MODELOS Y MÉTODOS ECONOMÉTRICOS DINÁMICOS: APRENDIENDO CON DATOS NO EXPERIMENTALES

Hernán Sabau García

1. EL MÉTODO ECONOMÉTRICO

En las ciencias sociales difícilmente existe el concepto de «experimento»[1]. Normalmente observamos el acontecer de la realidad, y simplemente registramos los sucesos que nos parecen relevantes para entenderla mejor. Pero no existe la posibilidad de observar bajo condiciones controladas que nos permitan distinguir de manera aislada los efectos de tal o cual fenómeno social sobre tal o cual otro. Ésta es una diferencia importante entre las ciencias sociales y las ciencias naturales y que afecta y determina nuestra manera de conocer y de aprender. Entender esto cabalmente puede evitar muchos errores de interpretación y, principalmente, mucha frustración. Creer que la «verdad» puede ser captada de manera objetiva a través de unas cuantas regresiones y pruebas de hipótesis estadísticas es, en el mejor de los casos, ingenuo. Interpretar el reproceso de la información económica como un proceso de conocimiento estadístico experimental, es un grave error.

El Método Econométrico debe entenderse como un proceso de aprendizaje acerca del objeto de estudio, en el que se va incorporando la información disponible, tanto teórica como empírica, para mejorar el conocimiento. Es perfectamente válida la utilización de modelos matemáticos con componentes aleatorios para expresar la información teórica, y el uso de modelos de corte estadístico para incorporar en ellos la información empírica. La formalidad de este procedimien-

1. Es importante subrayar aquí el término «difícilmente», a diferencia de «no» o «nunca», y referir al lector al artículo sobre Economía Experimental en este mismo volumen.

to de incorporar conocimiento teórico y empírico en un mismo modelo permite generar marcos de referencia y de interpretación que son útiles para entender mejor la realidad, para hacer predicción y, prediciendo bajo escenarios alternativos, planear y diseñar políticas.

Es importante destacar que las reflexiones de este artículo están motivadas principalmente por el trabajo con series temporales en Economía. Ha habido desarrollos interesantes en los últimos años en áreas del conocimiento económico en las que es posible contar con bases de datos más amplias, como son las encuestas, y hechas de manera secuencial en el tiempo sobre los mismos individuos o agentes, en la forma de panel, lo que permite una práctica econométrica más apegada a las bases estadísticas teóricas. En buena medida, la problemática de conocimiento econométrico en estas áreas tiene mejores condiciones. Tal es el caso de los estudios de pobreza o, más en general, de ingresos y gastos. Otra área que ha permitido una mejor práctica, al menos parcialmente, es la de las finanzas, en lo que se refiere a mercados, porque la posibilidad de contar con grandes masas de información permite partir los datos en subconjuntos que pueden ser tratados como conjuntos de datos muestrales secuenciales. Sin embargo, las reflexiones aquí vertidas se aplican también en general al trabajo con series financieras en las que los datos no se segmenten adecuadamente en las submuestras necesarias y suficientes.

En la siguiente sección delimitamos el concepto de «realidad» con el que podemos trabajar al construir modelos econométricos, introduciendo el Proceso Generador de Información. En la sección 3 postulamos las bases para construir modelos econométricos conjuntando información teórica y empírica. En la sección 4 proponemos formas válidas de interpretar el objeto de conocimiento y, finalmente, en la sección 5 hablamos de las estrategias a seguir en lo económico y lo estadístico para lograr modelos adecuados.

2. EL PROCESO GENERADOR DE INFORMACIÓN

La metodología econométrica dinámica parte de la base de que los fenómenos por estudiar, o sus características, son representables como variables susceptibles de observación y medición. La variable objeto de estudio la llamamos *variable endógena*, y la conceptualizamos como una variable aleatoria. Una variable aleatoria es una variable cuyo comportamiento es incierto, no sabemos de antemano qué valores va a tomar, pero sabemos que puede tomar valores en diversos rangos con diferentes probabilidades. Como la variable se realiza en el tiempo, es conveniente que le agreguemos un diferenciador del momento del tiempo al que pertenece. No es lo mismo hablar del consumo del mes de diciembre que del consumo del mes de julio. Así, conforme el tiempo va pasando, obtenemos una variable aleatoria distinta para cada momento o periodo. O, mejor dicho, la variable es la misma, pero cambia en el tiempo. A una variable aleatoria que podemos seguir y observar en el tiempo la llamamos un proceso estocástico.

