno”. “Pero esta formulación mediante conceptos probabilísticos tiene la ventaja de ser analítica y de poder derivar a partir de ella afirmaciones utilizando las reglas de la lógica”. En tanto en cuanto todas las leyes económicas que tenemos se pueden formular en términos probabilísticos no hay pérdida de generalidad en tal enfoque. En consecuencia, si se quiere aplicar la inferencia estadística para contrastar hipótesis económicas, tenemos que ello implica que la formulación económica representa una hipótesis estadística, es decir, afirmaciones referentes a ciertas distribuciones probabilísticas, lo cual conlleva la consideración de las relaciones teóricas como estadísticas. La creencia de que es posible utilizar la inferencia estadística sin este engarce estocástico en las relaciones teóricas “sólo puede basarse en la falta de precisión al formular problemas”.

1.3.- Teoría Económica y datos económicos

Los párrafos de resumen precedentes sirven para centrar el tema de los fundamentos estadísticos de la Econometría, pero es necesario incidir más en el trabajo de Haavelmo (1944). El enfoque probabilístico que en él se desarrolla está orientado a relacionar la teoría con los datos, pero en ello hay que puntualizar bien el concepto de modelo teórico. En tal contexto todo modelo teórico tiene asociado las condiciones experimentales, es decir, las reglas de medición de las variables que entran en el modelo, bajo las cuales la teoría es correcta. Tales normas de medición constituyen, por tanto, un diseño de experimentos reales.

Esta asociación de la teoría con el correspondiente diseño experimental también es de aplicación a la Economía, con lo que un modelo obtiene sentido económico solamente cuando se le asocia con un diseño experimental real, que describe (e indica cómo medir) un sistema de variables reales –no necesariamente observables- que tienen que identificarse con las correspondientes variables teóricas. Como consecuencia de esta identificación de variables teóricas y reales, el modelo teórico se convierte en una hipótesis a priori sobre los fenómenos reales. Tales hipótesis se componen de la implicación conjunta de una teoría y un diseño de experimentos. En este contexto es natural convenir que una teoría será denominada verdadera o falsa, según las hipótesis implicadas sean verdaderas o falsas, al ser contrastadas frente a los datos escogidos como las variables verdaderas.

El diseño experimental es un componente básico de cualquier teoría cuantitativa, aunque muchos economistas no describan sus diseños de experimentos explícitamente. No obstante, éstos pueden ser de dos tipo: “(1) experimentos que nos gustaría hacer para ver si ciertos fenómenos económicos reales –cuando son aislados artificialmente de “otras influencias”- verifican determinadas hipótesis y (2) la avalancha de experimentos que la Naturaleza está generando continuamente desde su propio e inmenso laboratorio, que nosotros meramente contemplamos como observadores pasivos” (Haavelmo 1994, pág. 14). En el segundo caso sólo se puede tratar de ajustar las teorías a la realidad que aparece frente a nosotros. En tal circunstancia tratamos de escoger una teoría y un diseño de experimentos, de modo que los datos resultantes serían aquellos que obtenemos mediante la observación pasiva de la realidad. En el primer caso, si no podemos limpiar los datos de las “otras influencias”, tendremos que introducir tales influencias en la teoría. Con ello, el enfoque adoptado por Haavelmo es que el alcance pretendido por la Teoría es el de

explicar los hechos observados sin dar en ello una connotación de descubrimiento de “la verdad”: “... cualesquiera que sean las “explicaciones” que se prefieran, no se debe olvidar que todas ellas son nuestras propias invenciones artificiales en la búsqueda por lograr un entendimiento de la vida real; no son verdades ocultas que han de ser descubiertas” (Haavelmo 1944, pág. 3).

Con ello, las relaciones que describe la Teoría Económica no pueden ser exactas y lo que se desea son “teorías que sin involucrarnos en contradicciones lógicas directas, afirmen que las observaciones se agruparán como norma general en un subconjunto limitado del conjunto de todas las observaciones concebibles, al tiempo que sea consistente con la teoría que de tanto en tanto una observación esté fuera de ese subconjunto” (Haavelmo 1944, pág. 40).

Fijado así el objetivo teórico, Haavelmo concluye que “por lo que sabemos, el esquema de probabilidad y variables aleatorias es hoy por hoy, el único esquema adecuado para formular tales teorías” (Haavelmo 1944, pág. 40).

1.4.- Formulación probabilística de las teorías económicas

En cuanto al realismo de las distribuciones probabilísticas utilizadas en el análisis de datos económicos, Haavelmo con mucha certeza afirma que no es necesario que las probabilidades sean reales para poder considerar las observaciones económicas como realizaciones de variables aleatorias, basta con que sean concebibles. Lo importante “no es si las probabilidades existen o no, sino si procediendo como si existieran (concibiéndolas en la mente), somos capaces de realizar afirmaciones sobre fenómenos reales que son correctas para fines prácticos” (Haavelmo (1944), pág. 43, el término entre paréntesis es añadido). En este sentido, Haavelmo insiste en que la experiencia demuestra que la pura noción hipotética de distribución de probabilidades es útil para derivar afirmaciones de que la probabilidad de que una observación futura caiga en un determinado subconjunto de valores es casi uno.

En este enfoque probabilístico aparece el tema de cómo es posible formular una teoría económica en términos probabilísticos si las observaciones pasivas de la

realidad proceden de circunstancias que no se mantienen estables. Tal cuestión se plantea por la creencia errónea de que el esquema probabilístico y de variables aleatorias no es bastante general; Haavelmo dedica las secciones 9 a 11 de su trabajo a explicar la generalidad del marco estadístico.

Así, Haavelmo (pág. 46) señala lo siguiente. Supóngase un conjunto de variables aleatorias, x_1, x_2, \dots, x_r , que se corresponden, por ejemplo, x_1 , con el consumo individual en un determinado bien y x_2, \dots, x_r , con las variables que influyen en las decisiones de dicho consumo. Sobre ellas tenemos una observación $(x_{11}, x_{21}, \dots, x_{r1})$ correspondiente al individuo 1, otra observación $(x_{12}, x_{22}, \dots, x_{r2})$ correspondiente al individuo 2 y así sucesivamente hasta una observación $(x_{1s}, x_{2s}, \dots, x_{rs})$ correspondiente al individuo s. El sistema de valores de dimensiones $r \times s$

$$\begin{pmatrix} x_{11} & x_{21} & \dots & x_{r1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{1s} & x_{2s} & \dots & x_{rs} \end{pmatrix} \quad (I.1)$$

puede verse como:

- a) Una muestra de tamaño s sobre una población r-dimensional.
- b) como s muestras de una observación cada una correspondientes a s poblaciones r-dimensionales, y,
- c) como una muestra de una única observación rs-dimensional.

De ellas c) es la interpretación más general y en la que como se verá más adelante, basa Haavelmo su desarrollo probabilístico inicial. En cualquier de estos casos, conocida la ley probabilística poblacional se pueden derivar las leyes probabilísticas correspondientes a funciones de las variables originales.

La conclusión es que la correspondencia entre variables económicas y observaciones pasivas de la naturaleza puede verse como la correspondencia entre universo poblacional y muestra de la teoría probabilística. Es decir, los datos observados son una realización del esquema teórico estocásticos que, en terminología

más actual, podemos denominar proceso generador de datos (PGD). Esta visión probabilística de Haavelmo ha sido plenamente asumida por la Econometría moderna, en la que el PGD es el punto de partida, tal como se discute más adelante, para derivar todos los modelos econométricos.

Así, en el enfoque de Haavelmo los datos tienen mucho que decir y lo que distintos autores han denominado “la revolución de Haavelmo” se puede resumir, como hace Morgan (1990) pág. 248, diciendo que implica cambiar de un enfoque una teoría inatacable y una práctica en la que los procesos de ajustes de los datos se realiza con el fin de convertir a los datos en correctos para tal teoría, a un enfoque consistente en encontrar el modelo adecuado para los datos observados mediante el uso de contrastes estadísticos.

1.5.- La distribución de Haavelmo

Para desarrollar lo anterior hagamos referencia a un conjunto de observaciones como en (1.1), pero en un contexto más general, es decir, sin especificar las variables económicas a las que se refieren y, sobre todo, sin determinar si se trata de series temporales u observaciones sobre agentes individuales. Así, supóngase que –en la terminología de Haavelmo (1944), pág. 69- x_1, x_2, \dots, x_n , es un conjunto de n fenómenos económicos, y $(x_{1t_i}, x_{2t_i}, \dots, x_{nt_i})$ es un vector de observaciones conjuntas de dichos n fenómenos. Sea

$$\begin{pmatrix}
 x_{1t_1} & x_{2t_1} & \dots & x_{nt_1} \\
 x_{1t_2} & x_{2t_2} & \dots & x_{nt_2} \\
 \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot \\
 x_{1t_N} & x_{2t_N} & \dots & x_{nt_N}
 \end{pmatrix} \tag{1.2}$$

un sistema de N observaciones correspondientes a los n fenómenos. El indicador $t_i = (i = 1, \dots, N)$ puede referirse a tiempo o a individuos. El sistema de nN valores precedente, puede considerarse –al igual que se hizo en el epígrafe anterior- “como un punto muestral E en el espacio muestral nN variables aleatorias $(x_{1t}, x_{2t}, \dots, x_{nt})$,

$t = t_1, \dots, t_N$, “ con una cierta función de distribución conjunta $P(\omega)$, en donde ω es un punto arbitrario en dicho espacio muestral. A tal supuesto Haavelmo lo denomina: “supuesto fundamental sobre la naturaleza de los datos económicos”, ya que “realmente es difícil concebir un caso en contradicción con este supuesto” (Haavelmo 1944, pág. 70).

Es decir, el sistema de (nN) valores observados que constituyen una muestra son vistos como una realización de nN variables aleatorias que siguen una determinada función de distribución conjunta, a la que Spanos (1989) denomina la distribución de Haavelmo.

En ella Haavelmo basa todos los problemas que aborda en su monografía, como identificación, estimación, contraste de hipótesis y predicción. Así, por ejemplo, en la pág. 85 hablando del problema de identificación concluye: “...todo recae en una única y misma cosa, a saber, el estudio de las propiedades de la distribución de probabilidad conjunta de las variables aleatorias (observables) en un sistema de ecuaciones estocásticas”.

En la Cowles Commission el “supuesto fundamental” de Haavelmo se vio como una gran aportación. A este respecto, Epstein (1987) citando una comunicación del 17 de junio de 1946 de Marschak a Willets, señala que Marschak al ser nombrado director de investigación de la Cowles Commission planificó inmediatamente “emplear todos los recursos de la comisión para desarrollar el trabajo de Tinbergen a la luz de las investigaciones de Haavelmo y de Mann y Wald”. En consecuencia con todo ello, tal como observa Morgan (1990) pág. 254, nota 17, la Cowles Commission quiso publicar el trabajo de Haavelmo (1944) y de Mann y Wald (1943) en un volumen, como una de sus monografías, pero la escasez de papel en ese tiempo de guerra les llevó a publicar ambos trabajos en *Econometrica*, pues en aquel tiempo su publicación se organizaba en la Cowles Commission. En cualquier caso, este reconocimiento sobre el trabajo de Haavelmo estuvo bien patente dentro de la Cowles Commission y se recoge, por ejemplo, en su monografía diez (Koopmans, 1950) al hablar en el capítulo segundo, sección 1.2 de la “Especificación de la distribución de todas las variables”.

En términos prácticos, la distribución de Haavelmo puede considerarse como lo que en el punto anterior se ha denominado el “proceso generador de los datos”.

Tal como han señalado diversos autores es sorprendente que a pesar de la influencia de Haavelmo en la Cowles Commission, la econometría teórica y aplicada posterior no ha tomado durante muchos años como punto central en la investigación a dicha distribución. Spanos (1989) apunta que probablemente se debió a cómo se percibió el resultado de la discusión entre Koopmans, perteneciente a la Cowles Commission, y los investigadores del NBER, sobre “Medición sin teoría”, a raíz de la publicación en 1946 del libro “Measuring Business Cycles” de Burns y Mitchell, pertenecientes al NBER⁵. Así, Koopmans, que consideraba el enfoque probabilístico de Haavelmo como una gran aportación, contribuyó sin pretenderlo a que tal enfoque se ignorase. No obstante, más recientemente ha habido una vuelta a Haavelmo como quedará claro en el epígrafe siguiente. Este redescubrimiento de Haavelmo está en la orientación econométrica denominada LSE. Artículos sobre la aportación de Haavelmo o la conexión de la metodología mencionada con él se encuentran en Spanos (1986), (1988), (1989) y Hendry et al. (1989). Así mismo, en Raymond (1992) se realiza una reflexión profunda sobre la evolución histórica de los modelos econométricos, en la que se analiza la influencia de la Cowles Commission y, en concreto, la importancia que en ella tuvieron los trabajos de Haavelmo.

1.6.- El proceso de reducción de Haavelmo

Haavelmo reconoce que “esta formulación del problema de contrastación de relaciones económicas es muy general”, en consecuencia señala que, con el fin de formular contrastes que no sean triviales, es necesario imponer restricciones adicionales...”.

Para ello indica que la función de distribución conjunta, dígame p_3 ⁶, de todas las variables se puede expresar como el producto de la función de distribución condicional conjunta (p_1) de un punto mN ($m < n$) variables (“dependientes”) con respecto las $(n-m)N$ restantes (“autónomas”, que se pueden considerar constantes en

⁵ Para el seguimiento de esta discusión véase, además de Burns y Mitchell (1946), Koopmans (1947), Vining (1949), Koopmans (1949) y Hastay (1951). Sobre esta discusión son interesantes los comentarios en Morgan (1990), que sugiere que no existía tanto desacuerdo entre las instituciones involucradas en la discusión y que esta tuvo también fines propagandísticos para asegurarse el soporte financiero que la fundación Rockefeller daba a ambas instituciones.

⁶ En los epígrafes siguientes a las funciones de distribución se les denomina por $D(\cdot)$. Aquí, sin embargo, se emplea la misma terminología, p_1, p_2, p_3 , de Haavelmo (1944).

muestras repetidas), por la función de distribuciones marginal (p_2) de este último conjunto de variables.

$$p_3 = p_1 \cdot p_2. \quad (1.3)$$

La función de distribución condicional conjunta implica que las variables dependientes y autónomas se relacionan a través de determinadas funciones f_{jti} –que pueden ser las mismas para cualquier t - en las que aparecen unas variables aleatorias, residuos, que podemos denominar ε_{jti} . Esto a su vez implica que conociendo las funciones f y realizando un supuesto sobre la función de distribución conjunta de las variables ε_{jti} condicional a las variables autónomas, se obtiene p_1 .

En este contexto, la función de distribución de Haavelmo se reduce y se obtiene una función p_3 más simple a través de restricciones sobre las formas f y sobre la función de distribución de las variables ε_{jti} . En este punto Haavelmo señala que “la principal tarea de la teoría económica es realizar elecciones provechosas de las formas f ” (pág. 71). Obsérvese que tales elecciones suponen también escoger un conjunto informativo condicional pertinente, aspecto sobre el que volveremos más adelante.

En cuanto al proceso reductor, Haavelmo pone el siguiente ejemplo (pág. 72). Supóngase que en (1.2) n es igual a dos y que x_2 es una variable autónoma. Con ello, las formas funcionales f se definen implícitamente sobre las variables x_{jti} y, en general serán del tipo.

$$f_{ti}(x_{1ti}, x_{1ti-1}, \dots, x_{1t1}; x_{2ti}, x_{2ti-1}, \dots, x_{2t1}; (X_0); \alpha_0, \alpha_1, \dots, \alpha_k; \varepsilon_{1ti}) = 0, \quad (i= 1, 2, \dots, N)$$

(1.4)

en donde $\alpha_0, \alpha_1, \dots, \alpha_k$ son constantes y (X_0) representa todas las condiciones iniciales, que si entran en una ecuación se suponen constantes. Si en tal caso las formas funcionales cumplen las restricciones de ser: (1) independientes de t_i , es decir, existe una única f , (2) lineales y (3) representar una relación estática, es decir, (3a) sólo variables x referencias a un mismo momento temporal entran en f y (3b) variables aleatorias x_{2ti} y ε_{1ti} son independientes de los valores previos de las variables x y de ε , tendremos:

$$x_{1ti} - \alpha_0 - \alpha_1 x_{2ti} - \varepsilon_{1ti} = 0. \quad (i=1, \dots, N) \quad (1.5)$$

Si ahora sobre las variables ε se introducen las restricciones de que (4a) se distribuyen independientemente de las variables x_{2ti} y (4b) $p_1(\varepsilon_{1t1}, \dots, \varepsilon_{1tN})$ es la función de densidad conjunta de las N variables ε . Entonces, para valores dados de x_{2ti} , las variables x_{1ti} tienen la siguiente función de densidad conjunta $p_1[(x_{1t1} - \alpha_0 - \alpha_1 x_{2t1}), (x_{1t2} - \alpha_0 - \alpha_1 x_{2t2}), \dots, (x_{1tN} - \alpha_0 - \alpha_1 x_{2tN})]$.

Con ello, sea cual sea la función de densidad de las variables p_2 , la función de densidad conjunta de todas las variables vendrá dada por p_3 en (1.3).

El ejemplo anterior pone de manifiesto que el problema de estimar los parámetros en la relación (1.4) se puede reducir el problema de estimar los parámetros de una función de densidad de dimensión N - p_1 -, “mediante un punto muestral asociado con tal ley probabilística”. En realidad, “no se puede dar una descripción más completa de las interconexiones entre un cierto número de variables aleatorias que aquella que está contenida en su distribución conjunta”, por lo que, “la distribución conjunta de todas las variables aleatorias observables en un sistema económico es la única base general para estimar los parámetros desconocidos del sistema” (pág. 88).

Para ilustrar la importancia de la Teoría económica en las formulaciones de f consideramos que un analista establece la relación entre consumo y renta de acuerdo con el ejemplo anterior, de modo que x_{1ti} y x_{2ti} son observaciones temporales de consumo y renta sobre un determinado agregado de individuos. La ecuación (1.5) es

una formulación simple sobre la determinación del consumo. Supóngase que para estimar tal relación dicho analista introduce la hipótesis (4.a), con lo que asumiendo normalidad en (4.b) obtiene que la estimación de los parámetros en la relación teórica es equivalente a estimar la esperanza matemática de x_{1ti} condicional a x_{2ti} .

Esta esperanza matemática condicional es:

$$E(x_{1ti}/x_{2ti}) = a_0 + a_1x_{2ti},$$

donde

$$a_0 = \mu_1 - a_1\mu_2 \quad y$$

$$a_1 = \sigma_{12}/\sigma_2^2,$$

siendo μ_1 y μ_2 las medidas poblacionales de x_1 y x_2 , σ_{12} la covarianza entre x_1 y x_2 , σ_2^2 la varianza de x_2 . Es decir, con los supuestos realizados por este hipotético analista los parámetros α_0 y α_1 del modelo (1.5) coinciden con los parámetros a_0 y a_1 derivados de los primeros y segundos momentos de la función de distribución conjunta de x_1 y x_2 .

Sin embargo, la teoría indica que la hipótesis (4.a) es incorrecta y, por lo tanto, el modelo (1.5) es inaceptable para este problema. Así, si se aborda ahora el problema de acuerdo con la teoría Keynesiada tendríamos⁷:

$$x_{1ti} = \alpha_0^* + \alpha_1^*x_{2ti} + \varepsilon_{ti} \quad y \quad (1.6.a)$$

$$x_{2ti} = x_{1ti} + x_{3ti}, \quad (1.6.b)$$

en donde x_{3ti} es la inversión. Es decir, esta teoría implica que: (1) el análisis debe realizarse teniendo en cuenta una tercera variable económica, la inversión y (2), los modelos teóricos (1.6.a) y (1.6.b) llevarán a las formas funcionales:

⁷ El desarrollo que sigue está tomado de Goldberger (1990).

$$x_{1ti} = (\alpha_0^* + \alpha_1^* x_{3ti} + \varepsilon_{ti}) / (1 - \alpha_1^*) \quad (1.7)$$

$$x_{2ti} = (\alpha_0^* + x_{3ti} + \varepsilon_{ti}) / (1 - \alpha_1^*). \quad (1.8)$$

Si se supone ahora que:

$$E(x_{3ti}) = \mu_3, \text{var}(x_{3ti}) = \sigma_3^2, E(\varepsilon_{ti}) = 0 \quad \text{y}$$

$$\text{Var}(\varepsilon_{ti}) = \sigma_\varepsilon^2$$

y se realiza la hipótesis (4.a) respecto x_{3ti} , es decir, $\text{Cov}(x_{3ti}, \varepsilon_{ti}) = 0$ y se supone normalidad en (4.b), se tienen que x_1 y x_2 siguen una función de densidad conjunta binormal.

En ella

$$E(x_{1ti}/x_{2ti}) = [\theta\alpha_0^* - (1 - \theta)\mu_3] + [\theta\alpha_1^* + (1 - \theta)] x_{2ti}, \quad (1.9)$$

donde $\theta = \sigma_3^2 / (\sigma_3^2 + \sigma_\varepsilon^2)$.

Obsérvese que la teoría Keynesiana está señalando que el parámetro de interés no es $a_1 = \sigma_{12}/\sigma_2^2$, la pendiente de la esperanza matemática formulada en (1.9), sino el parámetro α_1^* .

Este parámetro entra en la expresión de una esperanza matemática condicional, pero distinta a (1.9). En efecto, en (1.6) tomemos

$$E(x_{1ti}/x_{3ti}) = \alpha_0^* + \alpha_1^* E(x_{2ti}/x_{3ti}), \quad (1.10)$$

denominemos $x_{2ti}^* = E(x_{2ti}/x_{3ti}) = \pi_1 + \pi_2 x_{3ti}$. Dado que x_{2ti}^* es una función uno a uno de x_{3ti} , si $\pi_2 \neq 0$, se puede escribir

$$E(x_{1ti}/x_{2ti}^*) = \alpha_0^* + \alpha_1^* x_{2ti}^*. \quad (1.11)$$

Los parámetros a_0 y a_1 derivados de la función de distribución conjunta de x_1 y x_2 , definidos anteriormente se pueden expresar como:

$$a_0 = \theta \alpha_0^* - (1 - \theta) \mu_3 \quad y \quad (I.12)$$

$$a_1 = \theta \alpha_1^* + (1 - \theta), \quad (I.13)$$

con lo que cabe preguntarse cuál es el interés de condicionar respecto a x_2^* y no respecto a x_2 . La respuesta es inmediata si en el sistema (I.6) todos los parámetros constantes menos σ_3^2 tenemos que el parámetro a_1 variará con σ_3^2 , tal como se deduce de (I.12), mientras que α_1^* será estable, tal como se desprende de (I.10), dado que la $\text{Cov}(x_{3ti}, \varepsilon_{ti})$ es cero.