Podemos pensar en el tiempo en tres niveles: el pasado, que ya sucedió, ya se observó y resultó en valores determinados para la variable; el presente, que se está realizando y cuya realización es aleatoria; y el futuro, que está por realizarse y que es por ello también aleatorio. En el mismo momento en que se está realizando el presente, contamos con la información solamente del pasado: el presente se está generando y aún no tenemos información acerca del futuro, aparte de la que podamos inferir de lo que conozcamos hoy: el pasado o la historia. Por eso, es fundamental entender que, para hacer proposiciones, requerimos un conjunto de información, la historia, que es la información que los agentes económicos tienen disponible, hasta cada momento, para tomar decisiones. Cualquier elemento de dicho conjunto de información puede ser considerado como «dado» o conocido, es decir, no sujeto a aleatoriedad, para hacer proposiciones acerca de la variable endógena en el presente o futuro. Al irnos moviendo en el tiempo, el futuro se va haciendo presente y el presente se convierte en pasado y, obviamente, el conjunto de información va creciendo y todo lo que pertenece a él en ese momento ya no es aleatorio: ya es información, ya sucedió. Vamos aprendiendo si vamos capturando y sistematizando dicha información. En palabras simples: la historia de hoy la construimos con lo que sabemos hasta este momento, con la historia.

Al mecanismo que genera la información que se va observando lo llamamos el *Proceso Generador de Información* (*PGI*). El *PGI* acota nuestro concepto de «realidad» a lo que efectivamente podemos observar y medir y lo conceptualizamos, según lo expresado anteriormente, como un proceso estocástico para la variable endógena. Esto quiere decir que es posible hacer proposiciones probabilísticas para la variable endógena dada la información disponible. Tal vez, sin embargo, nos incomode pensar en una función de probabilidad, que podría ser una relación matemática compleja, para hacer estas proposiciones. La ventaja es que podemos caracterizar la función de probabilidad a partir de conceptos que nos son más familiares, como es la *expectativa* o *valor esperado* de la variable, también conocido como la *media*, que indica el valor alrededor del cual se concentra la distribución. El valor que tomará nuestra variable endógena hoy lo podemos descomponer en dos partes excluyentes: lo que esperamos y lo que no esperamos. Es decir,

$$Y_t = \mu_t + \varepsilon_t, \qquad \text{-(1)}$$

que descompone la variable endógena Y_t en un componente sistemático μ_t, lo esperado, y uno no sistemático ε_t, lo inesperado. Clamar la existencia del *PGI* no parece muy restrictiva a la luz de (1), ya que permite como casos particulares la aleatoriedad total (no podemos construir una expectativa acerca de la variable endógena) cuando $Y_t = \varepsilon_t$, así como la sistematicidad total cuando $Y_t = \mu_t$ y tenemos conocimiento perfecto de la variable en estudio.

Por lo general, no basta tener un indicador de localización de la distribución como la media. La *varianza* σ_t^2 de la distribución (o equivalentemente su *desviación estándar* σ_t) ofrece un indicador de la *dispersión* de la distribución alrededor de la media que permite hacer proposiciones probabilísticas para la variable dada la información disponible en cada momento[2]. Cuanto mayor sea la varianza (o desviación estándar), más dispersa o menos concentrada será la distribución.

Muchas de las distribuciones conocidas y bien comportadas se pueden caracterizar completamente con la media y la varianza, como es el caso de la distribución normal y, por el momento, no vale la pena que compliquemos el argumento tratando distribuciones más complejas. Supondremos, por tanto, que nos basta conocer la media y la desviación estándar para hacer proposiciones de probabilidad acerca de una variable. Estaremos en condiciones de hacer cualquier proposición probabilística acerca de la variable en estudio condicional sobre la información disponible si conocemos su distribución de probabilidad o, alternativamente, su media y varianza. Si conocemos μ_t y σ_t^2 podemos decir, por ejemplo, que la probabilidad de que nuestra variable endógena esté entre $\mu_t - 2\sigma_t$ y $\mu_t + 2\sigma_t$ será de aproximadamente 95%. Más específico: una variable aleatoria normal con media de 100 y desviación estándar de 10, sabemos que tomará valores entre 80(=100-2*10) y 120(=100+2*10) con aproximadamente 95% de probabilidad. El chiste es poder conocer esta media y esta varianza para hacer las proposiciones. Para esto nos sirve el Método Econométrico.

2. Formalmente, $\sigma_t^2 = E\left[\left(Y_t - \mu_t\right)^2\right]$, que denota que la varianza es una medida promedio ponderado de la distancia de cada uno de los posibles valores de la variable alrededor de su media, con ponderaciones iguales a las correspondientes probabilidades. El uso del cuadrado de estas desviaciones evita que, como medida de dispersión, las desviaciones positivas se compensen con las negativas.

3. EL MODELO ECONOMÉTRICO

El Método Econométrico se concentrará en construir un modelo para el *PGI* que sea operativo, en el sentido de hacer proposiciones concretas y cuantitativas, con carácter probabilístico, para la variable endógena. El punto de partida es la *Teoría Económica*, que sugiere una explicación para la variable endógena en función de otras variables. Las variables explicativas serán llamadas *variables exógenas*[3] X_t, también distinguidas en el momento t. Esto nos permite proponer un modelo teórico para el *PGI*

$$Y_t = X_t \mathrm{M} + u_t, \qquad \text{-(2)}$$

que denota que el valor esperado de la variable endógena depende de las variables exógenas, y donde M es un *parámetro*, que mide la intensidad de dicha relación: cuánto cambia el valor esperado de la variable endógena al cambiar la exógena. Esto reconoce que el modelo econométrico es un *modelo condicional* sobre un conjunto de información que incluye, por lo menos, a las variables exógenas, y también reconoce que se trata de un *modelo paramétrico*. El parámetro M es una constante que determina la media de la distribución. Este parámetro es habitualmente desconocido y usualmente la teoría sugiere si es positivo o negativo, generando relaciones directas o inversas entre las variables exógena y endógena. Nuestro modelo descompone la variable explicada en un *componente determinístico* $X_t M$, determinado o explicado por la teoría, y un *componente no determinístico* u_t, que no es explicado por la misma y cumple con que $E[u_t]=0$, es decir, esperamos que las desviaciones alrededor de lo que explicamos con la teoría y el razonamiento sean tanto hacia arriba como hacia abajo y, en promedio, se anulen. Muy pobre sería una explicación que admita desviaciones permanentes positivas o negativas[4].

El primer supuesto que se incorpora en (2) es el de *correcta especificación*, en el sentido de que la selección de variables exógenas es el

3. Trataremos a las variables exógenas como si se tratase de una sola; pues el trabajar con un vector de variables $X_t = (X_{t1}, X_{t2}, \ldots, X_{tN})$, que incluye N posibles componentes, como debiera ser el caso, no aporta cualitativamente mayor cosa a los argumentos y sí los hace más complejos. Por ello, aunque hablemos de «variables exógenas» y no de «variable exógena», matemáticamente lo haremos con una sola variable.

4. Una sutileza es que pueden existir desviaciones con patrones sistemáticos en el corto plazo alrededor de una explicación teórica adecuada, pero el requerimiento de valor esperado cero implica que dichas sistematicidades desaparecen en el largo plazo. Una sutileza, pero importante.

adecuado para explicar la variable endógena. Un segundo supuesto implícito en (2) es el de *exogeneidad*, que se refiere a la validez del condicionamiento de Y_t sobre X_t, es decir, establece que es propio suponer que los agentes conocen X_t al momento de decidir sobre o determinar Y_t[5].