El ejemplo anterior muestra el interés de formular la teoría de acuerdo con el sistema (I.6) y no con el modelo

$$x_{1ti} = E(x_{1ti}/x_{2ti}) + u_{ti},$$

ya que bajo la hipótesis recogida en (I.6) los parámetros de $E(x_{1ti}/x_{2ti})$ no son invariantes respecto a cambios de σ_3^2 . Si definimos un conjunto de parámetros estructurales, como aquel conjunto de parámetros en el que es posible que cambie uno permaneciendo constantes los restantes, tenemos que el objetivo de una teoría es formular relaciones estructurales, es decir, relaciones en las que los parámetros de interés sean parámetros estructurales.

Del ejemplo anterior se deduce que en el sistema Keynesiano α_1^* , que entra en la expectativa condicional (I.11), es un parámetro estructural respecto a cambios en la varianza de la inversión, mientras que a_1 , que entra en $E(x_{1ti}/x_{2ti})$ no es parámetro estructural.

En el proceso reductor se tiene en cuenta también la estructura de los datos, de modo que tal proceso sea lo más restrictivo posible, pero teniendo al mismo tiempo

fuertes razones para creer que el conjunto de hipótesis admisibles, que se contienen sobre los datos en el espacio probabilístico resultante, incluye la hipótesis verdadera. A esto último se puede denominar como el principio de adecuación estadística (véase, por ejemplo, Hendry et al. (1989) y Spanos (1989)). Este principio sobresale a lo largo de todo el capítulo IV en Haavelmo (1944) y supone establecer de forma explícita aquello “que se supone conocido sin lugar a dudas, y aquello que se desea contrastar” (pág. 66). Con ello, aquello que se incorpora como dado-restricciones respecto a la distribución de Haavelmo- en las formas funcionales f anteriores o en p_1 , debe ser compatible con los datos.

En este proceso reductor no sólo se ha pasado de la distribución de Haavelmo a una distribución p_3 extraordinariamente más simple, sino que tal función se ha factorizado de acuerdo con (I.3), de modo que para el análisis que se pretende entre las variables basta con fijarse en la función de distribución condicional conjunta, p_1 . Sobre en qué condiciones es posible basar el análisis en p_1 en vez de en p_3 se discutirá más adelante.

A la función p_3 Spanos (1989) la denomina el modelo estadístico que, en sus palabras, “constituye un resumen conveniente de la información muestral, conveniente en el sentido de que puede estimarse con los datos disponibles y nos permite considerar el modelo teórico en su contexto. Para que tal resumen sea adecuado, los supuestos estadísticos subyacentes que definen el modelo estadístico en cuestión deben ser aceptables para los datos” (pág. 415).

Es cierto que siempre cabrá preguntarse qué sentido tiene contrastar la significación de determinados coeficientes de las formas f cuanto el contexto teórico formal $-p_3$ ó p_1 – es erróneo. Como indica Haavelmo “hablando estrictamente, tal pregunta está siempre justificada cuando tratamos de explicar la realidad mediante un modelo teórico. Pero si seguimos dicha actitud hasta sus últimas consecuencias, nunca lograremos nada respecto la explicación de fenómenos reales” (pág. 81). La consecuencia es que la formulación del modelo estadístico es una cuestión de “conocimiento general e intuición”.

A diferencia de lo ocurrido en las tres o cuatro décadas inmediatamente posteriores a la publicación del trabajo de Haavelmo, en la actualidad los modelos econométricos se especifican a partir de la función p_3 ó p_1 , es decir, poniendo el

énfasis en “los supuestos probabilísticos de los datos y eligiendo adecuadamente el conjunto de información sobre el que se condiciona” (Spanos 1989, pág. 424). Tanto es así, que en el volumen Modelling Economic Series, Granger (1990), (ed.), en el que se recogen 17 trabajos correspondientes a una selección de entre los mejores econométricos actuales, Granger afirma; “la mayoría de los trabajos en esta colección discute cómo caracterizar o aproximar la distribución conjunta de Y_t –variables de interés- y X_t –otras variables utilizadas para explicar Y_t - condicional a su pasado, o a la distribución de Y_t condicional al pasado de Y_t y al pasado (y posiblemente también al presente) de X_t ” (los términos entre guiones son añadidos)⁸.

Desde 1944 la Teoría Econométrica ha avanzado mucho y en distintas orientaciones, pero hay una dirección de progreso que es de interés mencionar aquí. Tal dirección es la que contempla una especificación más general del modelo estadístico, consistente en una de las estas dos alternativas: (a) no realizar ningún supuesto sobre la forma funcional de la función de distribución, lo que deriva en modelos no paramétricos; (b) formular solamente ciertos aspectos de dicha función, concluyendo con modelos semi-paramétricos. Ambos modelos se comentarán más adelante, pero vale la pena subrayar aquí que en este tipo de Econometría la función de distribución de los datos tiene, si cabe, todavía mayor protagonismo.

Como conclusión a esta sección tenemos que los fundamentos estadísticos de los modelos econométricos se encuentran en la función de distribución conjunta de los datos. No obstante, esta función resulta ser excesivamente compleja y es necesario simplificarla para que la inferencia con una única observación de las variables sea operativa. Este proceso simplificado debe realizarse según las orientaciones de la teoría y debe concluir con una función de distribución capaz de recoger las principales características de los datos.

⁸ Obsérvese que Granger denomina X_t a lo que en la sección siguiente nosotros –siguiendo una terminología más generalizada en la literatura econométrica- denominamos Z_t .

II. PANORÁMICA GENERAL SOBRE LOS MODELOS ECONOMETRÍCOS

Para el desarrollo del proyecto docente no solamente es necesario dejar asentados los fundamentos estadísticos de la Econometría, sino que es conveniente describir una panorámica general que abarque la mayoría de los modelos econométricos, de modo que se vea cuál es su jerarquización y sobre todo las relaciones entre ellos. Para tal fin es muy conveniente enlazar con el último epígrafe de la sección anterior y empezar aquí dando una sistematización del proceso reductor de Haavelmo. Tras ello la presentación general de los diferentes modelos econométricos será más ilustrativa.

II.1.- Sistematización del proceso reductor

El proceso reductor de la sección anterior tiene que basarse en la teoría, y ésta es la forma que en el Cowles Commission se veía el papel de la teoría en la medición de los fenómenos reales. Un papel dentro de un contexto estocástico general capaz de explicar tales fenómenos reales. Curiosamente, como se ha señalado antes, posteriormente este último aspecto se olvidó en la investigación econométrica.

A su vez, el proceso reductor tiene que realizarse en función de las características de los datos, de modo que se concluya con un modelo estadístico capaz de explicar las observaciones pasivas que nos ofrece la Naturaleza. Resulta pues importante sistematizar tal proceso reductor. Para ello seguiremos a Ericsson y Hendry (1985), Spanos (1989) y Ericsson, Campos y Tran (1990), que realizan una organización de ideas sugeridas dentro de la tradición econométrica de la London School of Economics, véase, por ejemplo, Sargan (1964), Engle et al. (1983), Hendry y Richard (1982), (1983), Hendry (1983), Hendry et al. (1984), Hendry y Wallis (1984), Florens y Mouchart (1985a), (1985b), Ericsson y Hendry (1985), Spanos (1986), Hendry (1986), (1987), Gilbert (1986, 1989), Phillips (1988), Campos, Ericsson y Hendry (1990), Hendry y Ericsson (1991)⁹, etc. Este enfoque econométrico vendrá recogido en el texto anunciado desde hace bastante tiempo, pero todavía sin aparecer: Lectures on Econometric Methodology, del profesor Hendry (1992) y Cointegration, Error-Correction and the econometric analysis of non-stationary data, de Banerjee, Dolado, Galbraith y Hendry (1992).

⁹ La versión inicial de este trabajo data de 1983 con revisiones en 1985 y 1987 y el trabajo de Campos et al. (1990) se basa a la vez en material de estas versiones iniciales de Hendry y Ericsson (1991).

Los pasos a seguir en el proceso reductor los podemos clasificar de la siguiente forma:

- (0) El proceso generador de datos (PGD).
- (1) Marginalización del PGD respecto a las variables que no entran en el estudio del problema considerado.
- (2) Condicionamiento según la recursividad temporal.
- (3) Restricciones de homogeneidad.
- (4) Restricciones distribucionales.
- (5) Restricciones en la dependencia de los datos (memoria).
- (6) Condicionamiento contemporánea.

II.1.0.- El proceso generador de datos.

La estrategia de reducción que se desea sistematizar parte del proceso generador de datos, que en un sentido amplio puede definirse –véase Granger (1990) págs. 6 a 8- como la distribución conjunta de todas las variables que determinan y representan todas las decisiones de todos los agentes económicos. Estas decisiones básicas constituyen millones y millones de variables, por lo que en el análisis económico se procede a trabajar con variables que han sufrido un proceso previo de agregación sobre individuos –agregación transversal- o a lo largo del tiempo –agregación temporal.

En un determinado estudio empírico estas variables agregadas se pueden agrupar en tres conjuntos: (1) el vector Y_t , que recoge aquellas variables que se requieren explicar y que podemos denominar variables de interés, (2) el vector Z_t , que recoge aquellas variables que se utilizan para explicar Y_t y (3) S_t el vector correspondiente a las variables restantes.

Si se denomina W_t al vector

$$W_t = (Y_t', Z_t', S_t')'$$

tenemos que el verdadero proceso generador de datos definido anteriormente se puede aproximar por la distribución conjunta de todos los vectores W_t ;

$$D_w(W_1, \dots, W_T/W_0; \psi^{(t)}), \quad (II.2)$$

donde T es el número de observaciones sobre W , W_0 representa las condiciones iniciales y $\psi^{(t)}$ el vector de parámetros relevante, que para recoger el hecho de que sus elementos corresponden a parámetros que pueden cambiar por t , representamos el vector $\psi^{(t)}$. Esta función puede identificarse con la función de distribución de Haavelmo. Esta función es extraordinariamente amplia y compleja por lo que es necesario aplicar sobre ella el proceso reductor

II.1.1.- Marginalización del PGD respecto a las variables que no entran en el estudio del problema considerado.

Del vector W_t , sólo se van a considerar en el análisis las variables Y_t y Z_t que se pueden agrupar en el vector

$$X_t = (Y_t', Z_t')', \quad (II.3)$$

por tanto, el modelo econométrico sobre X_t se ha de basar en marginalizar $F_w(\cdot)$ respecto a las variables S_t . Con ellos se obtiene la función de distribución de $X = (X_1, X_2, \dots, X_T)$ como:

$$D_x(X_1, X_2, \dots, X_T/X_0; \theta^{(t)}) = \int D_w(W_1, \dots, W_T/W_0; \psi) dD_s\{S_0, \dots, S_T\}, \quad (II.4)$$

donde $\theta^{(t)}$ es un vector de funciones del vector de parámetros $\psi^{(t)}$.

II.1.2.- Condicionización según la recursividad temporal

En variables observadas a lo largo del tiempo, parece natural suponer que el futuro no influye en el pasado, con lo que utilizando la notación

$$X_t^j = (X_j', X_{j+1}', \dots, X_t')', \quad t > j \geq 1 \quad (II.5)$$

X_T^1 recoge todos los datos de la muestra, y

$$X_{t-h}^+ = \begin{pmatrix} X_0 \\ X_{t-h}^1 \end{pmatrix}, \quad (II.6)$$

es una submuestra que incluye las condiciones iniciales X_0 , con lo que se puede escribir

$$D_x(X_T^1/X_0; \theta^{(t)}) = D(X_T/X_{T-1}^1, X_0; \lambda_T) \cdot D(X_{T-1}^1/X_0; \delta^{(t)}) \quad (II.7)$$

y repitiendo este proceso de factorización respecto a $X_{T-1}^1, X_{T-2}^1, \dots, X_1$ se obtiene

$$D_x(X_T^1/X_0; \theta^{(t)}) = \prod_{t=1}^T D(X_t/X_{t-1}^+; \lambda_t) \quad (II.8)$$

En (II.8) el vector $\lambda^* = (\lambda'_1, \dots, \lambda'_T)'$, es una función de $\theta^{(t)}$.

Los pasos (1) y (2) en el proceso reductor son bastantes neutros en tanto en cuanto las variables S_t realmente no afecten al análisis, ya que una causalidad del pasado hacia el futuro, y no en sentido inverso, parece aceptable. En los pasos restantes las restricciones que se incorporan son menos neutras. Sin embargo, ciertas restricciones son imprescindibles para que sea posible la inferencia a partir de los datos económicos utilizados. En consecuencia, el tipo de restricciones que realmente se implementen deberán –tal como se insistió en la sección anterior- ser compatibles con los datos. Por tanto, el grado de restricción a aplicar en cada análisis, correspondiente a cada uno de los pasos siguientes del proceso reductor será distinto. El principio que debe guiar las reducciones es el de que el modelo estadístico resultante sea compatible con la teoría y adecuado para explicar las propiedades esenciales de los datos.

II.1.3.- Restricciones de homogeneidad

En (II.8) las funciones de densidad condicional son heterogéneas en el tiempo y dado que hay una única observación de todas las variables, la reducción es imprescindible.

Usualmente la restricción que se incorpora es que los parámetros λ_t no son dependientes del tiempo, es decir, $\lambda_t = \lambda$. Con ello,

$$D_x(X_T^1/X_0; \theta) = \prod_{t=1}^T D(X_t/X_{t-1}^+; \lambda). \quad (II.9)$$

Para que esta hipótesis de homogeneidad sea compatible con los datos es importante que en la marginalización no se hayan eliminado variables de interés. En efecto, la marginalización es siempre posible sean cuales sean las variables sobre las que se marginalice. No obstante, la marginalización debe hacerse respecto a las variables que no influyen en el análisis que se pretende, ya que de lo contrario los parámetros del modelo resultante pueden no ser invariantes y, por tanto, no serán parámetros estructurales. Para ilustrarlo cojamos un ejemplo de Goldberger (1991) sección 31.5. Así, considérese el modelo (II.23) con τ_j y L_j constantes y dos variables explicativas (z_1, z_2) . En tal caso el proceso generador de datos incluye a las variables (y, z_1, z_2) , con lo que con las hipótesis realizadas para obtener (II.23) se obtiene ahora

$$E(y_j/z_{1j}, z_{2j}) = c + \beta_1 z_{1j} + \beta_2 z_{2j}$$

y, en consecuencia

$$y_j = c + \beta_1 z_{1j} + \beta_2 z_{2j} + a_j.$$

Sin embargo, marginalizando respecto a z_2 se obtiene, bajo las hipótesis empleadas, que

$$E(y_j/z_{1j}) = c^* + \beta_1^* z_{1j}$$

con lo que el modelo

$$y_j = c^* + \beta_1^* z_{1j} + a_j^*$$

es también correcto. El problema con este último modelo es que el parámetro β_1^* , que es igual a $\beta_1 + \phi\beta_2$, donde ϕ es el parámetro de regresión de z_2 respecto a z_1 , no es invariante ante cambios en la covarianza entre z_1 y z_2 . Pero puede ocurrir que el PGD sea tal que cambios en dicha covarianza no alteran a la esperanza matemática de y condicional a z_1 y z_2 , en tal caso los parámetros de la regresión simple no serán invariantes mientras que los de la regresión ampliada sí.

Obsérvese que esta restricción recogida en (II.9) normalmente implicará que los coeficientes de las esperanzas matemáticas condicionales son constantes, pero abarca también la situación conocida en la literatura como modelos con coeficientes que varían de forma estocástica, pues entonces en tales modelos existen unos “meta-parámetros” que son constantes, y se recogen en el vector λ .

En modelos dinámicos la expresión (II.9) se cumple con la hipótesis de estacionariedad. La falta de estacionariedad tendrá que abordarse en función del tipo de no-estacionariedad presente en los datos. Así, si ésta consistiese en la presencia de un vector de medias determinísticas no estacionarias, esto implicaría la sustitución del subvector de parámetros en el vector λ que recoge las ordenadas en el origen, por funciones adecuadas de tiempo. Una no estacionariedad debida a raíces unitarias implicará muchas veces una reparametrización del modelo en términos de, modelos en diferencias, modelos con mecanismos de corrección del error, sistemas cointegrados (Phillips (1991)) o modelos estructurales los componentes no observables (Harvey (1989)). En los dos primeros casos la formulación del modelo es directamente sobre variables estacionarias y en el tercer caso es también posible una reformulación en tal sentido. Así, los modelos con raíces unitarias pueden reformularse y verse como modelos sobre variables transformadas que son estacionarias. En estos casos los fundamentos para la inferencia pueden ser más complejos, pero en tanto en cuanto el sistema se estime incorporando la información sobre las raíces unitarias presentes, los estimadores resultantes tienen propiedades asintóticas óptimas, y los contrastes de hipótesis pueden realizarse mediante los procedimientos habituales (véase Phillips (1991) y Banerjee et al. (1992)). La falta de estacionariedad en los segundos momentos inducirá, normalmente, que los parámetros del modelo dependan del tiempo y habrá que afrontarlo mediante esquemas variables en función de unos meta-parámetros estables.

Los modelos estructurales de componentes no observables son ampliamente discutidos y analizados en Harvey (1989). En este caso, la formulación del modelo es directamente sobre los niveles de variables X_t , que se definen en términos de componentes fundamentales como pueden ser, en los casos más sencillos, la tendencia y la estacionalidad. Estos componentes se modelizan como procesos estocásticos.

Para solucionar el problema de la no estacionariedad, el modelo se especifica en forma de modelo del espacio de los estados. Bajo el supuesto de normalidad de las densidades condicionales, el filtro de Kalman (Kalman (1960)) proporciona las correspondientes medias y varianzas condicionales, obteniéndose en consecuencia, las $D(X_t/X_{t-1}^+; \lambda)$ en (II.9). En el caso de que las densidades condicionales no sean normales, el filtro de Kalman aún puede ser utilizado, proporcionando estimaciones de error cuadrático medio mínimo entre las situaciones lineales de las medias y varianza condicionales.

II.1.4.- Restricciones distribucionales

Generalmente la restricción utilizada es la de expectativas condicionales lineales procedentes de una función de densidad normal. En bastantes casos en los que los datos no cumplen tal restricción, una transformación de los mismos, mediante el empleo de logaritmos, ratios, transformaciones de Box-Cox, puede hacer que los datos transformados cumplan la hipótesis de linealidad condicional normal. En tal caso $D(X_t/X_{t-1}^*; \lambda)$ en (II.9) se puede formular como

$$X_t/X_{t-1}^+ \sim N(\mu, \Omega) \quad (\text{II.10})$$

y el vector λ que se recoge en (II.9) incorpora los parámetros del vector μ y de la matriz Ω que no son redundantes.

En la formulación de las restricciones distribucionales es conveniente establecer dos niveles. (a) la forma de la función de la esperanza matemática condicional, con lo que a partir de ella se puede definir

$$a_t = X_t - E(X_t/X_{t-1}^+), \quad (\text{II.11})$$

que es una innovación respecto a X_{t-1}^+ y (b) la función de distribución condicional de las innovaciones. Ciertamente, si $D_x(X_t^1/X_0; \theta)$ sigue una función de distribución multinormal, se cumple que:

- (a) $E(X_t/X_{t-1}^+)$ es lineal, y
- (b) $a_t/X_{t-1}^+ \sim N(0, \Omega)$.

Pero para que se cumpla (a) -condicionalidad lineal- no es necesario la normalidad de $D_x(\cdot)$.

Lo anterior supone que la varianza condicional tiene una forma funcional muy concreta: es constante. Sin embargo, esto puede generalizarse. Así, para el caso del modelo de regresión la generalización se puede conseguir considerando que

$$h_t = E[(X_t - E(X_t/X_{t-1}^+))^2 | X_{t-1}^+],$$

con lo que en (b) se tiene que

$$\frac{a_t}{h_t^{1/2}} \sim N(0,1).$$

No obstante, en lo que sigue supondremos, generalmente, que $h_t = \sigma^2$.

Lo que se lleva dicho pone de manifiesto que la formulación de modelos no lineales y dinámicos a partir de $E(X_t/X_{t-1}^+)$ es posible dentro de este contexto que estamos desarrollando. Sin embargo, en la inferencia sobre tales modelos se presentan enormes problemas a no ser que la función no lineal y la estructura dinámica sean conocidas a priori. Por eso, cuando la linealidad condicional normal no se cumple en los datos –o en una transformación de los mismos– es, en general, muy difícil detectar la estructura no lineal y la función de densidad que los datos requieren. En tales casos se podría proceder sin realizar ningún supuesto sobre la forma funcional de la función de distribución conjunta, a través de modelos no paramétricos. Pero en dichas circunstancias si el número de variables es alto el análisis no-paramétrico requiere un elevadísimo número de observaciones, que no suele estar disponible en variables económicas. Por eso, el futuro de la Econometría no paramétrica es de aplicación muy limitada excepto, quizás, para series financieras de

observación en tiempo continuo, prácticamente. Sin embargo, entre la completa descripción de las funciones de distribución de los datos y la ausencia de supuestos sobre su forma funcional, existe una situación intermedia que consiste en formular ciertos aspectos de la función de distribución, por ejemplo, la esperanza matemática condicional, y no otros, pero continuando estando interesado en ellos, con lo que concluye en modelos semiparamétricos. La aplicación de la Econometría semiparamétrica requiere muchas observaciones, pero menos que la correspondiente formulación no-paramétrica, por lo que su implantación es factible, sobre todo con datos financieros intra-diarios o muestras de corte transversal sobre miles de agentes económicos. Ciertamente las aplicaciones econométricas semiparamétricas son muy escasas, no obstante Robinson (1991) se comentan ciertas aplicaciones de sus colaboradores. Sin embargo, un gran obstáculo actual para su aplicación es el desconocimiento de la técnica entre los economistas. En este sentido Delgado y Robinson (1992) presentan un repaso muy útil de los métodos no-paramétricos y semiparamétricos “con la esperanza de que a medida que la metodología sea más conocida y mejor entendida, pasará a ser más usada por los economistas en sus aplicaciones empíricas”.

El desarrollo que sigue será exclusivo sobre la Econometría paramétrica.

II.1.5.- Restricciones en la dependencia de los datos (memoria)

La hipótesis utilizada en el modelo de regresión clásico de que las observaciones son independientes, no se aplica, ciertamente, a las series temporales, por lo que tal hipótesis de independencia debe substituirse por alguna forma de independencia asintótica. A este tipo de restricciones Spanos (1986) las clasifica en una categoría que denomina “memoria”. En la presencia de raíces unitarias tal independencia asintótica se refiere a las variables transformadas que, como se ha indicado, son estacionarias.

Una forma de introducir esa independencia asintótica es truncado la longitud de los retardos sobre los cuales se condiciona en (II.9), es decir, aproximado en (II.10) X_{t-1}^+ por X_{t-1}^{t-l} , en donde l es el orden del mayor retardo considerado. Con ello (II.9) se convierte en

$$D_x(X_T^1/X_0; \theta) = \prod_{t=1}^T D(X_t/X_{t-1}^{t-l} ; \lambda). \quad (II.12)$$

Otra forma más general de introducir esa independencia asintótica es mediante la hipótesis de ergodicidad e invertibilidad. En procesos estacionarios gaussianos la ergodicidad implica que las autocorrelaciones o correlaciones entre variables tienden a cero a medida que el desfase entre las variables tiende a infinito y la invertibilidad implica que eso se puede obtener mediante formulaciones autorregresivas, posiblemente de orden infinito, pero convergentes.

Como cuestión de notación y cuando sea conveniente explicitar que el condicionamiento respecto al pasado cumple las hipótesis de ergodicidad e invertibilidad, pondremos el superíndice “e” sobre el conjunto respecto al que se condiciona.

Así, con tal supuesto, la expresión (II.9) toma la forma:

$$D_x(X_T^1/X_0; \theta) = \prod_{t=1}^T D(X_t/X_{t-1}^{+e}; \lambda). \quad (\text{II.13})$$

Con las hipótesis recogidas en II.1.1 a II.1.5 el modelo

$$X_t = E(X_t/X_{t-1}^{+e}) + e_t, \quad (\text{II.14})$$

que se deriva de la distribución conjunta $D_x(X_T^1/X_0; \theta)$, es operativo y puede estimarse con una única realización de variables.

Sin embargo, en $D_x(X_T^1/X_0; \theta)$ es posible continuar el proceso de condicionalización, pero realizándolo ahora sobre variables contemporáneas, tal como se ilustra en el punto siguiente.

II.1.6.- Condicionalización contemporánea.

Utilizando la notación introducida en (II.3), es posible condicionalizar Y_t , respecto al pasado, X_{t-1}^+ , y a los valores contemporáneos de Z_t . Así,

$$D_x(X_T^1/X_0; \theta) = D_x(Y_T^1, Z_T^1/X_0; \theta) = \prod_{t=1}^T D(Y_t, Z_t/X_{t-1}^{+e}; \lambda) =$$

$$\prod_{t=1}^T D(Y_t/Z_t, X_{t-1}^{+e}; \lambda_1) D(Z_t/X_{t-1}^{+e}; \lambda_2). \quad (\text{II.15})$$

Es decir, se ha factorizado $D(Y_t, Z_t/X_{t-1}^{+e}; \lambda)$ de la forma:

$$D(Y_t, Z_t/X_{t-1}^{+e}; \lambda) = D(Y_t/Z_t, X_{t-1}^{+e}; \lambda_1) D(Z_t/X_{t-1}^{+e}; \lambda_2). \quad (\text{II.16})$$

La descomposición que aparece en (II.15) es siempre posible y, por tanto, sin ningún tipo de pérdida de generalidad respecto a (II.13).

La utilidad de esta descomposición radica en que X_t es un vector de dimensión n y sólo se está interesado en explicar el comportamiento del subvector Y_t , de dimensión m , en función de otras k ($k = n - m$) variables Z_t . En tal situación, se puede pensar en no abordar la modelización a partir de $D(Y_t, Z_t/X_{t-1}^{+e}; \lambda)$ sino restringirse a realizar la modelización a partir de $D(Y_t/Z_t, X_{t-1}^{+e}; \lambda_1)$ ignorando la función de densidad marginal de Z_t . Esto simplifica el problema de modelización ya que en vez de operar con el sistema de n ecuaciones que se deriva de $D(Y_t, Z_t/X_{t-1}^{+e}; \lambda)$ se opera con un sistema $(n-k)$ ecuaciones. Sin embargo, tal proceder implicará, normalmente, una pérdida de generalidad y el problema que se plantea es señalar las condiciones bajo las que es posible utilizar solamente la función de densidad condicional, $D(Y_t/Z_t, X_{t-1}^{+e}; \lambda_1)$, sin que eso acarree una pérdida de información. Esto conduce al tema de exogeneidad que se trata en la sección II.2.

Con los supuestos incluidos en II.1.5 a II.1.5 se deriva el modelo (II.14), ahora añadiendo el supuesto de condicionalidad contemporánea se obtiene:

$$Y_t = E(Y_t/Z_t, X_{t-1}^{+e}) + e_t. \quad (\text{II.17})$$

II.1.7.- El proceso reductor en datos de corte transversal o de panel.

El proceso reductor ha sido descrito poniendo cierto énfasis en su aplicación a datos de series temporales, pero es igualmente aplicable a datos de corte transversal o de panel. Una interesante discusión para el caso de panel se encuentra en Spanos (1988).

Como ejemplo de proceso reductor con datos de corte transversal, pasamos a ilustrar la obtención de un modelo de regresión para dichos datos.

Supóngase que sobre un vector de n variables

$$X_j = (y_j, z_{1j}, \dots, z_{(n-1)j}) = (y_j, Z_j)$$

se dispone de N observaciones referente a N agente económicos, recogidos en X_N^1 .

La distribución conjunta de los datos será:

$$D_X(X_N^1; \theta),$$

en donde ya se ha producido la marginalización respecto a otras variables S_j que no entran en el análisis. El problema no tiene dimensión temporal por lo que no se aplica el paso (2). Respecto a las restricciones de homogeneidad, y de dependencia los correspondientes supuestos extremos son: las variables están idéntica e independientemente distribuidas. La hipótesis de independencia permitirá la siguiente simplificación:

$$D_X(X_N^1; \theta) = \prod_{j=1}^N D(X_j; \psi_j) \quad (\text{II.18})$$

y la de idénticamente distribuidas

$$\prod_{j=1}^N D(X_j; \psi_j) = \prod_{j=1}^N D(X_j; \lambda^+). \quad (\text{II.19})$$

Ahora bien, este supuesto de identidad de las distribuciones puede ser inadecuado para los datos si proceden de un muestreo estratificado. Así, es posible que las medias de las distribuciones $D(X_j; \psi_j)$ sean función de determinadas características de los agentes, por ejemplo, tamaño (τ), localización (L), etc. Suponiendo que tal función es una combinación lineal del tamaño y la localización, esta heterogeneidad se puede abordar sustituyendo los parámetros del vector λ^+ que

representan las medias comunes de cada variable en la hipótesis de distribuciones idénticas, por medias específicas a cada agente, m_j , que vienen dadas por

$$m_j = c_0 + c_1 \tau_j + c_2 L_j , \quad (II.20)$$

en donde m_j , T_j y L_j son vectores de dimensión n y los elementos de T_j y L_j son parámetros fijos en la distribución de cada agente. En este caso

$$D(X_j; \psi_j) = D(X_j; m_j, \lambda), \quad (II.21)$$

en donde el vector λ no incluye las medias de cada variable, que se recogen en m_j .

Por último falta especificar la condicionalización contemporánea y el tipo de distribución. Si ésta es normal y la teoría nos dice que los parámetros de interés son los de la $E(Y_j/Z_j)$ tendremos que la descomposición relevante es:

$$D(X_j; m_j, \lambda) = D(y_j/Z_j; m_j^*, \lambda_1) D(Z_j; \widetilde{m}_j, \lambda_2) , \quad (II.22)$$

en donde la hipótesis de normalidad asegura que λ_1 y λ_2 son vectores de variación libre al igual que m_j^* y \widetilde{m}_j ; en virtud del supuesto de normalidad y de (II.20). En consecuencia Z_j son variables débilmente exógenas –término que se define en la sección siguiente- por lo que para fines de inferencia sobre λ_1 basta con formular el modelo a partir de $D(Y_j/Z_j; m_j^*, \lambda_1)$. Este modelo será:

$$y_j = c_0^* + c_1^* \tau_j + c_2^* L_j + \sum_{i=1}^{n-1} \beta_i z_{ij} + a_j. \quad (II.23)$$

En el análisis de datos de corte transversal la hipótesis de que la distribución conjunta sea normal es con frecuencia inaceptable, ya que la marginal de Z , muchas veces variables dicotómicas, no lo puede ser. En tales casos la normalidad de la condicional puede ser aceptable, con lo que si la distribución marginal no es normal se puede tener que la $Var(y_j/Z_j)$ no sea constante. Así, si la varianza de a_j es función de tamaño τ_j y esa función es conocida, se pueden transformar las variables en (II.23) de modo que para las variables transformadas se cumpla la hipótesis de normalidad. Un tratamiento más complejo en esta dirección se encuentra en Judge et al. (1985). Obviamente la heteroscedasticidad puede ser menos precisa que lo que el enfoque anterior supone y Spanos (1990) propone un tratamiento paramétrico basado en una familia de distribuciones simétricas, pero leptokurticas. No obstante, este enfoque puede ser también restrictivo y un tratamiento más general puede consistir en la aplicación de métodos semiparamétricos, en los que, por ejemplo, se utilice la

hipótesis de condicionalidad lineal, pero sin precisar la forma funcional de la distribución de los residuos. En este sentido puede verse Delgado (1992).

Con los enfoques anteriores la heteroscedasticidad se trata directamente y los estimadores resultantes son eficientes. Una alternativa a lo anterior es la propuesta por White (1984), que sin pretender tratar la heteroscedasticidad, estima adecuadamente la varianza de los estimadores mínimo cuadráticos ordinarios – estimadores que ante la presencia de la heteroscedasticidad son consistentes, pero no eficientes- con lo que se pueden realizar contrastes t adecuados.

II.1.8.- Conclusión.

La discusión precedente ha puesto de manifiesto que los modelos econométricos se derivan de la distribución conjunta de los datos –“supuesto fundamental sobre la naturaleza de los datos económicos”- aplicando un proceso simplificador a partir de una distribución plenamente general. Este proceso incorpora las orientaciones de la teoría económica y a su vez tiene en cuenta las características fundamentales de los datos. Así pues, presenta un contexto general en el que encuadrar los temas de la Teoría Econométrica, dotándolos del fundamento estadístico necesario.

El procedimiento es válido tanto para datos de series temporales como de corte transversal o de panel

En los datos de series temporales es deseable, dada la dependencia que se encuentra en tales datos, una condicionalización basada en la recursividad temporal, con lo que introduciendo solamente los supuestos de homogeneidad, distribucionales y de memoria pertinentes, se puede formular el modelo recogido en (II.14), como modelo más general para el análisis. Este modelo genera unas innovaciones e_t y la desviación estándar de las mismas pasa a ser el punto de referencia con el que comparar las desviaciones estándares innovacionales de otros modelos más restrictivos.

Si a las hipótesis anteriores se añade la de condicionalización contemporánea se obtiene, para cualquier tipo de datos, el modelo (II.17), que en el contexto del análisis condicional contemporáneo para a ser hipótesis general mantenida sobre Y_t .

Todo este proceso pone de manifiesto que, tanto en (II.4) como (II.7) los residuos no son autónomos, sino derivado del proceso generador de los datos, es decir, de las características aleatorias de las variables económicas. En otras palabras, la aleatoriedad genuina o de partida no está en los residuos, que luego se la contagian a las variables, sino en las variables mismas, de modo que de ellas se derivan los residuos aleatorios.

II.2. Exogeneidad y determinación de los parámetros de interés.

En el punto II.1.6 se ha ilustrado cómo la condicionalización contemporánea simplifica el problema. No obstante, con tal condicionalización lo que se pretende principalmente es concentrar la atención en los parámetros de interés que determina la teoría económica, determinación que se realiza, normalmente, buscando aislar parámetros invariantes. Así, por ejemplo, el conjunto de parámetros que aparece en $E(X_t/X_{t-1}^{+e})$ en el modelo (II.4), puede contener elementos variantes, pero puede ocurrir que determinadas funciones de ellos no lo sean. Este intento de aislar invariantes se realiza utilizando los resultados que proporciona la teoría económica, siendo en ello básico la elección del conjunto informativo –dígase D_t - sobre el que se condiciona. Así, tanto (II.4) como (II.7) se pueden representar como:

$$X_t = E(X_t/D_{1t}) + e_{xt} \quad y \quad (II.24)$$

$$Y_t = E(Y_t/D_{2t}) + e_{yt}, \quad (II.25)$$

en donde $D_{1t} = (X_{t-1}^{+e})$ y $D_{2t} = (Z_t, X_{t-1}^{+e})$.

En (II.25) el conjunto informativo relevante no viene sugerido exclusivamente por la teoría, sino también por la naturaleza de los datos, que es quien básicamente ha determinado los supuestos de homogeneidad, distribucionales y de memoria en el proceso reductor. La modelización realizada en (II.24) y (II.25) se obtiene a partir de la naturaleza estocástica de las variables eligiendo apropiadamente el conjunto informativo condicionalmente D_t .

La condicionalización debida a la recursividad temporal no ofrece, como se comentó en su momento, especial controversia, ya que las condicionalizaciones

alternativas llevan a que el presente dependa del futuro. Sin embargo, en la condicionalización contemporánea las condicionalizaciones alternativas pueden importar en un mundo que puede cambiar hacia el futuro, y se trata de elegir la condicionalización con parámetros estructurales.

Para ilustrar el punto anterior volvamos al ejemplo de la sección II.1.7, y supongamos que n es igual a dos y que las m_j son constantes incorporadas en el vector λ . Con ello podemos realizar la descomposición,

$$D(X_j; \lambda) = D(Y_j/Z_j; \lambda_1) D(Z_j; \lambda_2) \quad (\text{II.26})$$

pero igualmente válido es

$$D(X_j; \lambda) = D(Z_j/Y_j; \phi_1) D(Y_j; \phi_2). \quad (\text{II.27})$$

Cuál es la descomposición de interés, tiene que venir determinado por la teoría. Si el mundo se va a mantener estable en el futuro no importa cuál de las dos descomposiciones sea la que se estime, ya que conocidos ϕ_1 y ϕ_2 se determinan λ_1 y λ_2 de forma única y viceversa. Pero si el mundo puede cambiar, de forma que cambie λ_2 , pero se mantenga constante λ_1 ¹⁰, entonces la teoría al apuntar que los parámetros de interés son λ_1 y no ϕ_1 está aislando parámetros invariantes y ya no es válido estimar ϕ_1 para luego calcular λ_1 . Es decir, en general, un objetivo que la teoría considera, entre otros, es aislar los componentes no constantes del vector λ , en el subvector λ_2 que sólo es relevante en la modelización de las variables que quedan fuera del modelo.

La condicionalización contemporánea recogida en (II.16) se realiza para derivar de ella, ignorando la función de densidad marginal de Z_t , el modelo (II.25). Para estudiar bajo qué condiciones del modelo (II.25) se obtiene sin pérdida de generalidad en cuanto a las variables Y_t , que en él se determinan, es necesario discutir

¹⁰ Esto puede ocurrir, por ejemplo, si cambia la varianza de z_j en su distribución marginal pero la covarianza entre y_j y z_j es proporcional a la varianza z_j .

los diferentes conceptos de exogeneidad que se formulan en Engle et al. (1983), que se corresponden a los diferentes objetivos que se pretendan en la modelización.

Este último es muy importante en el análisis econométrico. No existen variables con un status de exógenas. Tal propiedad depende del tipo de análisis que se quiere realizar. Que una variable es exógena implica que tal variable puede considerarse como dada para el propósito que se pretende, pero tal condición de poder ser considerada como dada depende del propósito en el que se esté involucrado. Así, una variable puede ser exógena para unos fines pero no para otros.

Con fines de inferencia condicional, por ejemplo, para contrastar si un determinado parámetro es cero, el concepto relevante es de exogeneidad débil. Para definirlo es necesario precisar sobre qué parámetro o parámetros, denominados de interés, φ , se quiere realizar la inferencia. Con ello se dice que las variables Z en (II.6) son débilmente exógenas para la inferencia sobre los parámetros de interés si:

- (a) los conjuntos paramétricos λ_1 y λ_2 no tienen elementos comunes y además son de variación libre, es decir, no están sometidos a restricciones cruzadas entre ellos,
- (b) el conjunto de parámetros de interés es función exclusivamente λ_1 .

Esta exogeneidad débil va referida por tanto a parámetros de interés, pudiendo darse el caso de que una variable sea exógena respecto a φ_i pero no respecto a φ_j .

Para fines de predicción la exogeneidad débil no es suficiente, ya que si Y_{t-j} influye en Z_t , tal como es posible en la distribución marginal de Z_t en (II.16), ésta no se podrá considerar como dada en la predicción de Y . La exogeneidad en la predicción requiere que la función de distribución conjunta se pueda factorizar de la forma:

$$D(Y_t, Z_t / X_{t-1}^{+e}; \lambda) = D(Y_t / Z_t, X_{t-1}^{+e}; \lambda_1) D(Z_t / X_{t-1}^{+e}; \lambda_2). \quad (\text{II.28})$$

Cuando en la función marginal (respecto al pasado) de Z_t no interviene, tal como aparece en (II.28), Y_{t-1}^+ , se dice que Y no causa, en la definición de Granger, a Z . Así, una variable es fuertemente exógena si, además de serlo débilmente, no viene causada, en el sentido de Granger, por Y . En términos de funciones de densidad la

exogeneidad fuerte implica la factorización (II.28) con las propiedades (a) y (b) respecto a λ_1 y λ_2 .

Con fines de determinación de políticas económicas –a nivel macro y a nivel de agentes económicos individuales- es necesario realizar simulaciones, es decir, aproximar cuál será el comportamiento de la variable de interés, si algunas de las variables relacionadas con ella – Z – siguen trayectorias distintas a las que se han observado hasta el momento. En la simulación, por tanto, está implícito un cambio en el proceso marginal generador de las variables explicativas. Si este cambio en el modelo de Z_t provoca cambios en los parámetros de la función de densidad condicional de Y_t , entonces las variables Z no se pueden considerar como dadas (exógenas) a efectos de simulación y control.

Para fines de simulación y control es necesario que las variables Z sean superexógenas ante la clase de intervenciones o cambios que se van a hacer en su función de densidad marginal. Así, una variable Z es superexógena si.

- (1) es débilmente exógena, y
- (2) todos los parámetros λ_1 de la función de densidad condicional de Y_t son invariantes ante la clase de intervenciones o cambios considerados en los parámetros λ_2 .

Si una variable es superexógena a ella no le afecta, por definición, la crítica de Lucas (1976). La contrastación de la superexogeneidad se puede hacer respecto a cambios sufridos en el pasado, véase, por ejemplo, Engle y Hendry (1989).

En resumen, si las variables Z_t son débilmente exógenas podemos estimar e inferir sobre λ_1 , sin pérdida de generalidad, a partir del modelo que se deriva de la función de densidad condicional y si son superexógenas podemos simular con tal modelo. En ambos casos no es necesario modelizar las variables Z_t . Con la exogeneidad fuerte se tiene que en el modelo condicional de Y_t se pueden tomar como dadas las predicciones de Z necesarias para predecir Y . En este caso para predecir Z se necesitaría utilizar el modelo que se deriva de su densidad marginal. La ventaja de la exogeneidad fuerte en la predicción no radica en ignorar la modelización de Z , sino en poderlo hacer de forma aislada respecto la modelización condicional de Y . Es decir, la simplificación, que para fines predictivos comporta la exogeneidad fuerte, está en no

tener que estimar un modelo de n ecuaciones referido a X_t , sino dos modelos, uno de m ecuaciones, referido a X_t y otro de $(n-m)$ ecuaciones, referido a Z_t .

Determinado el proceso reductor y esclarecidos los supuestos bajo los que la condicionalización contemporánea es posible sin pérdida de generalidad, sólo nos falta comentar, tal como se hace en la sección siguiente, los modos de formular la dependencia dinámica en las variables económicas, para poder presentar una panorámica general sobre modelos econométricos.

II.3. Formulaciones dinámicas alternativas.

En cuanto a la dimensión temporal de las variables económicas, el proceso reductor ha puntualizado los siguientes aspectos: (a) la condicionalización respecto al pasado se hacía de generalidad, (b) al igual que en cualquier tipo de datos es necesaria una restricción de homogeneidad, que en este caso suele ser la de estacionariedad tras la reparametrización que las raíces unitarias permiten y (c) la necesidad de restringir la dependencia temporal en las transformaciones estacionarias.

En esta sección es conveniente realizar algunos comentarios sobre (c). Tal como se indicó en su momento, la restricción de dependencia en la función de densidad, $D(X_t/X_{t-1}^+; \lambda)$ que aparece en (II.9) se puede hacer truncando los retardos que se recogen en X_{t-1}^+ y obteniendo así la función $D(X_t/X_{t-1}^+; \lambda)$ o suponiendo un conjunto de hipótesis más generales: ergodicidad e invertibilidad. Con tales hipótesis es posible una dependencia de orden infinito, pero convergente entre las variables, que puede recogerse mediante polinomios temporales racionales.

En el caso en que la dimensión del vector X_t sea la unidad y bajo el supuesto de estacionariedad y linealidad, el truncamiento temporal implica que el modelo que se deriva de la función de densidad es:

$$X_t = E(X_t/X_{t-1}^{t-l}) + a_t, \quad (II.29)$$

$$\text{donde } E(X_t/X_{t-1}^{t-l}) = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_l X_{t-l}, \quad (II.30)$$

es decir, X_t viene generada por un proceso autorregresivo de orden l , $AR(l)$. Sin embargo, las condiciones de ergodicidad e invertibilidad permiten derivar modelos más

generales. Bajo tales hipótesis, utilizando la terminología anterior, se tiene que $D(X_t/X_{t-1}^{+e}; \lambda)$ y con la hipótesis de convergencia exponencial el modelo resultante puede ser:

$$X_t = E(X_t/X_{t-1}^{+e}; \lambda) + a_t \quad (\text{II.31})$$

$$E(X_t/X_{t-1}^{+e}; \lambda) = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (\text{II.32})$$

es decir, X_t viene generada por un proceso autorregresivo de orden p y de medias móviles de orden q , ARMA (p, q).

Ciertamente el cociente polinomial

$$(1 - \phi_1 L - \dots - \phi_p L^p) / (1 - \theta_1 L - \dots - \theta_q L^q),$$

en donde L es el generador de retardos, se puede aproximar por un polinomio autorregresivo del tipo $(1 - \alpha_1 L - \dots - \alpha_r L^r)$, pero en tal caso la aproximación será de un orden r no inferior a $(p+q)$ y normalmente mayor, por lo que el principio de parquedad en la parametrización requerirá utilizar (II.31) y (II.32) en vez de (II.29) y (II.30).