Ahora bien, el modelo en (2) condiciona sobre el conjunto de información de las variables exógenas que sugiere la Teoría Económica, exclusivamente. La mayor parte de las veces la Teoría Económica funciona así: nos abstraemos del momento en el tiempo para pensar qué es lo sustantivo que lleva a los agentes económicos a actuar de tal o cual forma, en condiciones ideales o *ceteris paribus*. La realidad es que, cuando nos situamos en un momento en el tiempo, la información disponible para que los agentes tomen decisiones no incorpora solamente a las variables (teóricas explicativas) de ese momento, sino que incorpora además toda la información del pasado, sobre esas variables y cualesquiera otras, incluyendo la explicada: la historia. Por este motivo, para construir un modelo realista que tenga sentido en el tiempo, el conjunto de información disponible al momento debe agregar a la información del presente toda la que proviene del pasado, tanto de las variables exógenas como de la endógena. Esto replantea la descomposición en (2) como

$$Y_t = \mu_t\left(Y_{t-1}, X_t, X_{t-1}; \beta\right) + \varepsilon_t = \alpha Y_{t-1} + \gamma_0 X_t + \gamma_1 X_{t-1} + \varepsilon_t, \qquad \text{-(3)}$$

Donde reconocemos un componente *sistemático* $\mu_t\left(Y_{t-1}, X_t, X_{t-1}; \beta\right) = \alpha Y_{t-1} + \gamma_0 X_t + \gamma_1 X_{t-1}$[6] y uno *no sistemático* ε_t, y donde $E\left[\varepsilon_t\right] = 0$ establece la no sistematicidad de este último pues su valor esperado no depende en absoluto de toda la información dis-

5. Un tercer supuesto implícito es el de *distintividad* de las variables, es decir, que las variables exógenas realmente son variables que miden aspectos distintos de la realidad y no son redundantes entre sí. Si es posible obtener alguno(s) de los componentes de $X_t = \left(X_{t1}, X_{t2}, \ldots, X_{tN}\right)$ a partir de los restantes componentes, entonces dichos elementos son redundantes. Por ejemplo, si $X_{t1} = f\left(X_{t2}, \ldots, X_{tN}\right)$, entonces X_{t1} es redundante: no agrega nada a las demás variables. Al observar e incorporar información empírica, este supuesto toma el nombre genérico de *no multicolinealidad*. Al trabajar con una sola variable exógena este supuesto se cumple trivialmente.

6. Por simplicidad y sin pérdida de generalidad, hemos incluido la dependencia del pasado solamente en el momento inmediato anterior, tanto de la variable endógena como de la exógena y del *shock* externo no sistemático. Esto es suficiente para analizar las propiedades cualitativas del sistema y poner más rezagos simplemente hace más complejo el argumento sin aportar mucho. Que tantos rezagos de las variables puedan afectar el presente es, en términos generales, una materia empírica que se resuelve en la estimación del modelo.

ponible hasta el momento. Ahora tenemos varios parámetros que determinan el componente sistemático, y los agrupamos en el vector $\beta = (\alpha, \gamma_0, \gamma_1)$. Nuestro modelo, para el componente sistemático, podría tomar diversas formas funcionales y bastaría, en general, con que fuese continuo. Sin embargo, en ausencia de información acerca del *PGI* que nos lleve a pensar que la relación es más compleja, el punto de partida natural es suponer *linealidad* en los parámetros, por ser la relación más simple posible.

La diferencia entre el modelo teórico en (2) y el modelo econométrico en (3) es el conjunto de información sobre el que se condiciona. En el modelo teórico sólo se condiciona sobre las variables exógenas y, por ese motivo, se le suele llamar la *versión estática o de largo plazo* del modelo, mientras que (3) es un modelo teórico-empírico que, además de las relaciones económicas racionales, incorpora la dinámica que implica el que las variables toman dimensión en el tiempo, en el que se generan inercias, y por ello se le llama la *versión dinámica*, y corresponde a lo que realmente observamos. El comportamiento de largo plazo, que sin duda está presente en el día a día, obedece al comportamiento racional de los agentes, mientras que en el corto plazo este comportamiento está afectado por la inercia que introducen la costumbre y las instituciones. El modelo de corto plazo incorpora más información que el de largo plazo y, por tanto, lo contiene. Podemos, de hecho, relacionar los dos modelos a través de la descomposición