Es, por tanto, conveniente mantener las hipótesis de ergodicidad e invertibilidad como supuesto reductor y no solamente la de truncamiento temporal, como suele hacerse en la tradición econométrica de la London School of Economics (LSE).

Obsérvese que el modelo resultante de (II.31) y (II.32) podría verse como

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \eta_t, \quad (\text{II.33})$$

en donde η_t sigue un proceso de medidas móviles de orden q . Esta visión puede ser confusa porque el elemento residual, η_t , en (II.33) no es una innovación y por tanto $[\phi_1 X_{t-1} + \dots + \phi_p X_{t-p}]$ no es una esperanza matemática condicional respecto a un conjunto informativo relevante. Sin embargo, en (II.31) se destaca que tomando expectativas respecto a todo el conjunto informativo relevante, que es todo el pasado X que incluye también las innovaciones, se obtiene como residuo la innovación a_t .

En el caso que el vector X_t sea de dimensión mayor que uno, el problema de especificación dinámica con parquedad en la parametrización, destacando el conjunto informativo condicional relevante y la correspondiente innovación, es también inmediato. Lo ilustramos con el caso de n igual a dos, con una variable exógena.

En la función de densidad condicional $D(y_t/z_t, X_{t-1}^+; \lambda_1)$ el truncamiento temporal supone:

$$y_t = \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + \beta_0 z_t + \beta_1 z_{t-1} + \dots + \beta_s z_{t-s} + a_t. \quad (\text{II.34})$$

Si en (II.34) se cumple que

$$(1 - \alpha_1 L - \dots - \alpha_p L^p) = \alpha_r^*(L) \phi_b(L) \quad (\text{II.35})$$

$$(\beta_0 + \beta_1 L + \dots + \beta_s L^s) = \beta_d^*(L) \phi_b(L), \quad (\text{II.36})$$

es decir, si se da la restricción de factores dinámicos comunes de la literatura econométrica de la LSE, entonces se tiene un segundo modelo dado por:

$$y_t = \beta_0^* z_t + \dots + \beta_d^* z_{t-d} + \eta_t, \quad (\text{II.37})$$

donde para simplificar la exposición se ha supuesto que $\alpha_r^*(L) = 1$ y, en donde

$$(1 - \phi_1 L - \dots - \phi_p L^p) \eta_t = a_t. \quad (\text{II.38})$$

En ambos casos, la ecuación (II.34) o sistema formado por las ecuaciones (II.37) y (II.38), el modelo resultante proviene de

$$y_t = E(y_t/z_t, X_{t-1}^{t-1}; \lambda_1) + a_t, \quad (\text{II.39})$$

en donde el conjunto informativo es relevante y, por tanto, el residuo es una innovación. El sistema (II.37) y (II.38) se diferencia de (II.34) en que cumple la restricción de factores comunes.

El ejemplo anterior pone de manifiesto que para una explicación más intuitiva del modelo sobre y_t en el segundo caso, puede ser útil dividirlo en dos ecuaciones. Así, en la primera, (II.37), con $\beta_r^*(L) Z_t$ se recoge el efecto dinámico de la variable exógena y con η_t un efecto residual adicional, que se puede interpretar como el efecto de las variables omitidas. En la segunda ecuación se explica la generación de tal residuo. Sin embargo, esta explicación intuitiva no debe ocultar que el conjunto informativo condicional es el que se recoge en (II.39) y que el modelo genera auténticas innovaciones, dado el conjunto informativo considerado.

La conclusión es que una ecuación como (II.37) con un residuo que no es ruido blanco no es un modelo completo, éste requiere especificar el proceso generador de tal residuo. Si así se hace, ecuación (II.38), entonces se obtienen auténticas innovaciones y es posible determinar el conjunto informativo utilizado en su generación y, por tanto, el conjunto informativo en la expectativa condicional presente en el modelo de determinación de y_t .

La formulación de (II.37) y (II.38) supone restricciones respecto a (II.34). Sin embargo, puede haber situaciones en las que la teoría económica sugiera tales restricciones. Supóngase que la teoría sugiere determinadas características de suavidad y lentitud en el agotamiento de la función de respuesta de “y” respecto a z. Supóngase también que en la determinación de y_t entran otros factores como aspectos institucionales o progreso técnico, sobre los que no ha sido posible obtener mediciones, pero sobre los que se sabe seguro que no vienen generados por un proceso ruido blanco. Si en tales circunstancias $\beta_a^*(L)$ recoge bien las características de la función de y_t ante cambios Z_t , el modelo (II.37) y (II.38) en que se separan los efectos dinámicos específicos de las variables explicativas, de los efectos dinámicos residuales, puede ser especialmente útil. En este caso, habrá que contrastar que el modelo cumple las restricciones de factores comunes.

Obsérvese que en este ejemplo en donde los aspectos institucionales o progreso técnico tienen sobre y_t un efecto distribuido en el tiempo, si no se tiene ningún tipo de idea a priori sobre $\beta^*(L)$, habrá que abandonar el tipo de modelización recogido en (II.37) y (II.38). En tal caso, el elemento residual se podrá concebir como generado por

$$\eta_t = a_t/f(L),$$

donde $f(L)$ es un polinomio temporal. Así, multiplicando (II.37) por $f(L)$ se obtiene una formulación del tipo (II.34), que nos garantiza un elemento residual innovacional, que nos permite realizar un análisis empírico adecuado.

En este caso, cuando con las muestras, normalmente pequeñas, disponibles en economía se estimen los polinomios $\alpha_p(L)$ y $\beta_s(L)$, será difícil que la estimación recoja los factores comunes que $f(L)$ supone, sobre todo si el orden de $f(L)$ no es muy pequeño. En tal caso, la estimación habrá aproximado los verdaderos polinomios $\alpha_p(L)$ y $\beta_s(L)$ por otros muchos más sencillos, $\bar{\alpha}_m(L)$ y $\bar{\beta}_m(L)$, que probablemente no tendrán raíces comunes, pero que blanquearán la muestra, en el sentido de que los residuos no rechazan la hipótesis de venir generados por un proceso ruido blanco. En este caso $\bar{\alpha}_m(L)$ y $\bar{\beta}_m(L)$ permiten la modelización econométrica, pero no se les puede asignar significación económica.

Esta situación en la que es posible que se tenga información sobre la función de respuesta respecto a la variable explicativa, puede conducir a modelos como

$$y_t = g(L)z_t + \eta_t, \quad (\text{II.40})$$

donde

$$\eta_t = \frac{\theta(L)}{\phi(L)} a_t. \quad (\text{II.41})$$

es decir, modelos en los que la parquedad en la parametrización lleve a residuos ARMA como en (II.41) y no meramente a residuos AR como en (II.38).

En este caso, al igual que se dijo respecto al sistema (II.37) y (II.38), la formulación del modelo en dos ecuaciones puede ser de utilidad para entender el modelo, pero ello no debe ocultar que el modelo proviene de

$$y_t = E(y_t/X_{t-1}^{+e}; \lambda) + a_t,$$

en donde X_{t-1}^{+e} es el conjunto informativo relevante dada la información disponible.

Ciertamente se cumple que

$$\eta_t = \frac{\theta(L)}{\phi(L)} a_t \cong \frac{a_t}{f(L)},$$

con lo que multiplicando (II.40) por $f(L)$ se obtiene un modelo como (II.34). El problema está que el modelo del tipo (II.34) así obtenido puede incluir un número

excesivo de retardos, de forma que su estimación con los tamaños de muestra disponibles no sea aceptable. En tal caso las restricciones que (II.40) y (II.41) suponen sobre un modelo general del tipo (II.34) no se podrán contrastar.

La conclusión es que la estrategia de la LSE de abordar la restricción de dependencia mediante el truncamiento temporal es útil en muchos casos, pero no conlleva a excluir restricciones que supongan formulaciones dinámicas racionales. Asimismo, la parquedad en la parametrización puede llevar a sustituir $g(L)$, dígase de orden s , en (II.40) por $\omega_a(L)/\delta_c(L)$, en donde $(a + c) < s$. Un tratamiento más extenso del tema se encuentra en Espasa y Cancelo (1993), capítulo tercero.

Hasta este momento, en esta sección, se ha abordado el tema de truncamiento temporal en la media condicional $E(X_t/X_{t-1}^+)$, pero como se ha mencionado anteriormente, para describir la densidad condicional de algunos fenómenos económicos también es necesario suponer que los segundos momentos condicionales no son constantes y utilizar formulaciones operativas que sean compatibles con los datos. Así, si la varianza condicional, h_t , evoluciona con el tiempo, también es necesario truncar la dependencia temporal. Una de las especificaciones más sencillas propuestas en la literatura para modelizar h_t es el modelo GARCH (p, q) que, para el caso de que la dimensión del vector X_t sea la unidad, viene dado por

$$h_t = \omega + \sum_{i=1}^q \alpha_i x_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j};$$

ver Engle (1982) y Bollerslev (1986).

Se han propuesto en la literatura otras especificaciones para modelizar la evolución de la varianza. Considérese, por ejemplo, el siguiente proceso sin estructura en la media condicional.

$$x_t = \varepsilon_t \sigma_t$$

en donde ε_t es un ruido blanco y σ_t viene dado por

$$\log \sigma_t^2 = \omega + \alpha \log \sigma_{t-1}^2 + \eta_t$$

y η_t es también ruido blanco. Este modelo es conocido como modelo de volatilidad estocástica y fue propuesto por Taylor (1986).

En este caso, no se cumple la restricción de que la densidad condicional sea normal, pero se puede obtener condicionalidad lineal, mediante la transformación $\log X_t^2$. (Recuérdese la discusión en la sección II.1.4. sobre las restricciones distribucionales).

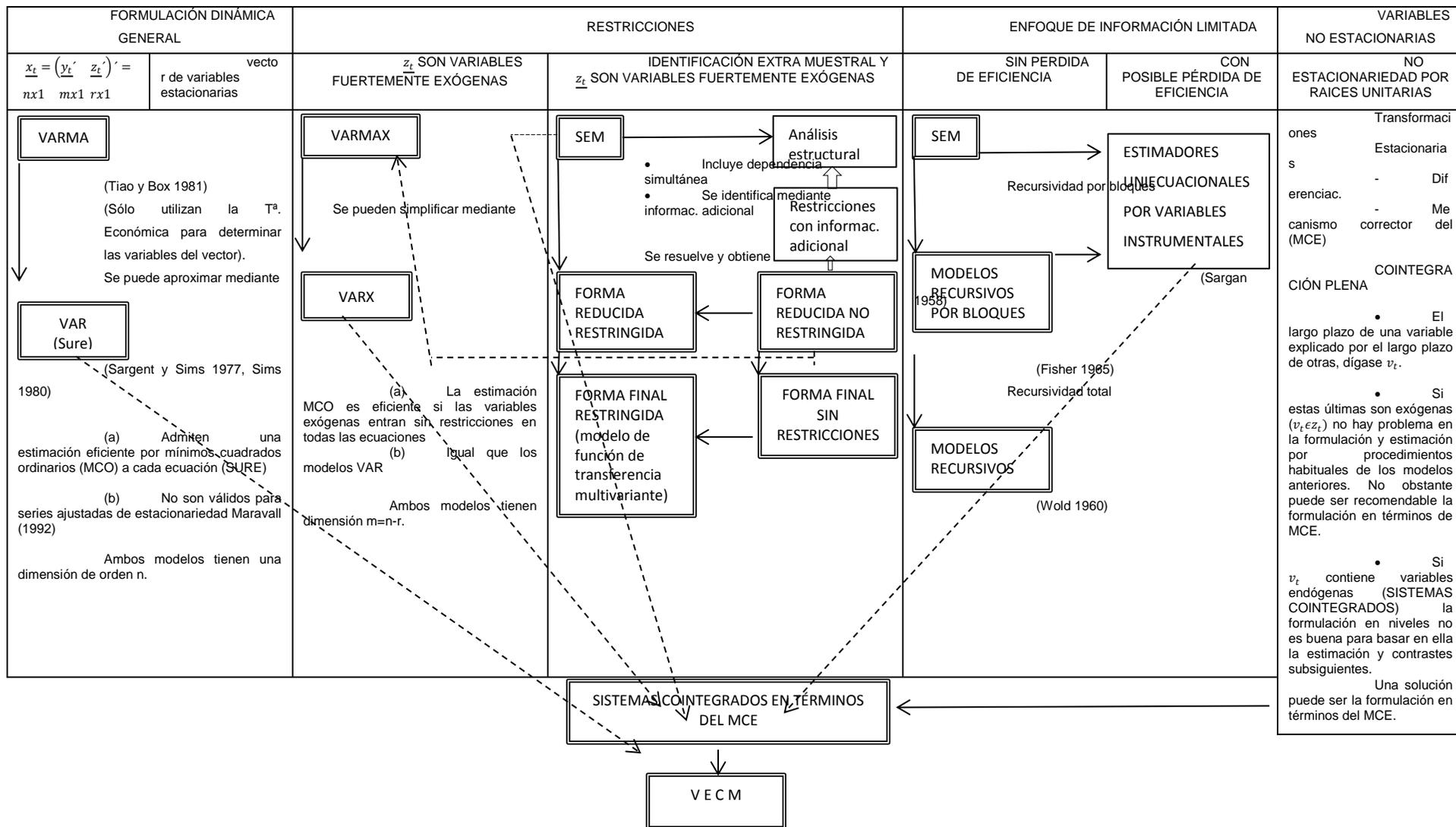
II.4. Esquema integrador de los modelos econométrico más usuales.

Con todo el preámbulo realizado en los tres epígrafes anteriores es posible presentar un esquema sintetizador de los modelos econométricos lineales más usuales, que se recoge en el cuadro 1.

II.4.1. Esquema integrador de los modelos econométrico más usuales.

Con las restricciones (1) a (5) del proceso reductor comentado en la sección II.1 se concluía con el prototipo de modelo recogido en (II.14). La información teórica que se utiliza en su formulación se reduce a especificar los elementos del vector X , por lo que implícitamente se está señalando los elementos del vector S_t sobre los que se está marginalizando. Aceptado el supuesto de marginalización respecto a S_t (II.14) es el modelo más general con el que se va a trabajar. Debido a la escasa implicación que tiene la teoría económica en la formulación de estos modelos, su desarrollo se ha producido fundamentalmente en el campo de la Estadística, en la rama de Series Temporales, y su introducción se encuentra en Quenouille (1957).

MODELOS ECONÓMICO LINEALES



Modelos VAR con mecanismo de corrección del error (King et al. 1991)

Bajo las hipótesis de linealidad y ergodicidad recogidas en (II.14) se tiene que

$$E(X_t/X_{t-1}^{+e}) = \Phi_1 X_{t-1} + \dots + \Phi_p X_{t-p} - \Theta_1 e_{t-1} - \dots - \Theta_q e_{t-q},$$

en donde

$$\Phi_j = \begin{pmatrix} \phi_{11}^j & \phi_{12}^j & \dots & \phi_{1n}^j \\ \vdots & & & \vdots \\ \phi_{n1}^j & \dots & \dots & \phi_{nn}^j \end{pmatrix} \quad \text{y}$$

$$\Theta_h = \begin{pmatrix} \theta_{11}^h & \dots & \theta_{1n}^h \\ \vdots & & \vdots \\ \theta_{n1}^h & \dots & \theta_{nn}^h \end{pmatrix},$$

con lo que denominado

$$\Phi_p(L) = I - \Phi_1 L - \dots - \Phi_p L^p, \text{ y} \quad (\text{II.42.a})$$

$$\Theta_q(L) = I - \Theta_1 L - \dots - \Theta_q L^q \quad (\text{II.42.b})$$

se tiene que

$$\Phi_p(L)X_t = \Theta_q(L)e_t, \quad (\text{II.43})$$

que es la formulación compacta del modelo ARMA multivariante de orden (p,q) o modelo VARMA (p,q). Una presentación de estos modelos orientada a su implementación práctica se encuentra Tiao y Box (1981) y un estudio a fondo de los mismos en Lütkephol (1991).

De (II.43) se obtiene, bajo la hipótesis de invertibilidad, que

$$[\Theta_q(L)]^{-1} \Phi_p(L)X_t = e_t,$$

con lo que utilizando la aproximación

$$[\Theta_q(L)]^{-1} \Phi_p(L) \cong \Phi_r(L),$$

se tiene

$$\Phi_r(L)X_t = e_t \quad (II.44)$$

que es un modelo autorregresivo vectorial y se denomina VAR (r). Este modelo se deduce directamente de (II.14) empleando la hipótesis de truncamiento temporal en la restricción de dependencia del proceso reductor. Esta obtención directa del VAR a partir de (II.14) no explicita la condición de invertibilidad, lo que ha llevado en muchas ocasiones a un uso indebido de estos modelos.

El uso de los modelos VAR en economía ha sido defendido por Sargent y Sims (1977) y Sims (1980) y durante la década de los ochenta han encontrado una respuesta relativamente amplia dentro de la profesión.

Esta popularidad de los modelos VAR frente a los modelos VARMA, escasamente utilizando en economía, se debe, en gran parte, a que la estimación de un VAR sin restricciones se puede realizar sin pérdida de eficiencia mediante la aplicación de mínimos cuadrados ordinarios, ecuación por ecuación.

Como se ha indicado antes, estos modelos incorporan la restricción de invertibilidad, por lo que, como demuestra Maravall (1992), no son válidos sobre series X_t que hayan sido previamente ajustadas de estacionalidad, y que tales series no son invertibles, es decir, no admiten una representación autorregresiva convergente.

En los modelos VAR y VARMA la dimensión del sistema es amplia, n , pues no se distingue entre variables endógenas y exógenas. En terminología econométrica, que se comenta posteriormente, estos son modelos de forma reducida, pero de la forma reducida más general posible. Ciertamente en datos de corte transversal estos modelos no tienen cabida, ya que no hay historia pasada sobre la que condicionar.

II.4.2. Modelos con la restricción de variables exógenas.

Si la teoría económica determina los parámetros de interés dentro del vector λ de la función de distribución conjunta que aparece en (II.13), es posible que en el vector X_t existía un subvector de variable exógenas. Excepto allí donde sea necesario, se utilizará simplemente el término “exógenas” sin precisar si son exógenas débiles, fuertes o superexógenas. La presencia de variables exógenas conlleva que sea posible la condicionalización contemporánea sin pérdida de generalidad, con lo que el prototipo de modelo que se obtiene es el recogido en (II.17), que se diferencia de (II.14) en que ahora en el conjunto informativo condicional se añade Z_t . Con eso la expectativa $E(Y_t/Z_t, X_{t-1}^+e)$ diferirá de (II.42) para el subvector Y_t en X_t , en que ahora, además de lo explicitado en (II.42), aparecen también términos $\beta_0 Z_t$, donde β_0 es una matriz $m \times k$. Con ello agrupando los distintos términos adecuadamente se obtiene

$$\Phi_p(L)Y_t = \beta_s(L)Z_t + \Theta_q(L)e_t. \quad (\text{II.45})$$

En la literatura estadística de series temporales a estos modelos se les denomina VARMAX, en donde la X que aparece en tal denominación hace referencia a que el modelo VARMA, respecto a las variables endógenas Y, está ampliado con variables explicativas exógenas¹¹.

En la derivación que se ha realizado del modelo VARMAX para las variables Y_t queda muy claro que es un modelo restringido respecto al modelo VARMA correspondiente a X_t , ya que el primero incorpora la restricción de exogeneidad. Sin embargo, la terminología VARMAX, o la de “modelos multivariantes de función de transferencia con múltiples inputs o múltiples outputs” empleada en Jenkins (1979), puede inducir confusión pues presentan el modelo VARMAX sobre Y_t como un modelo más general que el modelo VARMA sobre Y_t (“modelos estocásticos multivariantes con múltiples outputs” según Jenkins (1979)). Esto último es obvio. La confusión puede venir por el hecho de que el modelo VARMA sobre Y_t se obtiene marginalizando el proceso generador de los datos respecto a Z_t , aspecto claramente inadecuado, que se pone de manifiesto al considerar los modelos VARMAX en los que aparecen las variables Z_t .

¹¹ En la literatura de series temporales se suele utilizar la letra X para representar el vector de variables que se toma como dado en el análisis. Nosotros por el contrario venimos utilizando la letra Z.

El comentario anterior sirve para ilustrar la importancia del paso de la marginalización en el proceso reductor que se realiza en todo análisis aplicado. Esta es una decisión que afecta a todas las etapas restantes del proceso reductor y un cambio sobre el conjunto de variables sobre las que se ha marginalizado, implica revisar todo el proceso reductor. No hacerlo puede llevar a finalizar el análisis con la elección del modelo VARMAX sobre el modelo VARMA respecto a las variables Y_t , en una comparación entre ambos, cuando la hipótesis de exogeneidad del VARMAX sea inaceptable. Considerar el modelo VARMAX como el más general puede suponer una negligencia en el analista, introduciendo una restricción de exogeneidad sin percatarse de ello.

Si en (II.17) el término $E(Y_t/Z_t, X_{t-1}^{+e})$ se reemplaza por $E(Y_t/Z_t, X_{t-1}^{-e})$ haciendo uso de la hipótesis de truncamiento temporal en la restricción de dependencia se concluye con el modelo

$$\Phi_p(L)Y_t = \beta_s(L)Z_t + e_t, \quad (II.46)$$

que se denomina modelo VARX.

Al igual que lo dicho para los modelos VAR y VARMA, se tiene que en la práctica económica los modelos VARX son muchos más populares que los VARMAX.

En (II.46) se puede despejar Y_t con lo que se obtiene:

$$Y_t = \Phi_1 Y_{t-1} + \dots + \Phi_p Y_{t-p} + \beta_s(L)Z_t + e_t, \quad (II.47)$$

que, en terminología econométrica, es un modelo de forma reducida propiamente dicho, ya que incluye variables exógenas. En él, el residuo es ruido blanco, pero si existen factores comunes entre $\Phi_p(L)$ y $\beta_s(L)$ se podría formular (II.47) con un residuo autorregresivo, de forma similar a como se discutió anteriormente para el caso uniecuacional.

De (II.45) se obtiene

$$Y_t = \Phi_1 Y_{t-1} + \dots + \Phi_p Y_{t-p} + \beta_s(L)Z_t + \Theta_q(L)e_t, \quad (II.48)$$

en donde el residuo sigue un proceso vectorial de medias móviles. Ciertamente dichas medias móviles se pueden aproximar mediante un esquema autorregresivo y eliminarlo del componente residual ampliando adecuadamente $\Phi_p(L)$ y $\beta_s(L)$, de forma similar a la discutida en la sección II.3. La utilización del esquema resultante –que será del tipo (II.47)- o de (II.48) se reduce a un problema de parquedad en la parametrización.

El modelo (II.47) es un sistema de m regresiones dinámicas, al que le es aplicable la denominación SURE (seemingly underlated regression equations) introducida en Zellner (1962). En la terminología de la LSE, véase Hendry et al. (1984) a los modelos recogidos en (II.47) se les denomina autorregresivos y de retardos distribuidos, AD (autoregressive-distributed lag); autorregresivos porque entran retardos en Y_t y de retardos distribuidos porque las variables exógenas entran contemporánea y retardadamente.

En el caso uniecuacional, $m=1$, se tendría cambiado la terminología de los parámetros por otra más acorde con la empleada dentro de la tradición de la LSE,

$$y_t = \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + \sum_{j=1}^k (\beta_{j0} z_{jt} + \beta_{j1} z_{jt-1} + \dots + \beta_{js} z_{jt-s}) + e_t, \quad (II.49)$$

que es la hipótesis mantenida en la que se basa la metodología de la LSE.

Con datos de corte transversal el modelo (II.45) pierde su dimensión dinámica y queda reducido a

$$Y_t = \beta z_t + e_t, \quad (II.50)$$

en donde el subíndice t se refiere a individuos. El sistema (II.5) es un sistema de forma reducida estático formando m modelos de regresión clásicos.

II.4.3. Modelos identificados con restricciones teóricas: MODELOS ESTRUCTURALES.

Los modelos del epígrafe anterior han sido derivados a partir del siguiente esquema:

$$Y_t = E(Y_t/Z_t, X_{t-1}^{+e}, X_0) + e_t, \quad (\text{II.50})$$

que podemos representar más escuetamente como

$$Y_t = \mu_t + e_t, \quad (\text{II.51})$$

en donde μ_t , que es una esperanza matemática condicional, es ortogonal a las innovaciones e_t . Si denominamos a μ_t a la parte sistemática del modelo, tenemos que en (II.51) la parte sistemática es ortogonal a las innovaciones. De forma similar, se han derivado los modelos del epígrafe II.4.1- Por su modo de derivación a ambos modelos podemos denominarlos modelos estadísticos, pues proceden directamente de una distribución condicional.

El tema de identificación, tal como se planteó inicialmente en Frisch y en Haavelmo (1944), es más amplio de lo que se entiende por tal en los libros de Econometría. Para entrar en él, conviene distinguir, tal como hace Spanos (1989), a quien seguimos en este punto, entre identificación estadística e identificación estructural. Si se dispone de un conjunto de observaciones y de una clase de modelos estadísticos, la identificación estadística se refiere a la estimación (elección) de un modelo particular, dentro de esa clase, el cual constituye un resumen adecuado de la información muestral. En este sentido la identificación estadística está directamente relacionada con el concepto de identificación en la metodología Box-Jenkins.

Los parámetros de μ_t y de la función de densidad de e_t en (II.51) son función de los parámetros de la función de densidad condicional $D(Y_t/Z_t, X_{t-1}^{+e}, X_0; \lambda_1)$. Esta última a través de las restricciones de homogeneidad, distribucionales y de memoria ha sido formulada de forma que tenga en cuenta las peculiaridades de los datos, por lo que los modelos del tipo (II.50) están identificados estadísticamente, con lo que su estimación a partir de la muestra es factible.

Por otra parte, los modelos teóricos, a los que denominaremos modelos estructurales, pueden tener una formulación distinta a la representada en (II.50) y en el epígrafe I.6 se vio, siguiendo la terminología de Haavelmo, que venían representados por unas funciones f_{ti} como la recogida en (I.4). Utilizando la terminología desarrollada en las últimas secciones, tenemos que tales modelos teóricos f_{ti} se pueden escribir como

$$Y_t = F(\widetilde{X}_t, X_{t-1}^+; \gamma) + a_t, \quad (\text{II.52})$$

en donde \widetilde{X}_t indica que en la ecuación para y_i ($j = 1, \dots, m$) el vector \widetilde{X}_t se obtiene excluyendo de X_t la variables y_{jt} .

Si nos fijamos en una ecuación de este sistema, que sin pérdida de generalidad puede ser la primera, partiendo Y_t como $(y_{jt} \widetilde{Y}_t)'$ tenemos que

$$y_{1t} = f_1(\widetilde{Y}_t, Z_t, X_{t-1}^+; \gamma_1) + a_{1t}, \quad (\text{II.53})$$

en donde $f_1(\cdot)$ puede no ser la $E(y_{1t}/\widetilde{Y}_t, Z_t, X_{t-1}^+)$, en cuyo caso, la parte sistemática de (II.53) $-f_1(\cdot)-$ no es ortogonal con a_{1t} . Es decir, los modelos estructurales se formulan según los dictámenes de la teoría sin una preocupación por la ortogonalidad, característica que se da en los modelos estadísticos, por lo que, la parte sistemática de un modelo estructural puede no ser ortogonal con la innovación. Si no existe tal ortogonalidad, la estimación de los parámetros del modelo ha de realizarse con variables instrumentales.

Si F es lineal existe una reformulación del modelo

$$y_t = E(Y_{1t}/\mathbb{D}_t) + \varepsilon_t, \quad (\text{II.54})$$

en la que \mathbb{D}_t es ortogonal a ε_t y en donde los parámetros de $E(y_{1t}/\mathbb{D}_t)$ son los de $f_1(\widetilde{Y}_t, Z_t, X_{t-1}^+)$. En el ejemplo de la función de consumo keynesiana de la sección I.6, tenemos que la ecuación estructural (II.53) viene dada por (I.6.a), para cuya correspondiente función $f_1(\cdot)$ se demuestra en (I.9) que no es una esperanza matemática condicional. La esperanza $E(y_{1t}/\mathbb{D}_t)$ introducida en (II.54) viene en el ejemplo recogida en (I.11).

Los parámetros estructurales se podrán estimar si se dispone de suficientes instrumentos. Sin que aquí se entre en ello, el problema de identificación estructural es un problema de disponer de suficientes instrumentos para la estimación de parámetros.

Otra forma de ver el problema de identificación estructural es el siguiente. Teniendo en cuenta las peculiaridades de los datos se ha formulado el modelo estadístico (II.5), que es el que puede estimarse con la información disponible. Si el interés no está en el modelo estadístico, sino en el estructural será necesario una reparametrización de aquél. En consecuencia la identificación estructural se refiere a que los parámetros estructurales γ se pueden definir de forma única en términos de los parámetros del modelo estadístico, λ_1 . Por tanto, la identificación estructural es un problema previo a la estimación.

Bajo la restricción de linealidad el modelo estadístico general al que hemos llegado es el recogido en (II.45). En él, la matriz Φ_0 , tal como se define en (II.42.a), es la unidad. Ahora bien, la teoría puede indicar una relación de dependencia simultánea en las variables Y_t , lo que implicaría que elementos fuera de la diagonal principal de Φ_0 tendrían que tomar valores que no fuesen nulos. Piénsese, por ejemplo, en un modelo anual de determinación de precios y salarios. Ahora bien, si Φ_0 es distinta de la unidad en el modelo estructural, éste tendrá que incluir restricciones en otros parámetros de las restantes matrices para que sea posible obtener los parámetros estructurales a partir de los del modelo estadístico. Si el número de restricciones son exactamente las necesarias para poder obtener los parámetros estructurales si dice que el modelo está exactamente identificado. Pero el número de restricciones puede ser superior, y el modelo estar sobreidentificado. En este caso el modelo estructural supone pues restricciones adicionales sobre los modelos estadísticos del epígrafe anterior.

Con las restricciones de identificación se están señalando los instrumentos disponibles para la estimación, con lo cual al Teoría Económica al indicar dichas restricciones está apuntando los instrumentos disponibles para la estimación estructural. La validez de tales dictámenes de la Teoría se podrá inferir contrastando

las restricciones que el modelo estructural impone al estadístico. Sobre este punto es importante señalar que tal contraste sólo tendrá sentido si se realiza sobre un modelo estadístico adecuado para los datos.

A los modelos estructurales, por la simultaneidad que permiten en la matriz Φ_0 , se les denomina modelos de ecuaciones simultáneas, SEM (simultaneous equation models). En la literatura econométrica tales modelos se suelen formular con una terminología para los parámetros diferente a la empleada en los dos epígrafes anteriores. Estos modelos se pueden formular así:

$$B_0 Y_t + B_1 Y_{t-1} + \dots + B_p Y_{t-p} + C(L)Z_t = u_t, \quad (II.55)$$

en donde

$$C(L) = C_0 + C_1 L + \dots + C_S L^S \quad y$$

u_t puede ser ruido blanco o venir generado por un proceso estacionario, en cuyo caso se denominará a_t a la correspondiente innovación y tendrá una matriz Σ de varianzas y covarianzas. La identificación del modelo supone que existen suficientes restricciones en las matrices de parámetros estructurales y en Σ .

Premultiplicando (II.55) por B_0^{-1} y despejando Y_t tenemos:

$$Y_t = P_1 Y_{t-1} + \dots + P_p Y_{t-p} + \pi^*(L)Z_t + v_t; \quad (II.56)$$

si u_t era simplemente estacionario, también lo será v_t , en cuyo caso, denominaremos e_t a sus correspondientes innovaciones.

Al sistema expresado en (II.56) se le denomina modelo de forma reducida (RF). Si se compara (II.56) con los que la sección anterior vemos que no son más que modelos VARX o VARMAX. Por tanto, los modelos de forma reducida son lo que hemos denominado modelos estadísticos. Al mismo tiempo, los modelos, que en terminología tomada de la literatura de series temporales, hemos denominado VARX o VARMAX no son más que modelos de forma reducida en la terminología econométrica. Obsérvese, sin embargo, que el modelo SEM en (II.55) incorpora

restricciones en los parámetros para poder estar identificado, con lo que al resolver el sistema se obtendrá una forma reducida restringida a diferencia de las formas reducidas del epígrafe anterior que no incluían restricciones.

La discusión llevada a cabo en este epígrafe sobre modelos estructurales es válida para modelos sobre datos de corte transversal, si bien, en tal caso las formulaciones SEM y de forma reducida serán mucho más sencillas ya que no tendrán una dimensión dinámica. En concreto, ambos tipos de modelos tomarán, respectivamente, las siguientes formas:

$$B_0 Y_t + C_0 Z_t = a_t,$$

$$Y_t = \pi_0 Z_t + e_t$$

donde u_t y e_t son innovaciones y el subíndice t en todas las variables se refiere a individuos y no al tiempo.

II.4.4. Modelos recursivos

En los epígrafes anteriores se ha señalado que en un modelo simultáneo general, además de posibles relaciones desfasadas temporalmente entre sus variables, existe relación contemporánea entre todas o parte de las variables endógenas, recogidas en la matriz B_0 . Por ello se dice que el modelo es simultáneo, o directamente simultáneo. En este sentido, los modelos de forma reducida o los modelos VARMA/VARMAX son sólo indirectamente simultáneos, ya que la relación contemporánea entre variables endógenas no aparece en la parte funcional, sino en la matriz de varianzas y covarianzas de las innovaciones.

A estos modelos indirectamente simultáneos y con un elemento residual ruido blanco se les denomina, tal como se indicó antes, modelos SURE. En Zellner (1962) se demuestra que en estos modelos, salvo que cumplan al menos una de las dos siguientes condiciones:

a) que todas las ecuaciones tengan exactamente las mismas variables explicativas,

b) que la matriz de varianzas y covarianzas de los residuos –que vienen generados por un proceso ruido blanco- sea diagonal,

la estimación individual de las ecuaciones lleva a pérdidas de eficiencia frente a una estimación conjunta. Obsérvese que un modelo VAR corresponde al primero de los casos mencionados, por lo que estos modelos pueden estimarse eficientemente utilizando para cada ecuación estimadores mínimo cuadráticos ordinarios.

Un tipo especial de modelo estructural es aquél en que se puede dar dependencia contemporánea, pero ésta es de forma tal que no hay retroalimentación. Esto implica que las ecuaciones del modelo se pueden ordenar de forma tal que, parafraseando a Fisher (1965) página 609, una variable endógena de orden superior nunca influye en una variable endógena de orden inferior, ni contemporáneamente ni desfasada en el tiempo. A estos modelos se les denomina modelos recursivos.

En la literatura econométrica los modelos simultáneos fueron tratados inicialmente por Haavelmo (1943) y con posterioridad en el seno de la comisión Cowles se abordaron los problemas de identificación y estimación de dichos modelos; sus resultados se publicaron en las monografías diez (Koopmans, 1950) y catorce (Hood y Koopmans, eds.; 1953) de dicha comisión.

Poco después de aparecer la última monografía citada surgen, como un enfoque alternativo a los modelos estructurales simultáneos, los modelos recursivos, propuestos por Wold (1954, 1960 y 1964). La tesis de Wold consiste en que en un modelo la causalidad debe estar encadenada, ya que en el mundo real una variable no puede ser al mismo tiempo causa y efecto de otra.

Sin embargo, y a pesar de los argumentos de tipo lógico en favor de los modelos recursivos, en Economía se da el problema de que en la inmensa mayoría de los casos las variables que entran en un sistema no se pueden observar a intervalos de tiempo tan pequeños como para la recursividad sea un concepto operativo. Así, los modelos macroeconómicos basados en datos de las contabilidades nacionales, que suelen construirse con periodicidad anual o como mucho trimestral, no pueden

formularse en términos recursivos. Es cierto que cuanto más se desagrega en el tiempo un modelo, menor es el problema de simultaneidad y más admisible resulta la hipótesis de recursividad; pero en contrapartida, la complejidad dinámica se incrementa considerablemente.

Las restricciones que el modelo recursivo impone sobre el modelo SEM de la sección II.4.3 y sus implicaciones sobre la estimación se pueden ver, por ejemplo, en Espasa (1973)¹².

Un modelo estructural dinámico recursivo se caracteriza por el hecho de que siempre es posible ordenarlo de tal forma que la variable endógena de la primera ecuación, Y_1 , dependa sólo de su pasado, de las variables exógenas y de los valores presentes y pasados de su propia innovación, y además su innovación no está relacionada con ninguna otra innovación, presente o pasada, que aparezca en el modelo. Esto implica que la variable Y_1 puede estudiarse de manera independiente del resto de variables endógenas, es decir, puede ser modelizada de forma consistente y eficiente especificando y estimando únicamente la ecuación que a ella se refiere, e ignorando las demás ecuaciones.

La segunda ecuación del modelo, referida a la variable Y_2 , ha de reflejar que ésta depende de su propio pasado, del presente y pasado de Y_1 , de las variables exógenas, de los valores presentes y pasados de su propia innovación y de las innovaciones pasadas de Y_1 . A esto hay que añadir que su innovación a_2 ha de ser independiente de cualquier otra innovación del modelo. Si esto se cumple, entonces para construir un modelo para Y_2 basta con considerar únicamente su ecuación, a Y_1 se le trata como una variable exógena de las contenidas en Z , ya que de hecho es exógena.

En general, en la ecuación j -ésima correspondiente a la variable Y_j han de entrar únicamente:

¹² En la transcripción de este artículo se produjeron bastantes erratas, por lo que el lector interesado debe consultar la fe de erratas a dicho artículo, publicada en Cuadernos de Economía, 1973, núm. 2.

- los valores pasado de Y_j ;
- los valores presentes y pasado de Y_1, Y_2, \dots, Y_{j-1} ;
- las variables exógenas Z ;
- los valores presente y pasados de a_j ;
- los valores pasado de a_1, a_2, \dots, a_{j-1} ;

y además

- la innovación a_j han de estar incorrelacionada con cualquier otra innovación del modelo.

Si estos requisitos se cumplen, el modelo para Y_{jt} se obtiene considerando exclusivamente su ecuación de comportamiento, ya que en ella Y_1, Y_2, \dots, Y_{j-1} son variables exógenas.

La principal consecuencia de la recursividad es que en un modelo recursivo la estimación conjunta de todo el sistema proporciona exactamente el mismo resultado que si se estima cada ecuación por separado. Esta propiedad es consecuencia del hecho de que un sistema recursivo la causalidad siempre es unidireccional; por lo tanto si el interés se centra en una variable Y_j , no es necesario estimar las ecuaciones correspondientes a Y_1, Y_2, \dots, Y_{j-1} , ya que éstas son variables fuertemente exógenas en la determinación de Y_j , y tampoco es necesario incluir las ecuaciones de Y_{j+1}, \dots, Y_n , pues estas variables no intervienen en la determinación de Y_j .

A priori resulta difícil saber si las matrices de coeficientes y la matriz de varianzas de las innovaciones cumplen las restricciones necesarias para que el modelo sea recursivo. En la práctica la recursividad suele ser un resultado que se obtiene a posteriori a raíz de estimar un modelo multiecuacional, se observa que los resultados apuntan hacia un modelo recursivo, en vez del modelo con causalidad multidireccional que se había planteado en un principio.

Aunque hasta ahora se ha comentado la recursividad en el marco de modelos SEM, esta propiedad no se da exclusivamente en los modelos estructurales, ya que también los modelos multivariantes de series temporales pueden presentarla.

Como ya se ha señalado los sistemas económicos tal y como se observan no pueden, en general, representarse como modelos recursivos. Sin embargo, sí resulta más verosímil que se dé un supuesto alternativo menos restrictivo, propuesto en Fisher (1965), y que se ha denominado recursividad por bloques.

En otras palabras, un modelo recursivo por bloques es aquél en que cada bloque de ecuaciones individualmente considerado constituye un modelo simultáneo, pero por grupos de ecuaciones la estructura es recursiva; esto implica que para fines de estimación, un bloque cualquiera puede tratarse aisladamente sin pérdida de eficiencia. Cuando se da este tipo de recursividad la dimensión de la simultaneidad se reduce enormemente, lo que simplifica bastante la labor del analista.

La recursividad implica que partiendo del vector $X_t = (Y_t' Z_t')$, en donde las variables Z_t se consideran exógenas, en la modelización de las variables de interés Y_t se encuentra que cuando tales variables aparecen como explicativas en las m ecuaciones del sistema se pueden considerar como fuertemente exógenas. Debido a tales características de exogeneidad en las variables Y_t , el sistema en el que éstas se determinan puede considerarse, sin pérdida alguna de eficiencia, ecuación por ecuación.

III. EL PLANTEAMIENTO UNIECUACIONAL DE LA MODELIZACIÓN ECONOMÉTRICA

III.1. Procedimientos de información limitada y variables instrumentales.

El análisis de los modelos recursivos ha traído a un primer plano la modelización de los fenómenos económicos basada en considerar solamente la ecuación referida a la variable de interés. En la modelización econométrica se pueden considerar dos situaciones extremas.

1) Una en la que se da causalidad multidireccional entre las variables económicas, por lo que se requiere, salvo las dos excepcionales indicadas anteriormente al hablar del modelo SURE, la modelización conjunta de todo el sistema. Ciertamente esto es siempre necesario si el modelo tiene un componente de medias móviles sobre todo el vector de innovaciones. Si se persiguen fines predictivos será necesario estimar todo el sistema, conjuntamente o no, según las características de éste.

Cuando lo que se pretende es solamente estimar los parámetros del modelo u éste no tiene parte de medias móviles, conviene distinguir los modelos de forma reducida, por ejemplo VARX, en los que la multidireccional contemporánea se recoge exclusivamente en la matriz de varianzas y covarianzas de las innovaciones, y los modelos estructurales, en los que tal interrelación aparece también en la parte funcional del modelo.

En el primer paso si se estima una ecuación ignorando el resto se obtienen unos resultados que, excepto en los modelos en los que todas las ecuaciones tienen exactamente las mismas variables explicativas –como ocurre en los VAR- son consistentes pero ineficientes; en el segundo, los resultados son también, en general, inconsistentes.

2) En la sección II.4.4 se estudió una clase especial de modelos. Los modelos recursivos, en los que, debido a sus características particulares, la estimación uniecuacional proporciona los que, debido a sus características particulares, la estimación uniecuacional proporciona exactamente los mismos resultados a los que conduciría el estudio del sistema en conjunto.

Es de destacar que el analista no tiene ningún tipo de control sobre la recursividad del sistema; se encuentra con que las variables que analiza, tal y como se observan, tienen la característica de recursividad, y por lo tanto basta para fines de estimación, con que se centre su esfuerzo en la ecuación que le interesa, o se encuentra con que el sistema de ecuaciones para esas variables no es recursivo y se ve obligado a tratar dicho sistema en conjunto¹³.

Desgraciadamente, en la práctica es corriente encontrarse con alguno(s) de los siguientes problemas:

- 1) Existe incertidumbre sobre la dimensión multivariante del estudio.
- 2) No se dispone de información teórica suficiente y fiable para todas las ecuaciones del modelo.
- 3) Algunas ecuaciones no se pueden formular, especialmente por dificultades a la hora de disponer de datos sobre todas las variables explicativas.
- 4) El coste del análisis multiecuacional no se justifica en función de sus beneficios esperados.

De ahí que al analista aplicado le sea de especial interés poder llegar a una situación intermedia entre las dos que se han considerado hasta ahora. Esto le lleva a preguntarse si existe algún tipo de planteamiento parcial, consistente en fijarse en un número reducido de ecuaciones ignorando el resto, que proporcione resultados aceptables; y en caso afirmativo, cómo se puede medir el coste asociado a dicho planteamiento en relación al enfoque multiecuacional general. Este planteamiento, que en la literatura se denomina de información limitada, tiene su versión extrema cuando el investigador sólo se fija en una ecuación ignorando las demás, dando origen al planteamiento uniecuacional.

El objetivo último de este enfoque es que el analista pueda escoger entre realizar un estudio complejo que optimice la información disponible, o un análisis más sencillo, donde la sencillez tendrá un coste determinado que será aceptable o no en función del tipo de resultados requeridos.

¹³ Y como se discutió en el epígrafe anterior la característica de recursividad normalmente se pone de manifiesto a posteriori, una vez que se ha estimado el sistema.

Los estimadores que se manejan en el análisis estadístico-econométrico de series temporales se valoran, entre otras, en función de dos características, la consistencia y la eficiencia. El tratamiento multiecuacional del sistema garantiza la obtención de estimadores consistentes y eficientes; en cambio, si el sistema no es recursivo, el planteamiento uniecuacional dará lugar, salvo las excepciones señaladas, a estimadores ineficientes e incluso, según el tipo de modelos, inconsistentes.