$$Y_t = \mu\left(Y_{t-1}, X_t, X_{t-1}; \beta\right) + \varepsilon_t = f\left(X_t; M\right) + g\left(Y_{t-1}, X_{t-1}; \beta\right) + \varepsilon_t = f\left(X_t; M\right) + u_t,$$

donde el componente sistemático se puede descomponer en un componente (sistemático) determinístico $f\left(X_t; M\right)$ (de largo plazo) y un componente (sistemático) no determinístico $g\left(Y_{t-1}, X_{t-1}; \beta\right)$ (puramente friccional o de corto plazo), es decir,

$$\mu\left(Y_{t-1}, X_t, X_{t-1}; \beta\right) = f\left(X_t; M\right) + g\left(Y_{t-1}, X_{t-1}; \beta\right),$$

o bien, el componente no determinístico u_t se puede descomponer en un componente sistemático (no determinístico) $g\left(Y_{t-1}, X_{t-1}; \beta\right)$ y uno no sistemático ε_t, es decir,

$$u_t = g\left(Y_{t-1}, X_{t-1}; \beta\right) + \varepsilon_t.$$

El modelo econométrico es, entonces, un *modelo paramétrico condicional* que propone formas teórico-empíricas para la descomposición del *PGI* dada en (1). Esta propuesta incluye como extremos a la ignorancia total, cuando $\mu\left(Y_{t-1}, X_t, X_{t-1}; \beta\right) = 0$ o $Y_t = \varepsilon_t$, y al conocimiento perfecto, cuando podemos hacer proposiciones exactas para la variable en estudio y $Y_t = \mu\left(Y_{t-1}, X_t, X_{t-1}; \beta\right)$ y $\varepsilon_t = 0$, y donde el vector de parámetros es conocido. Tenemos también el caso intermedio en que somos incapaces de explicar la variable en estudio teóricamente, la teoría es inútil, pero los datos sí proporcionan un modelo para pronóstico, es decir, si $\mu\left(Y_{t-1}, X_t, X_{t-1}; \beta\right) = g\left(Y_{t-1}, X_{t-1}; \beta\right)$ o bien $f\left(X_t; M\right) = 0$. En tal caso el modelo contiene sólo conocimiento empírico y sirve para pronosticar, pero es poco útil desde una perspectiva económica o política (con la posible excepción de la evacuación: sálvese quien pueda).

Las descomposiciones anteriores muestran un vínculo trivial entre los modelos teórico y econométrico en (2) y (3). El vínculo no trivial entre ambos se da cuando pensamos bajo qué condiciones el modelo dinámico en (3) alcanza un estado de equilibrio. Para que esto sea posible, sería indispensable que los estímulos externos se mantengan constantes, esto es, que $X_t = X_{t-1} = X_{t-2} = \ldots$ Entonces también podría la variable endógena alcanzar un estado de equilibrio $Y_t = Y_{t-1} = Y_{t-2} = \ldots$ sólo perturbado por el componente aleatorio que recoge las desviaciones inesperadas en el tiempo, reduciendo (3) a

$$Y_t = \alpha Y_t + \gamma_0 X_t + \gamma_1 X_t + \varepsilon_t = \alpha Y_t + \left(\gamma_0 + \gamma_1\right) X_t + \varepsilon_t,$$

o bien, resolviendo para Y_t, bajo la condición de que $\alpha < 1$,

$$Y_t = X_t \mathrm{M} + u_t,$$

donde el multiplicador o parámetro de largo plazo M está dado ahora por

$$\mathrm{M} = \frac{\gamma_0 + \gamma_1}{1 - \alpha},$$

que lo relaciona con los parámetros de corto plazo. La condición $\alpha < 1$ es necesaria para que exista la relación de equilibrio en (2), que debe interpretarse como un «estado» de equilibrio, puesto que es probabilística dada la presencia del componente u_t que siempre, aún en equilibrio, provocará desviaciones aleatorias alrededor del nivel esperado de equilibrio X_tM. Esto requerirá que $E\left[u_t\right] = 0$ para que no haya desviaciones permanentes de la variable endógena con respecto a su media teórica[7].