El problema de inconsistencia que aparece con los modelos estructurales no se puede, en el caso de modelos estructurales sobreidentificados, resolver a partir de estimadores consistentes de la forma reducida no restringida, pues, para tales modelos, de dicha forma reducida no se pueden obtener los parámetros estructurales directamente. Las restricciones se pueden poner por el procedimiento de distancia mínima propuesto por Malinvaud (1966) sección 19.3, aunque dicho procedimiento, al igual que el de la máxima verosimilitud, requiere la optimización de una función no lineal, incluso para modelos lineales. Sin embargo, podemos emplear una técnica especial, denominada estimación por variables instrumentales que da estimadores lineales para procesos lineales. El método de variables instrumentales, que fue introducido por Reiersol (1941, 1945), aplicado y desarrollado por Geary (1949) y Durbin (1954), y formalizado en su enfoque actual por Sargan (1958).

El método de variables instrumentales, en sus múltiples variantes, ha recibido mucha atención en la literatura, ya que es de utilidad siempre que la parte sistemática de un modelo no sea ortogonal con el término de error. Además, su aplicación tampoco está libre de dificultades. No obstante, para los propósitos de esta memoria su característica principal es la de permitir un tratamiento uniecuacional de variables que pertenecen a un sistema. En Hendry (1976) se deriva un sistema de ecuaciones normales del que se obtienen prácticamente todos los tipos de estimadores conocido para el modelo SEM con residuos ruido blanco o autorregresivos de primer orden. En Espasa (1975) esto se generaliza, mediante métodos espectrales, para sistemas con cualquier tipo de residuos con la única restricción de que sean estacionarios.

La elección de los instrumentos con los que obtener las variables instrumentales es una labor ardua ya que se suele disponer de poca información al

respecto. En este sentido el analista encontrará en la Teoría Económica, como ya se indicó anteriormente, una orientación importante¹⁴.

Tal y como se acaba de comentar, esta elección está plenamente vinculada a las restricciones con las que se ha especificado el modelo SEM. Así, ciñéndonos al caso de restricciones cero en los parámetros de las matrices estructurales en un modelo estático, el número de instrumentos realmente disponibles para estimar los parámetros correspondientes a las variables endógenas que entran como explicativas en una ecuación dada, equivale exactamente al número de restricciones cero que afectan a las variables exógenas en esa ecuación. Obsérvese, por tanto, que el tener suficientes restricciones para que la ecuación esté identificada equivale a tener suficientes instrumentos para aplicar el procedimiento de variables instrumentales; además, la sobreidentificación implica más instrumentos que los estrictamente necesarios, lo que normalmente conducirá a obtener una variable instrumental más correlacionada con la variable endógena que instrumentaliza y, por tanto, a proporcionar un estimador más eficiente.

En cuanto a la ineficiencia relativa a un conjunto de información mayor de los estimadores de información limitada, no hay ninguna manera de compensar la pérdida de información que, normalmente, resulta del tratamiento uniecuacional en modelos que no son recursivos. El no tener en cuenta las interrelaciones entre todas las variables endógenas del modelo forzosamente conlleva una visión limitada y parcial del sistema económico, lo que se traduce, en general, en una pérdida de precisión de los estimadores resultantes.

En ese sentido el coste de un planteamiento uniecuacional se puede medir como la pérdida de precisión, en términos de varianza de la innovación, que se produce al pasar de un tratamiento multiecuacional al uniecuacional.

Al analista le corresponde así la decisión de intercambiar eficiencia por simplificación del análisis; si prefiere la eficiencia, necesariamente deberá plantearse

¹⁴ Un excelente trabajo sobre la problemática y posibles soluciones de la selección de instrumentos en Fisher (1965).

un modelo multivariante; si por el contrario espera que la ganancia en eficiencia no compense la mayor complejidad del modelo multivariante, puede decirse por la solución uniecuacional.

Sin embargo, conviene señalar que la mayor eficacia del modelo completo supone que éste se puede especificar adecuadamente, ya que un error con una ecuación del mismo distinta a la ecuación de interés, puede conducir a estimadores inconsistentes. Muchas veces la formulación del modelo completo es muy compleja pero la del uniecuacional no, con lo que la mencionada pérdida de eficiencia es en esos casos más teórica que real. En tales circunstancias en las que el modelo global no se puede especificar la máxima eficiencia factible es la que se obtiene con el enfoque de información limitada. Este enfoque en cualquier caso debe verse como un procedimiento sólido frente al riesgo de errores de especificación en el resto del modelo global.

III.2. Precisiones sobre el uso de modelos econométricos

Como conclusión a lo tratado hasta este momento se puede decir que en la formulación de un modelo econométrico con n variables es importante explicar las variables exógenas, pues ello permite reducir la dimensión multiecuacional. Si de las n variables que se consideraron al principio k son exógenas, entonces basta con plantear el subsistema formado por las $n-k$ ecuaciones correspondientes a las variables endógenas.

Este procedimiento de condicionar respecto a las variables exógenas se realiza principalmente a partir de los dictámenes de la Teoría económica, que establece cuales son los parámetros de interés. Por tanto, este condicionamiento no es algo plenamente basado en los datos, por lo que autores como Sims, véase Sims (1980). Lo ponen en entredicho y sugieren operar a través de los modelos VAR sin establecer a priori si hay variables exógenas.

Cuando el análisis se realiza con fines predictivos, la reducción de la dimensionalidad, pasando de n a $n-k$ ecuaciones, sólo se logra si las variables exógenas son fuertemente exógenas. Bien entendido que aun en este caso habrá que estimar el sistema de k ecuaciones respecto a las variables exógenas, para generar predicciones sobre las mismas que entrarán como inputs (valores dados) en la predicción realizada con el sistema de $n-k$ ecuaciones correspondientes a las variables endógenas.

La ventaja habrá consistido en operar con dos sistemas de menor dimensión, en vez de con un gran sistema de dimensión n . En la práctica, la predicción de las variables exógenas se realiza mediante modelos univariantes, con lo que su sistema de k ecuaciones conjuntas se sustituye por k modelos univariantes.

La predicción obliga a operar con todo el sistema de $n-k$ ecuaciones, ya que aunque sólo se esté interesado en la variable Y_h , si ésta depende de otras variables endógenas es necesario utilizar todo el sistema, la forma reducida restringida, en el que se determinan las variables endógenas ($n-k$ ecuaciones) para obtener una predicción óptima de Y_h .

Sin embargo, si el objetivo es estimar los parámetros de un número reducido de ecuaciones, que con frecuencia puede ser una única ecuación, cabe adoptar un enfoque de información limitada que incluso puede llegar a ser uniecuacional. Este enfoque comportará, en general, una pérdida de eficiencia que es absolutamente inevitable; y además planteará problemas de consistencia, que tendrán que ser resueltos mediante la aplicación de procedimientos de variables instrumentales. No obstante, las pérdidas de eficiencia pueden muy bien ser aceptables dado que gracias al enfoque de información limitada habrá sido posible estimar una relación económica de interés inmersa en un sistema multiecuacional que puede ser muy complejo de formular.

En resumen, las características de exogeneidad que un sistema multiecuacional concreto posee constituyen una propiedad específica del mismo, que

permite siempre operar con el subsistema de las $n-k$ ecuaciones correspondientes a las variables endógenas. En tanto en cuanto tales características de exogeneidad son correctas no generan distorsiones en las conclusiones obtenidas a partir del sistema reducido de $n-k$ ecuaciones. Adicionalmente, el analista puede plantearse aislar su ecuación (o ecuaciones) de interés del sistema reducido.

Las posibilidades y costes de realizar tal aislamiento dependen de la finalidad del análisis: predicción o estimación. Pagando como coste una pérdida de eficiencia el análisis uniecuacional es, en general, posible para fines de estimación de los parámetros. Para fines predictivos o de simulación el aislamiento de una ecuación sólo es posible si sus variables explicativas son fuertemente exógenas y superexógenas, respectivamente.

Habiendo insistido en los posibles costes que entraña el análisis uniecuacional, conviene señalar una ventaja importante del mismo. En efecto, el enfoque multiecuacional tiene las propiedades de eficiencia comentadas, siempre que el analista sea capaz de formular adecuadamente todo el sistema; en caso contrario, los errores cometidos en la especificación de una ecuación se trasladan, en general, a las restantes. Por tanto, cuando sobre la ecuación de la variable de interés Y_h se tiene bastante seguridad sobre su formulación, pero en la formulación de las restantes ecuaciones que completan el sistema de variables endógenas que intervienen en la explicación de Y_h se tienen grandes incertidumbres o incluso ignorancia sobre cómo proceder, el enfoque uniecuacional es más seguro, pues para su estimación se requiere simplemente ser capaz de formular un conjunto válido de instrumentos. En tales casos, el enfoque uniecuacional tiene una solidez y firmeza, frente a errores de especificación en las otras ecuaciones, del que el enfoque multiecuacional carece y hace que aquél sea preferible a éste.

III.3. El principio “de lo general a lo particular” en la modelización econométrica.

Con el proceso reductor hemos concluido con dos tipos básicos de modelos:

a) Modelos que emplean solamente la condicionalización derivada de la recursividad temporal (modelo II.14 ó II.24)

$$X_t = E(X_t/D_{1t}) + e_{xt},$$

que luego se concretaba en el modelo VARMA (p,q) , que era un modelo de forma reducida muy general o sin condicionalización contemporánea.

b) Modelos que emplean también la condicionalización contemporánea (modelo II.17 ó II.25)

$$Y_t = E(Y_t/D_{2t}) + e_{yt},$$

que hemos visto que se concretaban en el modelo VARMAX, que era un modelo de forma reducida con variables exógenas.

Además, si la teoría proponía relaciones funcionales entre las variables que implicaban simultaneidad, pero se disponía de suficientes restricciones para la identificación estructural del modelo, se obtenía, como un caso particular de los anteriores, el modelo SEM, cuya formulación general se recoge en (II.55).

En tanto en cuanto el modelo SEM ya implica reducciones sobre los dos anteriores, que pueden ser contrastadas, vamos a ceñir la discusión siguiente a los modelos (II.24) ó (II.25). Ambos suponen una reducción razonable del PGD, excluyendo o incorporando la condicionalización contemporánea, pero contienen una formulación dinámica muy general. Por tanto, en el proceso de modelización econométrica ambos modelos pasan a ser hipótesis mantenidas de partida, bajo el supuesto o no de condicionalización contemporánea, y en tal proceso se intentará obtener una versión restringida de la correspondiente hipótesis mantenida, que sea compatible con los datos. Es decir, partiendo de un modelo general se intentará llegar a uno particular que tenga soporte en los datos.

El modelo resultante podrá considerarse como un modelo parco en la parametrización, pero derivado de una hipótesis mantenida que recoge las características de los datos y las implicaciones de la teoría.

El principio de lo general a lo particular en la modelización dinámica lineal se defiende en Anderson (1971) y ha sido difundido a través de la denominada metodología de la LSE, véase Sargan (1964) y Hendry et al. (1984). En el contexto dinámico y lineal este procedimiento de lo general a lo particular parece el procedimiento natural (véase Phillips (1988) pág. 351) y el que puede conducir a mejores resultados empíricos. Sin embargo, este procedimiento no tiene un traslado claro e inmediato a un contexto fuertemente no lineal. En dicho marco perfilar la naturaleza no lineal proponiendo un esquema que lo aproxime constituye normalmente el avance más rentable en tal tipo de modelización. No obstante, estas propuestas de aproximación de un determinado comportamiento no-lineal son difíciles de realizar, y en ellas la estrategia de formulaciones no-lineales generales no parece muy recomendable, véase Granger y Teräsvirta (1992). En bastantes problemas la aproximación lineal por tramos de un comportamiento no-lineal puede ser útil, véase Cancelo y Espasa (1991) y Álvarez et al. (1992). Si en dichos casos, se dispone de muchas observaciones se puede ser general formulando el número de tramos y permitiendo a su vez una especificación dinámica amplia en cada uno de ellos. En Espasa y Cancelo (1991) se encuentra una aplicación de ese procedimiento a la modelización diaria.

Una ilustración altamente pedagógica del interés de proceder de lo general a lo particular en la modelización econométrica lineal se encuentra en Hendry et al. (1984), sección 2.6. La aplicación de este principio implica no detener el proceso reductor en la hipótesis mantenida (II.24) o (II.25), sino continuarlo hasta obtener un modelo más parco en parámetros, pero todavía aceptables para representar las características de los datos.

El ejemplo que consideran Hendry et al. (1984) se refiere a un contexto de información limitada, concretamente uniecuacional, utilizando en la restricción de memoria el truncamiento temporal. Así, el modelo es pues del tipo (II.34), en donde se supone que el truncamiento para $p=1$ es el válido. Con ello, la aplicación del proceso

reductor de Haavelmo conduce al siguiente modelo como hipótesis mantenida o general:

$$y_t = \alpha_1 y_{t-1} + \beta_0 z_t + \beta_1 z_{t-1} + a_t. \quad (\text{III.1})$$

Este modelo, en el que se supone que z es débilmente exógena, siendo tan sencillo contiene a otros nueve como casos particulares del mismo, tal como se refleja en el cuadro 2 construido a partir del cuadro que presentan Hendry et al. (1984) pág. 1042. Todos esos nueve casos suponen restricciones respecto a (III.1) y, por tanto, son contrastables. Es más, la distintas alternativas recogidas en el cuadro 2, suponen estructuras de respuesta dinámica y relaciones a largo plazo muy diferentes.

Los nueve casos del cuadro 2 se generalizan para un truncamiento dinámico superior a uno y para más variables. El ejemplo pone de manifiesto lo erróneo que es modelizar una relación entre variables a partir de un esquema preconcebido, modelo de indicador adelantado, modelo de ajuste parcial, etc., sin insertarlo dentro de un proceso reductor en el que se contrastan las restricciones que dicho modelo incorpora respecto a una hipótesis mantenida que aproxima adecuadamente el proceso generador de los datos.

Este principio de lo general a lo particular es también aplicable a modelos con estructuras dinámicas racionales, tal como se apuntó en el epígrafe II.3. Un comentario más extenso al respecto se encuentra en Espasa y Cancelo (1992), capítulo III.

Cuadro 2

Modelos derivado de		
$y_t = \alpha_1 y_{t-1} + \beta_0 z_t + \beta_1 z_{t-1} + a_t$		
Modelo derivado (denominación)	Ecuación	Restricción
1) Regresión estática	$y_t = \beta_0 z_t + a_t$	$\alpha_1 = \beta_1 = 0$
2) Modelo univariante de series temporales	$y_t = \alpha_1 y_{t-1} + a_t$	$\beta_0 = \beta_1 = 0$
3) Modelo series diferenciadas	$\Delta y_t = \beta_0 \Delta z_t + a_t$	$\alpha_1 = 1; \beta_1 = -\beta_0$
4) Modelo con indicador adelantado	$y_t = \beta_1 z_{t-1} + a_t$	$\alpha_1 = \beta_0 = 0$
5) Modelo de retardos distribuidos	$y_t = \beta_0 z_t + \beta_1 z_{t-1} + a_t$	$\alpha_1 = 0$
6) Modelo de ajuste parcial	$y_t = \alpha_1 y_{t-1} + \beta_0 z_t + a_t$	$\beta_1 = 0$
7) Modelo de factores comunes (errores autorregresivos)	$\begin{cases} y_t = \beta_0 z_t + u_t \\ u_t = \alpha_1 u_{t-1} + a_t \end{cases}$	$\beta_1 = -\beta_0 \alpha_1$
8) Modelo con mecanismo de corrección del error	$\Delta y_t = \beta_0 \Delta z_t + (1 - \alpha_1)(y - z)_{t-1} + a_t$	$1 = \alpha_1 + \beta_0 + \beta_1$
9) Modelo de forma reducida sin incluir la condicionalización contemporánea	$y_t = \beta_1 z_{t-1} + \alpha_1 y_{t-1} + a_t$	$\beta_0 = 0$

III.4. El planteamiento de información limitada en sistemas cointegrados.

Cuando un sistema de ecuaciones incorpora variables endógenas no estacionarias que están cointegradas¹⁵, los problemas que aparecen respecto a sistemas estacionarios son varios. Phillips (1991) realiza un excelente tratamiento de esta cuestión. Los principales resultados de dicho trabajo se pueden resumir así:

1) Es importante que la estimación de sistemas cointegrados se realice incorporando la información sobre las raíces unitarias presentes. Con ello los estimadores tienen propiedades asintóticas óptimas, y los contrastes de hipótesis pueden realizarse mediante los procedimientos habituales.

2) Por el contrario, si las raíces unitarias se estiman como unos parámetros más del sistema, se obtiene que las distribuciones asintóticas necesarias para la inferencia no son distribuciones de tipo estándar, con lo que es necesario tabular tales distribuciones¹⁶. Esto es cierto, tal como puntualiza Sims et al. (1990), para las variables que no se pueden expresar en una reparametrización del modelo como variables $I(0)$ con media nula. En dichas tabulaciones hay que tener en cuenta el efecto de los parámetros molestos (nuisance parameters), que además tienen que ser estimados. En tales casos no se dispone de una teoría asintótica óptima aplicable.

3) En la estimación e inferencia de sistemas cointegrados lo importante no es la forma precisa de especificar el modelo, por ejemplo en la forma de mecanismo de corrección del error, sino la incorporación de la información sobre raíces unitarias, que también se puede realizar especificando las variables dependientes de cada ecuación en niveles y las explicativas en niveles y diferencias.

4) La reformulación de modelos de información limitada, incluso modelos uniecuacionales, es posible y relativamente fácil en sistemas en los que sólo aparecen las relaciones a largo plazo, es decir, que no incorporan una estructura dinámica

¹⁵ Si no hay ninguna relación de cointegración entre ellas, diferenciando las variables se obtiene un sistema estacionario.

¹⁶ Una alternativa es la empleada por Ferreti y Romo (1992) para un modelo univariante AR(1) consistente en utilizar métodos de remuestreo.

transitoria. En dichos casos los modelos de información limitada se formulan incluyendo las variables explicativas en niveles y en diferencias presentes y pasadas. Con tales modelos se pueden obtener estimaciones óptimas de las relaciones a lo largo o relaciones de cointegración.

5) Si estos sistemas incluyen formulación dinámica transitoria con variables endógenas contemporáneas además de las relaciones de largo plazo, la reformulación necesaria para la construcción de modelos de información limitada óptimos puede ser compleja, así como requerir también diferenciaciones futuras de las variables explicativas endógenas. En tales casos Phillips a su vez menciona la estimación semiparamétrica.

6) Para la estimación óptima de las relaciones de cointegración sólo se requieren estimadores consistentes de los parámetros que recogen la estructura dinámica transitoria, incluso en el caso en que estos parámetros y los de las relaciones de cointegración no tengan campos de variabilidad independientes entre sí.

7) Con fines de estimación e inferencia el enfoque uniecuacional sólo es válido si las variables explicativas son fuertemente exógenas.

8) Si se aplica un enfoque uniecuacional cuando no se cumple la propiedad de exogeneidad anterior, se tiene que los estimadores resultantes incorporan los inconvenientes señalados en el punto segundo. Por ello en sistemas cointegrados el tratamiento simultáneo es más apremiante que en el contexto clásico de variables estacionarias.

Respecto a los modelos VAR, la presencia de variables no estacionarias y cointegradas tiene las siguientes implicaciones:

a) El problema de no estacionariedad no se puede solucionar formulando un modelo VAR sobre variables diferenciadas. Esta formulación es incorrecta ya que omite los términos –mecanismos de corrección del error- que recogen las relaciones de largo plazo entre los niveles de las variables.

b) Si se formula el modelo VAR sobre las variables en niveles, el modelo no incorpora la información sobre las raíces unitarias presentes entre las variables, sino que dichas raíces se estiman. En tal caso los estimadores obtenidos tienen los problemas señalados más arriba en el punto segundo.

c) Una solución consiste en formular el modelo VAR sobre series diferenciadas incluyendo los términos correspondientes a los mecanismos de corrección del error. Al modelo resultante se le denomina en ocasiones VECM (Vector Error Correction Mechanism). Un tratamiento exhaustivo de su estimación por el método de la máxima verosimilitud, que permite contrastar hipótesis sobre el número de vectores de cointegración existentes, puede encontrarse en Johansen (1988 y 1991).

III.5. Convergencia hacia el modelo VECM.

Como comentario final se puede señalar que la teoría de la cointegración ha proporcionado una terminología que se ha aceptado y sobre todo asumido tanto en econometría como en macroeconomía, lo que ha servido para establecer una comunicación mucho más fluida entre ambas disciplinas. Además, la cointegración está sirviendo para unificar distintas metodologías econométricas. En Granger (ed., 1990) se recogen una serie de trabajos conducentes a ilustrar el estado actual de la econometría de series temporales, y en ellos se presentan las principales metodologías existentes en la actualidad. Pues bien, las siguientes metodologías: (a) de modelos uniecuacionales dinámicos con mecanismo de corrección del error (LSE), (b) de modelos VAR, o (c) de ecuaciones simultáneas, están mostrando una fuerte convergencia con el desarrollo de la teoría de la cointegración.

En efecto, la teoría de la cointegración ha puesto de manifiesto que la formulación de modelos simultáneos puede exigir la estimación de raíces unitarias, lo que conduce a problemas en la estimación y sobre todo en la contrastación a partir de tales estimaciones. Por eso, la formulación en términos de mecanismos de corrección del error es mucho más interesante, ya que con ellos los problemas anteriores desaparecen.

Sin embargo, la formulación que se emplea en la metodología de la LSE de modelos de información limitada puede ser, en presencia de variables cointegradas, extraordinariamente más compleja, con lo que la formulación del sistema multiecuacional parece más importante. Este sistema puede ser del tipo VECM

mencionado anteriormente, que supone una ampliación de los VAR tradicionales mediante la incorporación de mecanismos de corrección del error.

Todos estos puntos aparecen tratados con profundidad en los excelentes trabajos de Phillips (1991) y Johansen (1988, 1991 y 1992).

IV. DISEÑO Y EVALUACIÓN DE MODELOS ECONOMÉTRICOS

IV.1. Implementación el principio de adecuación estadística en la modelización econométrica.

De la discusión realizada en las secciones anteriores sobresale lo que se entiende, en esta Memoria, por modelización paramétrica. Para precisar este término emplearemos las palabras de Hendry y Ericsson (1991), pág. 18. Así, la modelización paramétrica en econometría consiste en caracterizar las propiedades de los datos, cumpliendo los siguientes requisitos:

- a) mediante relaciones paramétricas simples
- b) que sean interpretables a la luz de los conocimientos que proporciona la Teoría Económica,
- c) permaneciendo “razonablemente” constantes a lo largo del tiempo, y
- d) proporcionando una explicación al tipo de resultados obtenidos por modelos precedentes.

Esta última característica (d) no ha aparecido explícitamente en la discusión anterior, si bien el aspecto subyacente en ella, correspondiente a una estrategia progresiva de investigación, ha sido comentado en el punto I.1. Sobre ello volveremos más adelante.

En la construcción de un modelo econométrico está presente el proceso reductor descrito en la sección II, en el que se introducen supuestos que implican, en general, restricciones sobre el comportamiento de las variables observadas, por lo que son susceptibles de contrastación. En consecuencia, para asegurarnos que el modelo empleado en un determinado estudio cumple con el principio de adecuación estadística definido en I.6, será necesario evaluarlo. Asimismo, en el proceso de diseño de un modelo habrá que contrastar si la característica que una determinada reducción supone es compatible con los datos.

Con ello tenemos que en el aprendizaje econométrico, siendo importantes las técnicas de estimación, que permiten obtener intervalos de valores para los parámetros de interés, no constituyen el único tema central de la disciplina. Las técnicas de estimación han dominado excesivamente en una orientación

econométrica, que podemos clarificar de incorrecta, pero muy extendida, a la que Spanos (1986) denomina “enfoque econométrico de libro de texto” y Gilbert (1986) “enfoque econométrico AER” (“average economic regression”). En la elección de esas siglas por parte de Gilbert pesa mucho la discusión entre Friedman-Schwartz y Hendry-Ericsson sobre la demanda de dinero en el Reino Unido (véase Friedman y Schwartz, 1982, 1991) y Hendry y Ericsson (1991) y la discusión metodológica a micrófono abierto entre Hendry, Leamer y Poirier plasmada en la revista *Econometric Theory*, v.6, 1990, págs. 171 a 261, sobre el subjetivismo que subyace en la modelización econométrica (véase también Pagan (1987) y Phillips (1988) para una evaluación crítica de dichas metodologías).

En tal orientación econométrica, el objetivo de la Econometría es estimar una supuesta relación teórica exacta, a la que se le añade un término de error, generalmente de tipo ruido blanco, que recogería los errores de medida en la variable dependiente así como los efectos de las variables omitidas. En este enfoque la teoría da cuenta de toda la información que hay en los datos, excepto un cierto error de carácter no sistemático, que es el único elemento aleatorio en el modelo econométrico. Es decir, los supuestos estocásticos se realizan exclusivamente sobre los errores. Como señala Spanos (1986), en el capítulo introductorio, esto supone utilizar la Econometría para proveer cifras a nuestras propias construcciones rechazando el objetivo de explicar los fenómenos de interés.

En contraposición a lo anterior, en el enfoque econométrico que se deriva de lo que llevamos dicho, se tiene que los temas centrales de la disciplina vienen determinados por una metodología de “diseño, estimación y evaluación” necesaria para la construcción de modelos. Como señalan Ericsson y Hendry (1985), pág. 261, este enfoque está dentro del espíritu de la metodología Box-Jenkins para el análisis univariante. No obstante, en el enfoque econométrico en el que esta memoria se entronca, se pone cierto énfasis en el modelo contemplado más general, con el fin de establecer una media para la varianza innovacional. Además, la Teoría Económica y, como se explica más adelante, los modelos previamente existentes juegan también un papel importante.

En consecuencia, para perfilar la enseñanza de la Econometría, es necesario estructurar la problemática de diseño y evaluación de modelos. Tales actividades en el proceso de modelización giran en torno a un conjunto de criterios asociados a características que, en las distintas etapas del proceso reductor, se supone que deben satisfacer los modelos y que se contrastan mediante una batería de estadísticos. Siguiendo la tradición establecida por Hendry y Richard (1982) y recogida, por ejemplo, en Ericsson y Hendry (1985), Hendry y Ericsson (1991) y Ericsson et al. (1990) a quienes seguimos y en los que basamos el desarrollo en este punto, los pasos en el proceso reductor pueden suponer la pérdida de información, por lo que los contrastes se discutirán relacionándolos con la información que está disponible al analista y refiriéndolos al proceso reductor. Al mismo tiempo se pretenderá ilustrar el modo de utilización de los mismos con fines de diseño o evaluación. Básicamente, desde el punto de vista del diseño, satisfacer un criterio se interpreta como que no hay pérdida de información por adoptar la reducción asociada al mismo, por lo que se le considerará compatible con los datos. Desde el punto de vista de la evaluación un criterio es una hipótesis nula, sobre la que se desea determinar su validez.

Los criterios que se van a aplicar en la modelización econométrica son:

1. Los errores del modelo deben ser innovaciones y corresponder a la distribución que se ha supuesto. Que los errores no sean innovaciones implica que tienen un comportamiento sistemático que puede explicarse a partir del conjunto informativo condicionante y el modelo, por tanto, es incorrecto. En efecto, el error de un modelo viene dado por

$$e_t = Y_t - E(Y_t/\mathbb{D}_t), \quad (\text{IV.1})$$

en donde \mathbb{D}_t es el conjunto informativo condicionante. En (IV.1) e_t es por construcción ortogonal a \mathbb{D}_t , por lo tanto, si el modelo es correcto para los datos analizados, los errores calculados (estimados) con dichos datos deben ser innovacionales. Es decir, no pueden ser predecibles ni contener estructura alguna que sea explicable en función \mathbb{D}_t .

Este criterio se basa en el hecho de que partiendo del principio de que la aleatoriedad está en las variables económicas, los errores de los modelos son elementos genuinos sino derivados. Su derivación es de forma que son desviaciones sobre una esperanza matemática condicional respecto a un conjunto informativo que se considera relevante, por tanto, son ortogonales a él. Para que el modelo formulado sea adecuado a los datos empleados, los errores obtenidos con tales datos deben cumplir la propiedad de ortogonalidad que da justificación al modelo.

Este es, por tanto, uno de los principales criterios empleados en un buen proceso de modelización, independientemente de la estrategia que en él se emplee o de los datos sobre los que se aplique.

2. Las variables explicativas débilmente exógenas deben tener ciertamente dicha propiedad.

De lo contrario la distribución condicional respecto a la que se ha deducido el modelo contemplado supone una pérdida de información respecto a la correspondiente distribución conjunta.

3. Los parámetros del modelo deben ser razonablemente constantes y, en consecuencia, el modelo debe ser apto para la predicción.

4. El modelo debe ser consistente con la teoría que explica el comportamiento de las variables consideradas.

En términos generales esto resulta algo vago, pero implica que el modelo debe ser capaz de reproducir los resultados de la Teoría en la que supuestamente se ha basado. Una implicación clara de este criterio se refiere a las relaciones teóricas de largo plazo entre variables económicas. Si la Teoría establece una relación de equilibrio a largo plazo entre las variables de un modelo, por ejemplo, uniecuacional, es decir, se trata de variables plenamente cointegradas, entonces los residuos del modelo han de ser estacionarios.

La consistencia con la teoría requiere que el modelo no puede recoger valores en los parámetros que la teoría descarta.

5. El modelo debe ser compatible con el sistema de medida.

Esto implica que el modelo no puede generar predicciones que sean imposibles porque no respetan determinadas restricciones que definen a los datos. Por

ejemplo, la restricción presupuestaria, el que los precios no pueden ser negativos, que el empleo no puede ser superior a la población activa, etc.

En los casos que el sistema de medida requiera que una variable tome siempre valores positivos, tal característica puede asegurarse mediante el empleo de logaritmos. No obstante, imponer esa propiedad en un modelo que no se deriva del proceso generador de los datos puede conducir a que el modelo estimado no cumpla los criterios (1) a (3).

6. El modelo debe comprender, manteniendo una parquedad razonable en su estructura, los resultados obtenidos por modelos previos.

Es decir, se pretende que el modelo resultante además de ser adecuado según los criterios anteriores, sea una representación suficiente de resultados empíricos de modelos previos. La idea es la siguiente. Si el modelo bajo consideración, dígase α , es una derivación adecuada del proceso generador de los datos, debe contenerse en él una explicación sobre los resultados obtenidos con modelos previos, dígase β . Si α no es capaz de explicar por qué se obtienen los resultados de β , es evidencia de que α no recoge adecuadamente las características del proceso generador de los datos, ya que de éste se derivó en modelos β con unas determinadas características, sobre las que el modelo α no es capaz de dar razón. Con este principio a la modelización econométrica se le exige ir más allá de la obtención de modelos compatibles con los datos que se puedan interpretar de acuerdo con los resultados de la Teoría, se le requiere además que el modelo propuesto englobe los avances anteriores. De esta forma, la investigación empírica no sólo es útil para entender el mundo real, sino que se enmarca en una dinámica de avance progresivo del conocimiento.

El principio de comprensión (“encompassing”) implica que el modelo α es capaz de anticipar problemas con los que los modelos β se van a encontrar en el futuro. Así, puede explicar por qué los modelos β van a generar predicciones poco ajustadas a cambios futuros, por qué sus parámetros no van a ser estables, etc. Que posteriormente se confirmen tales aspectos es una de las mayores garantías que se pueden obtener sobre α .

Ciertamente el principio de comprensión sólo tiene sentido que se aplique sobre los modelos β que han pasado los criterios precedentes. Por otro lado, dado que un modelos general siempre comprenderá a los modelos que sean casos particulares del mismo, no tiene especial sentido cumplir con el principio de comprensión mediante el diseño de un modelo muy amplio que incorpore a los modelos previos. En línea con Hendry y Richard (1989), se trata de obtener un modelo comprensivo de los anteriores, en el sentido de poder explicar sus resultados, pero incorporando al mismo tiempo una parquedad razonable en su estructura. En este sentido, como afirman Hendry y Ericsson (1991), el principio de comprensión determina los límites aplicables en el proceso reductor, tanto para establecer la hipótesis mantenida, como para lograr un modelo particular a partir de ella. Es decir, el principio de comprensión supone introducir límites en el grado de parquedad en la parametrización, en aras a poder dar cuenta de resultados previamente establecidos en otros modelos. Obsérvese que este ligero alejamiento del principio de parquedad no es arbitrario, sino que se deriva del hecho de pretender incorporar la información contenida en los datos empleados en modelos previos.

Una ilustración muy útil del principio de comprensión se encuentra en Ericsson y Hendry (1985), pág. 259, que pasamos a comentar.

Supongamos que se tienen dos modelos uniecuacionales diferentes procedentes de las siguientes hipótesis

$$H_1 : E(y_t/Z_{1t}) = Z'_{1t}\beta_1 \quad (IV.2)$$

y

$$H_2 : E(y_t/Z_{2t}) = Z'_{2t}\beta_2 \quad (IV.3)$$

en las que se establece además que los correspondientes residuos e_{1t} y e_{2t} son ruido blanco con una distribución normal con varianzas σ_j^2 . Para simplificar la exposición se supondrá que Z_{1t} y Z_{2t} no tienen variables comunes.

Si ante la existencia de ambos modelos se contempla que el conjunto informativo relevante viene dado por $X_t = (y_t, Z_{1t}', Z_{2t}')'$, tenemos que la función de densidad conjunta de los datos se puede factorizar de la siguiente forma:

$$D(y_t / Z_{1t}, Z_{2t}; \cdot) D(Z_{1t} / Z_{2t}; \cdot) D(Z_{2t}; \cdot) \quad (IV.4)$$

El modelo (IV.2) implica que Z_2 es irrelevante para explicar “y” dada Z_1 . Ahora bien, a partir de la función de densidad conjunta de (Z_1, Z_2) y bajo el supuesto de normalidad se tiene que

$$Z_{1t} = \pi Z_{2t} + \varphi_{1t}, \quad (IV.5)$$

$$E(\varphi_{1t} \varphi_{1t}') = \Omega.$$

Utilizando (IV.2) y (IV.5) se obtiene:

$$y_t = \beta_1' \pi Z_{2t} + (e_{1t} + \beta_1' \varphi_{1t}) = \beta_2' Z_{2t} + e_{2t}. \quad (IV.6)$$

Es decir (IV.6) es lo que de acuerdo con el modelo (IV.2) debe encontrarse en el modelo (IV.3). Por tanto, el modelo (IV.2) comprenderá al (IV.3) si cumple que

$$H_a : \beta_2 = \pi' \beta_1 \quad (IV.7)$$

y

$$H_b : \sigma_2^2 = \sigma_1^2 + \beta_1' \Omega \beta_1. \quad (IV.8)$$

A la hipótesis H_a se le denomina comprensión paramétrica y a la H_b comprensión de varianza. Se puede demostrar que denominación en la varianza ($\sigma_1^2 < \sigma_2^2$) es condición necesaria para un modelo comprenda en varianza a otro, pero no es condición suficiente.

Como se ha dicho, si con el modelo (IV.2) se pretende que se está realizando una caracterización adecuada del proceso generador de los datos, debe permitir derivar de él los resultados del modelo (IV.3). En tal caso, los resultados del modelo (IV.3), β_2 y σ_2^2 , no deben diferir de forma significativa, de los que se derivan del modelo (IV.2), $\pi' \beta_1$ y $(\sigma_1^2 + \beta_1' \Omega \beta_1)$. Una diferencia significativa implicaría que (IV.2) no comprende a (IV.3) y que, por tanto, no constituye una derivación adecuada del proceso generador de los datos. Sobre contrastes de comprensión puede verse, Mizón y Richard (1986), Mizon (1984), Mackinnon (1983) y Pesaran (1982).

En el proceso de modelización los criterios anteriores tienen, tal como se destaca en diferentes trabajos de Hendry y se recoge muy adecuadamente en Ericsson et al. (1990), dos tipos de aplicaciones. Una constructiva de descubrimiento, empleada en el proceso de diseño de modelo, y otra destructiva, de justificación, empleada en la tarea de evaluación de un modelo. En esta última se trata de ver si un criterio, una característica o reducción que incorpora o se supone en el modelo, (hipótesis nula) se puede considerar válida (no se rechaza) dentro de un conjunto de información en el que se pueden contemplar formulaciones menos restrictivas (hipótesis alternativa).

En el aspecto constructivo o de diseño del modelo, si un criterio no se cumple, la acción a emprender en la tarea de reespecificación no es, normalmente, en absoluto obvia. La razón está en que dos contrastes suelen tener potencia respecto a diversas alternativas y es una cuestión de habilidad por parte del constructor del modelo el descubrir, entre las diferentes posibilidades que tales alternativas suponen, la dirección de progreso más adecuada para la modelización de los datos que está analizando. Así, por ejemplo, rechazar la hipótesis de ruido blanco para los residuos a través del estadístico Durbi-Watson, que contempla como hipótesis alternativa que tales residuos tienen autocorrelación de primer orden, no implica en

absoluto que lo que los datos necesitan es ampliar el modelo con una estructura autorregresiva que tales residuos tienen autocorrelación de primer orden, modelos con factores comunes en el cuadro 2. Pueden deberse, tal como señala Sargan (1964), a un error en la especificación dinámica del modelo, de modo que la restricción de factores comunes respecto la hipótesis mantenida, que se comentaba al introducir el cuadro 2, no sea correcta. Pero incluso todavía más, la especificación dinámica puede ser correcta y tratarse de una variable omitida. La conclusión es clara, en cuanto a la labor de re-especificación a emprender con motivo de haberse rechazado uno de los criterios analizados, no se trata, tal como afirma Spanos (1989), de reformular el modelo de modo que se “cure” el “síntoma” que se recoge en la hipótesis alternativa para la que el contraste ha sido diseñado. Rechazar la hipótesis nula no implica que la alternativa sea cierta. Tal rechazo tiene indicaciones mucho más vagas. Supone que el proceso reductor empleado para llegar al modelo contrastado no es válido, por lo que la acción procedente es revisar los diferentes pasos en dicha reducción y optar por posturas menos restrictivas. Siguiendo con el símil anterior, detectado un determinado síntoma, la cura aplicar debe diseñarse a partir de un estudio de todo el historial reductor que configura el modelo sometido a la consulta del experto.

Para relacionar los criterios anteriores con los conjuntos de información disponibles al analista, conviene sistematizar la información según proponen Hendry y Richard (1982). Así, se pueden distinguir cuatro conjuntos de información: (I) los datos de propio modelo analizando, (II) la procedente de la Teoría Económica, (III) la correspondiente al sistema de medida de los datos y (IV) la relativa a los datos de modelos alternativos. Los conjuntos (I) y (IV) se pueden, a su vez, subdividir en lo referente; (a) al pasado, (b) presente y (c) futuro. Con tal sistematización en el cuadro 3, construido a partir de los correspondientes cuadros en Hendry y Ericsson (1991) y Ericsson et al. (1990), se relacionan los criterios anteriores con el conjunto de información al que hacen referencia y se señala la restricción del proceso reductor a la que están asociados.

Determinados criterios, hipótesis nulas, se pueden enfrentar ante diferentes alternativas dando lugar a distintos contrastes. Un resumen de ello se encuentra en el cuadro 4, construido a partir de las mismas referencias empleadas para el cuadro 3.

CRITERIOS EN LAS ESTRATEGIAS DE DISEÑO Y EVALUACIÓN DE MODELOS		
CRITERIOS	CONJUNTO DE INFORMACIÓN CON EL QUE ESTÁ RELACIONADO	CARACTERISTICA ASOCIADA EN EL PROCESO REDUCTOR
<p>1.- Los errores del modelo deben ser innovacionales y corresponder a la función de distribución que se ha supuesto.</p> <p>2.- Variables explicativas débilmente exógenas</p> <p>3.- Parámetros constantes (predicciones adecuadas)</p> <p>4.- El modelo debe ser consistentes con la Teoría.</p> <p>5.- El modelo debe ser compatible con el sistema de medida de los datos.</p> <p>6.- El modelo debe comprender ("encompass") los resultados de modelos previos.</p>	<p>I.- DATOS EMPLEADOS EN EL MODELO</p> <p>I.a.- EL PASADO</p> <p>I.b.- EL PRESENTE</p> <p>II.c.- EL FUTURO</p> <p>II.- PROPOSICIONES Y RESULTADOS DE LA TEORÍA ECONÓMICA</p> <p>III.- CARACTERÍSTICAS DEL SISTEMA DE MEDIDA DE LOS DATOS</p> <p>IV.- DATOS EMPLEADOS EN MODELOS PREVIOS</p>	<p>Restricciones en la dependencia de los datos (Memoria) y en la forma de la función de distribución.</p> <p>Restricciones en la condicionalización contemporánea.</p> <p>Restricciones de homogeneidad.</p>

**HIPÓTESIS ALTERNATIVAS Y ESTADÍSTICOS EN LA CONTRASTACIÓN DE CRITERIOS EN EL DISEÑO Y EVALUACIÓN DE
MODELOS**

HIPÓTESIS NULA (criterio)	HIPÓTESIS ALTERNATIVA	REFERENCIA SOBRE EL ESTADÍSTICO PROPUESTO PARA CONTRASTE
1.- ERRORES INNOVACIONALES “ “ “ “	Autocorrelación residual de primer orden Autocorrelación residual de orden j Residuos ARCH de orden j Heteroscedasticidad residual en función de los regresores Restricciones inválidas sobre los parámetros RESET (“regression equation specification error test”) de orden j	Durbin y Watson (1950), (1951) Box y Pierce (1970), Ljung y Box (1978), Godfrey (1978), Harvey (1981, pág. 173) Engle (1982) White (1980, pág. 825), Nicholls y Pagan (1983), White (1984). Johnson (1963, pág. 126) Ramsey (1969)
2.- ERRORES GAUSSIANOS 3.- VARIABLES EXPLICATIVAS DEBILMENTE EXÓGENAS 4.- PARAMETROS CONSTANTES “	Asimetría y exceso de curtosis q variables instrumentales no son independientes de los errores Parámetros no constantes en las submuestras Predicciones incompatible en una submuestra de q observaciones	Jarque y Bera (1980) Sargan (1958, 1964), srgan (1980, pág. 1136), Engle et al. (1983) Fisher (1922), Chow (1960, pág. 595 a 602) Chow (1960, pág. 594-5)
5.- CONSISTENCIA TEÓRICA - cointegración 6.- COMPRENSIÓN DE MODELOS PREVIOS	Ausencia de cointegración Ajuste relativamente pobre	Engle y Granger (1987) Hendry y Richard (1982)
- dominación en la varianza - comprensión de varianza - comprensión paramétrica	Varianza residual inexplicable Variables adicionales significativas	Cox (1961, 1962), Pesaran (1974), Hendry (1983) Jonsonston (1963, pág. 126), Mizon y Richard (1986)

IV.2. Resumen sobre el proceso de modelización.

Como conclusión se puede decir que a los modelos econométricos se llega, tal como se viene insistiendo en esta memoria, en dos etapas. Primero, a través de un proceso reductor se obtiene un modelo amplio, que se convierte en hipótesis mantenida, y, luego, se busca restringir tal esquema general para obtener un modelo particular. En la obtención de éste se ha podido optar por un enfoque de información limitada, que en su versión extrema conduce a la modelización uniecuacional. Este enfoque supone aceptar una pérdida de información en aras a simplificar o incluso a poder hacer factible la modelización de las variables de interés. Esa pérdida de información conlleva, en general, un coste de eficiencia en el modelo resultante, pero garantizado, dentro de un proceso adecuado de modelización, la consistencia de la estimación. El modelo particular alcanzado debe cumplir una serie de criterios, listados en el cuadro 3, que garanticen, que dentro del contexto informativo elegido (limitado o completo), no existe pérdida de información debido a las reducciones y restricciones empleadas. La información disponible se organiza en el cuadro 3 en distintos conjuntos, con lo que los criterios citados se vinculan a un conjunto informativo específico, al tiempo que se señala la característica del proceso reductor a lo que están asociados.

En el cuadro 4 se listan distintos contrastes que sirvan para investigar si el modelo cumple o no con los diferentes criterios. Como señalan Hendry y Ericsson (1991) se dirá que un modelo es “congruente” si satisface los criterios mencionados y, como derivación de ello, capta las principales características de los datos y ofrece inferencias, de confianza, sobre cuestiones económicas.

El enfoque econométrico descrito está encaminado a que se puedan obtener conclusiones sensatas y sostenibles sobre la realidad económica. Sin embargo, como muy bien advierten Ericsson y Hendry (1985), págs. 260-61, es posible que se produzcan conflictos entre los criterios señalados. Así, por ejemplo, el criterio de compatibilidad con el sistema de medida puede ser difícil de cumplir sin mantener que los errores se derivan de funciones de distribución truncadas con bandas que varían convenientemente con el tiempo, o el criterio de consistencia teórica puede entrar en conflicto con el de dominación en la varianza en modelos

sujetos al principio de parquedad en su formulación. En esas situaciones, tal como señalan dichos autores, la solución que se proponga dependerá de los objetivos del análisis, pero también de la habilidad del investigador en implementar sugerencias creativas que favorezcan un resultado.

El contexto econométrico presentado en esta memoria ofrece un marco sólido para que la aplicación econométrica concluya con resultados que tengan amplias garantías de objetividad, a la vez que la fundación y sistematización del proceso de modelización permite clarificar y estimular el tipo de aportaciones personales creativas, que el analista empírico debe realizar.

IV.3. Conclusiones para la enseñanza de la Econometría.

El programa que se describe en las secciones siguientes se basa en la idea de que la aplicación econométrica sólo será fructífera, si surge de una acción que se ha desarrollado a partir de una concepción teórica general capaz de ordenar los problemas y sistematizar las vías de solución de los mismos. Así, para que un economista pueda progresar en el arte de la econometría aplicada, necesita tener asimilada una fundamentación teórica en la que se engarcen todos los instrumentos que utiliza en sus aplicaciones.

En consecuencia, la enseñanza de la econometría no puede realizarse proponiendo una secuencia de instrumentos prácticos cada vez más complejos, despreocupándose de la conexión teórica que entre ellos existe, sino que desde el principio deben presentarse los fundamentos estadísticos en los que la tarea econométrica cobra su sentido y justificación, así como el planteamiento que de ellos se deriva para la modelización econométrica. La enseñanza de la Econometría tiene poco sentido, si alguno, si no está orientada hacia la aplicación práctica, y, precisamente, para que tal aplicación pueda hacerse con confianza y produciendo resultados con garantías de objetividad, es necesario que se base en una formación teórica adecuada. Esta unión indisoluble de orientación práctica a partir de una sólida base teórica es la que determina los programas que se presentan en las secciones siguientes.

V. BIBLIOGRAFÍA

Álvarez, L.J., J.C. Delrieu y A. Espasa, (1992), "Aproximación lineal por tramos a comportamientos no lineales: estimación de señales de nivel y de crecimiento", ponencia presentada al International Workshops on Seasonal Adjustment Methods and Diagnostics, Washington. Working Paper, universidad Carlos III de Madrid.

Anderson, T.W., (1971), *The Statistical Analysis of Time Series*, Wiley, Nueva York.

Banerjee, A. y D.F. Hendry, (1992), "testing integration and cointegration: an overview", *Oxford Bulletin of Economics and Statistics*, 54, 3.

Banerjee, A., et al. (1992), *Cointegration, Error-Correction and the econometric analysis of non-stationary data*.

Bollerslev, T., (1986), "Generalized Autoregressive Conditional Heteroscedasticity", *Journal of Econometrics*, 31, 307-327.

Bowden, R. y D. Turkington, (1984), *Instrumental Variables*, Cambridge university Press, Cambridge.

Box, G.E.P. and D.A. pierce, (1970), "Distribution of Residual Autocorrelations in autoregressive-integrated Moving Average Time Series Models", *Journal of the American Statistical Association*, 65, 322, 1509-1526.

Brunner, K. y A.H. Meltzer, (eds. 1976), "Carnegie-Rochester Conference Series on Public Policy: The Phillips Curve and Labor Markets", *Suplement to the Journal of Monetary Economics*, v.1.

Burns, A.F. y W.C. Mitchell, (1946), *Measuring Business Cycles*, New York: National Bureau of Economic Research.

Campos, J., N.R. Ericsson y D.F. Hendry, (1990), "An analogue model of Phase-averaging procedures", *Journal of Econometrics*, v. 43, nº 3, págs. 275-292.

Cancelo, J.R. y A. Espasa, (1991^a), "Forecasting Daily Demand for Electricity with Multiple-Input Nonlinear Transfer Function Models: acase of study", *Working Paper 91-21*, Departamento de Economía de la Universidad Carlos III de Madrid.

Ciomba, P. (1910), *Grundriss einer Oekonometric und die auf der Nationaloekonomie aufgebaute natürliche Theorie der Buchhaltung*, Lwow (Lernberg).

Cox, D.R. (1961), "Tests of Separate Families of Hypotheses", *Journal of the Royal Statistical Society, Series B*, 24, 2, 406-424.

Chow, G.C. (1960), "Tests of Equality between Sets of Coefficients in Two Linear Regressions", *Econometrica*, 28, 3, 591-605.

Delgado, M., (1992), "Semiparametric Generalized Least Squares in the multivariate nonlinear regression model", Econometric Theory, v.8, págs. 203-222.

Delgado, M.A. y P.M. Robinson (1992), "Nonparametric and semiparametric methods for economic research", Journal of Economic surveys, v.6, nº 3, págs. 203-249.

Duesenberry, J., G. Fromm, L.R. Klein y E. Kuh, (1965), The Brookings Quarterly Econometric Model of the USA, North-Holland, Amsterdam.

Durbin, J., (1954), "Errors in Variables", Review of the Institute of International Statistics, v. 22, 23-31.

Durbin, J. and G.S. Watson, (1950), "Testing for Serial Correlation in Least Squares Regression. I", Biometrika, 37, 3 and 4, 409-428.

Durbin, J. and G.S. Watson, (1951), "Testing for Serial Correlation in Least Squares Regression. II", Biometrika, 38, 1 and 2, 159-178.

Engle, R.F., (1982), "Autoregressive Conditional Heteroscedsticity with Estimates of the Variance of United Kingdom Inflation", Econometrica, 50, 4, 987-1007.

Engle, R. y C.W.J. Granger (1987), "Cointegration and Error Correction: Regression, Estimation and Testing", Econometrica, vol. 55, 251-276.

Engle, R.F., G.W.J. Granger, J. Rice y A. Weiss, (1986), "Semiparametric estimates of the relation between weather and electricity sales", Journal of the American Statistical Association, 81, págs. 310-320.

Engle, R.F. y D. Hendry (1989), "Testing SuperExogeneity and Invariance", Discussion Paper 8951, Department of Economics, University of California at San Diego.

Engle, R.F.; D. Hendry y J.F. Richard (1983), "Exogeneity", Econometrica, vol. 51, 277-304.

Epstein, R.J., (1987), A History of Econometrics, North-Holland.

Ericsson, N.R. (1991), "Cointegration, Exogeneity and Policy Analysis; an Overview", International Finance Discussion Papers, num 415, Boards of Governors of the Federal Reserve System.

Ericsson, N.R., J. Campos y H.A. Tran (1990), "PC-GIVE and David Hendry's Econometrics Methodology", Revista de Econometría, v. X, nº 1, págs. 7 a 117, abril.

Ericsson, N.R., y D.F. Hendry, (1985), "Conditional Econometric Modeling: An application to New House Prices in the United Kingdom", Chapter 11 in A.C. Atkinson and S.E. Fienberg (eds) A Celebration of Statistics: The ISI Centenary Volume, New York: Springer-Verlag, 251-285.

Escribano, A., A. Espasa y D. Peña, (1992), "Econometric Models and Multiple Time Series", Proceedings of the American Statistical Association.

Espasa, A. (1973), "Modelos Macroeconómicos: Simultaneidad y Recursividad, Estimación y Observaciones Escasas", Cuadernos de Economía, vol. 1, num 1, 3-26. En Cuadernos de Economía, vol. 1, num 2, se encuentra una lista de erratas de este artículo.

Espasa, A. (1975), A wage-prices inflation model for United Kingdom, 1950-1970: Its specification and estimation by classical and spectral methods, Tesis doctoral no publicada, London School of Economics and Political Science.

Espasa, A., y J.R. Cancelo, (1989), "Modelos univariantes en el análisis económico", Revista Española de Economía, v. 6, nº 1-2, 85-107.

Espasa, A. y J.R. Cancelo, (1991), "Model Bases Measures of Contemporaneous Economic Growth", Working-Paper 91-22, Departamento de Economía, Universidad Carlos III de Madrid.

Espasa, A. y J.R. Cancelo, (1992), Métodos cuantitativos para el análisis de la coyuntura económica, Alianza Editorial.

Espasa, A. y D. Peña, (1990), "Los modelos ARIMA, el estado de equilibrio en variables económicas y su estimación", Investigaciones Económicas, v. XIV, 191-211.

Ferreti, N.E. y Romo, J.J., (1992), "Bootstrapping unit root AR(1) models", Working Paper 92-22, Universidad Carlos III de Madrid.

Fisher, R.A., (1922), "The Goodness of Fit of Regression Formulae, and the Distribution of Regression Coefficients", Journal of the Royal Statistical Society, 85, 4, 597-612.

Fisher, R.A., (1925), Statistical Methods for Research Workers, Edimburgo, Oliver y Boyd.

Fisher, F.M., (1965), "The Choice of Instrumental Variables in the Estimation of Economy-Wide Econometric Models", en J. Duesenberry et al. (eds., 1965), capítulo 15.

Florens, J.P. y M. Mouchard, (1985a), "Conditioning in Dynamic Models", Journal of Time Series Analysis, vol. 6, nº 1, págs. 15 a 34.

Florens, J.P. y M. Mouchard, (1985b), "A Linear Theory for Noncausality", Econometrica, 53, 1, 157-175.

Friedman, M. y A. J. Schwartz, (1982), Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices and Interest Rates, 1867-1975, Chicago, University of Chicago Press.

Milton Friedman and Anna J. Schwartz -----, (1991), "Alternative Approaches to Analyzing Economic Data", American Economic Review, marzo, págs, 39-49.

Frish, R., (1926), "Sur un problème d'économie pure", Norsk Matematisk Forenings skrifter.

Frish, R., (1936), "Note on the term "Econometrics"", Econometrica, vol. 4, págs. 95.

Frish, R., (1933), "Editorial", Econometrica, vol. 1, págs. 1 a 4.

Geary, R.C., (1949), "Determination of Linear Relations between Systematic Parts of Variables with Errors of Observations, the Variances of Which are Unknown", Econometrica, vol. 17, 30-58.

Gilbert, C.L., (1986), "Professor Hendry's Econometric Methodology", Oxford Bulletin of Economics and Statistics, 48, 3, 283-307.

Gilbert, C.J., (1989), "LSE and the British Approach to Time Series Econometrics", Oxford Economic Papers, vol. 25, University of Chicago Press, Chicago.

Godfrey, L.G., (1978), "Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables", Econometrica, 46, 6, 1293-1301.

Goldberger, A.S., (1991), A Course in Econometrics, Harvard University Press, Cambridge.

Granger, C.W.J., (1990), Modelling Economic Series, Clarendon Press, Oxford.

Granger, C.W.J. y T. Teräsvirta, (1992), Modelling Dynamic Nonlinear Relationships.

Haavelmo, T., (1943), "The Statistical Implications of a System of Simultaneous Equations", Econometrica, vol. 11, 1-12.

Haavelmo, T., (1944), "The Probability Approach in Econometrics", Econometrica, 12, Supplement, 1-118.

Harvey, A.C., (1981), The Econometric Analysis of Time Series, Oxford, Phillip Allan.

Harvey, A.C., (1989), Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press, Cambridge.

Hastay, M., (1951), review of T.C. Koopmans (ed), "Statistical Inference in Dynamic Economic Models", Journal of the American Statistical Association, 46, 388-90.

Hendry, D.F., (1976), "The Structure of Simultaneous Equation Estimators", Journal of Econometrics, vol. 4, 51-88.

Hendry, D.F., (1983), "Econometric Modeling: The "Consumption Function" in Retrospect", Scottish Journal of Political Economy, vol. 30, 193-406.

Hendry, D.F., (1986), "Econometric Modelling with Cointegrated Variables: an overview", Oxford Bulletin of Economics and Statistics, 48, 3, 201-212.

Hendry, D.F., (1987), "Econometric Methodology: A Personal Perspective", Capítulo 10 en T.F. Bewley (ed.), Advances in econometrics: Fifth World Congress, Cambridge: Cambridge University Press, v. 2, 29-48.

Hendry D.F., (1992), Lectures on Econometric Methodology, Oxford University Press.

Hendry, D.F. y N.R. Ericsson (1991), "An econometric analysis of U.K. money demand in Monetary Trends in the United States and the United Kingdom by Milton Friedman and Anna J. Schwartz, American Economic Review, marzo, págs. 8 a 38.

Hendry, D.F., E.Z. Leamer y D.J. Poirier (1990), "The ET dialogue: a contribution on econometric methodology", Econometric Theory, v. 6, págs. 171 a 261.

Hendry, D.F., A.R. Pagan, y J.D. Sargan, (1984), "Dynamic Specification", Capítulo 18 en Z. Griliches y M.D. Intriligator, (eds), Handbook of Econometrics, Amsterdam: North-holland, v. 2, 1003-1100.

Hendry, D.F. y J.F. Richard, (1982), "On the formulation of Empirical Models in Dynamic Econometrics", Journal of Econometrics, vol. 20, nº 3, 3-33. Recogido posteriormente en el capítulo 14 de Granger (1990).

Hendry, D.F. y J.F. Richard, (1983), "The Econometric Analysis of Economic time Series", (con discusión), International Statistical Review, vol. 51, 111-163.

Hendry, D.F. y J.F. Richard, (1989), "Recent Developments in the Theory of Encompassing", Capítulo 12 en B. Cornet y H. Tulkens (eds.), Contributions to operations Research and Economics: The Twentieth Anniversary of CORE, Cambridge, Massachusetts: MIT Press, 393-440.

Hendry, D.F., A. Spanos y N.R. Ericsson, (1989), "The Contributions to Econometrics in Trygve Haavelmo's", The Probability Approach in Econometrics, social okonomen, 43, 11, 12-17.

Hendry, D.F. y K.F. Wallis, (eds. 1984), Econometrics and Quantitative Economics, Basil Blackwell.

Hood, W.C. y T.C. Koopmans, (1953), (eds.), Studies in Econometrics Method, Cowles Commission Monograph 14, New York: Wiley.

Jarque, C.M. and A.K. Bera, (1980), "Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals", Economic Letters, 6, 3, 255-259.

Jenkins, G.M., (1979), Practical Experiences with modelling and forecasting time series, Gwilym Jenkins and Partners, Lancaster.

Johansen, S., (1988), "Statistical analysis of cointegration vectors", Journal of Economic Dy. and Control, v. 12, págs. 235-254.

Søren Johansen, (1991), "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models", Econometrica, 59, págs. 1551 a 1580.

Johansen, S., (1992), "Cointegration in partial systems", Journal of Econometrics, vol. 52, págs. 389-402.

Johnston, J., (1963), Econometric Methods, New York, McGraw-Hill.

Judge, G.G., W.E. Griffiths, R. Hill, H. Lütkepohl y T. Lee, (1985), The Theory and Practice of Econometrics, John Wiley & Sons, Nueva York.

Kalman, R.E., (1960), "A ner approach to linear filtering and prediction problems", Journal of Basic Engineering, Transactions, ASME. Series D 82, 35-45.

King, R.G., C.I. Plosser, J.H. Stock, M.W., Watson, (1991), "Stochastic Trends and Economic Fluctuations", American Economic Review, vol. 81, 819-840.

Koch, P.D. y R.H. Rasche, (1988), "An Examination of the Commerce Department Leading-Indicator Approach", Journal of Business and Economic Statistics, v. 6, 167-187.

Kolmogorov, A., (1933), Foundations of the Theory of Propability, publicado en 1950 por Chelsea, Nueva York.

Koopmans, T.C., (1947), "Measurement Without Theory", Review of Economic Statistics, 29, 161-72.

Koopmans, T.C., (1949), "Reply to Rutledge Vining", Review of Economics and Statistics, 31, 86-91.

Koopmans, T.C., (1950), (eds.), "Statistical Inference in Dynamic Economic Models", Monografía 10 de la Cowles Comission for Research in Economics, Wiley, Nueva York.

Levy, P., (1937), Théorie de láddition des variables aléatoires, París.

Ljung, G.M. y G.E.P. Box, (1978), "On a measure of lack of fit in time series models", Biometrika, 66, págs. 67-72.

Lucas, R.J., (1976), "Econometric Policy Evaluation: a Critique", en K. Brunner y A.H. Meltzer (eds, 1976), 19-46.

Lütkepohl, H., (1991), Introduction to Multiple Time Series Analysis, Springer-Verlag.

Mackinnon, J.G., (1983), "Model specification test against non-nested alternatives", Econometric Reviews, vol. 2, nº 1, págs. 85-158.

Malinvaud, E., (1966), Statistical Methods of Econometrics, North-Holland.

Mann, H.B. y A. Wald, (1943), "On the Statistical treatment of linear Stochastic Difference Equations", *Econometrica*, v. 11, págs. 173 a 220.

Maravall, A., (1992), "Stochastic Linear Trends: Models and Estimators", mimeo, European University Institute.

Mentz, R.P., E. de Alba, A. Espasa y P.A. Morettin, (eds. 1989), Statistical Methods for Cyclical and Seasonal Analysis, Interamerican Statistical Institute, Panama.

Mizon, G.E., (1984), "The Encompassing Approach in Econometrics", en D. Hendry y K. F. Wallis, (eds. 1984), 135-172.

Mizon, G.E. y J. F. Richard, (1986), "The Encompassing Principle and its Application to Testing Non-Nested Hypothesis", *Econometrica*, vol. 54, 657-678.

Morales, E., J.F. Izquierdo y A. Espasa, (1991), "Análisis coyuntural de los precios al consumo en las comunidades autónomas españolas: aplicación a Castilla-León", Documento de Trabajo 91-08, Universidad Carlos III de Madrid.

Morgan, M.S., (1990), *The History of Econometric Ideas*, Cambridge university Press.

Neyman, J. (1937), Lectures and Conferences on Mathematical Statistics, Washington.

Nicholls, D.F. and A.R. Pagan, (1983), "Heteroscedasticity in Models with Lagged Dependent Variables", *Econometrica*, 51, 4, 1233-1242.

Pagan. A., (1987), "Three econometric methodologies: critical appraised", Journal of Econometric Surveys, n. 1, págs. 2-24.

Pesaran, M.H., (1974), "On the General Problem of Model Selection", *Review of Economic Studies*, 41, 2, 153-171.

Pesaran, M.H., (1982) "Comparision of local power of alternative tests of non-nested regression models", *Econometrica*, septiembre, vol. 50, págs. 1287 a 1305.

Pesara, M.H., (1987) "Econometrics", in *Econometrics, the New Palgrave*, págs. 1 a 34, J. Eatwell, M. Hilgate y P. Newman (eds.)

Phillips, P.C.B., (1988), "Reflections on Econometric Methodology", Economic Record, 64, 187, 344-359.

Phillips, P.C.B., (1991), "Optimal Inference in Cointegrated Systems", *Econometrica*, 59, 2, 283-306.

Quenouille, M.H., (1957), The Analysis of Multiple Time Series, Griffin, Londres.

Ramsey, J.B., (1969), "Tests for Specification Errors in Classical Linear Least-squares Regression Analysis", Journal of the Royal Statistical Society, Series B, 31, 2, 350-.71.

Raymond, J.L., (1992), "Análisis coyuntural y modelos macroeconómicos", documento de trabajo 88/92, Fundación Fondo para la Investigación Económica y Social, Madrid.

Reiersol, O., (1941), "Confluence Analysis by Means of Lag Moments and Other Methods of confluence Analysis", *Econometrica*, v. 9, 1-24.

Reiersol, O., (1945), "Confluence Analysis by Means of Instrumental Sets of Variables", *Arkiv for Matematik, Astronomi och Fysik*, Almquist and Wicksells Boktryckeri-AB, Uppsala, 1-119.

Revilla, P., R. Rey y A. Espasa, (1989), "Characterization of Production in Different Branches of Spanish Industrial Activity, by Means of Time Series Analysis", trabajo presentado al Congreso Europeo de la *Econometric Society*, Munich, Agosto, 1989. Aparecido posteriormente como Working-Paper 91-28, Departamento de Economía, Universidad Carlos III de Madrid.

Robinson, P.M., (1991), "Applications of Semiparametric Modelling in Economics", *Revista Española de Economía*, v. 8, nº 1, págs. 53-60.

Samuelson, P.A., T.C. Koopmans y J.R.N. Stone, (1954), "Report of the evaluative committee for *Econometrica*", *Econometrica*, vol. 22, págs. 141 a 146.

Sargan, J.D., (1958), "The Estimation of Economic Relationships Using Instrumental Variables", *Econometrica*, 26, 3, 393-415.

Sargan, J.D., (1964), "Wages and Prices in the United Kingdom: a study in econometric methodology", en P. Hart, G. Mills y J. Whitaker (eds), *Econometric Analysis for National Economic Planning*, Butterworths, Londres.

Sargan, J.D., (1980), "The Consumer Price Equation in the Post War British Economy: an Exercise in Equation Specification Testing", *Review of Economic Studies*, vol. 47, 113-135.

Sargent, T.J. y C.A. Sims, (1977), "Business Cycle Modelling without Pretending to have too much a priori economic theory", en Sims (1977).

Sims, C.A., (1980), "Macroeconomics and Reality", *Econometrica*, vol. 48, 1-48. Posteriormente publicado como capítulo 8 en Granger (1990).

Sims, C.A., J.H. Stock y M.W. Watson, (1990), "Inference in Linear Time Series Models with some unit roots", *Econometrica*, v. 58, nº 1, 113-144.

Spanos, A., (1986), *Statistical Foundations of Econometric Modelling*, Cambridge University Press, Cambridge.

Spanos, A., (1988), "Error autocorrelation revisited: the AR(1) case", *Econometric Reviews*, 6, 2, págs. 285-294.

Spanos, A., (1989), "On rereading Haavelmo: a retrospective view of econometric modelling", *Econometric Theory*, vol. 5, págs. 405-29.

Spanos, A., (1990), "The simultaneous equations model revisited: statistical adequacy and identification, Journal of Econometrics.

Taylor, S.J., (1986), Modelling Financial Time Series, John Wiley, Chichester.

Tiao, G.C. y G.E.P. Box, (1981), "Modeling multiple time series with application", Journal of the American Statistical Association, vol. 76, 802-816.

Vining, R., (1949), "Koopmans on the Choice of Variables to be Studied and of Methods of Measurements", Review of Economic and Statistics, 31, 77-86.

White, H., (1980), "A Heteroskedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", Econometrica, 48, 4, 817-838.

White, H., (1984), Asymptotic Theory for Econometricians, Academic Press, Inc., Economic Theory, Econometrics, and Mathematical Economics.

Wilks, S.S., (1943), Mathematical Statistics, Princeton.

Wilks, S.S., (1937), Statistical Inference, Princeton.

Wold, H., (1954), "Causality and Econometrics", Econometrica, vol. 22, 162-177.

Wold, H., (1960), "A Generalization of Causal Chain Models", Econometrica, vol. 28, 443-463.

Wold, H., (1964), Econometric Model Building: Essays on the Causal Chain Approach, North-Holland.

Zellner, A., (1962), "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias", Journal of the American Statistical Association, vol. 57, 348-368.