7. Esto no permite que las variables endógenas y exógenas presenten comportamientos permanentemente desiguales. Engle y Granger llamaron a esto co-integración,

Hasta ahora sólo hemos hablado de *y* parametrizando la media. Esto no es suficiente, en general, para que hagamos proposiciones probabilísticas: falta la varianza. Por regla general y también por sencillez, y en ausencia de conocimiento explícito en contrario, se suele partir de suponer que la varianza es constante, a lo que se le denomina el supuesto de *homoscedasticidad*[8]. Entonces

$$Var\left[y_t\right]=E\left[\left(y_t-\mu_t\right)^2\right]=E\left[\varepsilon_t^2\right]=\sigma^2.$$

La parametrización del modelo incluye ahora a los parámetros tanto de la media como a la varianza, es decir, el vector de parámetros del modelo es $\theta=\left(\beta,\sigma^2\right)=\left(\alpha,\gamma_0,\gamma_1,\sigma^2\right)$. Ya aclaramos anteriormente que muchas de las distribuciones conocidas y bien comportadas dependen solamente de estos dos primeros momentos y, una vez estimados éstos, todos los momentos de orden superior se estiman en función de ellos, por lo que basta con estimar los parámetros de media y varianza para tener la imagen completa de la distribución. Tal es el caso de la distribución normal. No es particularmente complejo, sin embargo, generalizar el planteamiento a momentos de orden más alto si esto se requiere (cf. Sabau, 1988).

Concluyendo: si conocemos la forma de la distribución y los parámetros, entonces, dada la información disponible, podríamos hacer cualquier tipo de proposición probabilística condicional con respecto a la variable endógena. Supongamos, por ejemplo, que sabemos que el ingreso de un cierto grupo económico en el tiempo está dado por un *PGI* como (3),

$$Y_t=0.6Y_{t-1}+0.7X_t+0.4X_{t-1}+\varepsilon_t,$$

donde Y_t es el ingreso del grupo económico y X_t es el monto de lo que invierte, ambos en el periodo t. Esto es, conocemos los parámetros $\alpha=0.6$, $\gamma_0=0.7$ y $\gamma_1=0.4$, en términos de la notación en (3). Suponemos, además, que el proceso es homoscedástico, con varianza $\sigma^2=\$^2 250{,}000$, o bien desviación estándar de $\sigma=\$500$. La información disponible hasta el momento es que el ingreso del periodo anterior fue de \$98,339 (Y_{t-1}), y que en el mismo periodo se invirtieron

y es la condición necesaria y suficiente para que el sistema pueda representar un estado de equilibrio.

8. Realmente no es complejo proponer que la varianza, al igual que la media, dependa de las variables exógenas, lo que podría provenir de algún planteamiento teórico, y también que contenga componentes dinámicos. Entonces se dice que el proceso es heteroscedástico.

\$25,000 (X_{t-1}). También sabemos que el ingreso que se esperaba el periodo pasado era de \$97,791. El grupo económico quiere hacer propuestas para su ingreso para distintas alternativas de planes de inversión (X_t). Si invierten \$25,000, como en el periodo anterior, el ingreso esperado será de \$86,503[9], con lo que su ingreso esperado caería sustancialmente por debajo del de este periodo y con 95% de probabilidad el ingreso estará entre \$85,503 y \$87,503, sumando y restando 2 desviaciones estándar del componente no sistemático. Si incrementan su inversión para evitar esta caída hasta \$40,000, el ingreso esperado será de \$97,003[10], con un intervalo de 95% de \$96,003 a \$98,003, aún parcialmente por debajo del ingreso de este periodo. Necesitarán invertir por encima de los \$42,000 este periodo para mantener su nivel de ingreso esperado en el corto plazo. Para el largo plazo, el parámetro relevante es el multiplicador dado por M=(0.7+0.4)/(1-0.6)=2.75. Con esto, manteniendo un ritmo de inversión como el actual de \$25,000, el grupo vería caer su ingreso esperado en el medio plazo hasta \$68,750 (=2.75*\$25,000), mientras que incrementar y mantener la inversión a ritmos de \$42,000 llevaría el ingreso esperado hasta niveles de \$115,500 (=2.75*\$42,000).

Esto proporciona elementos al grupo económico para hacer sus decisiones en materia de inversión, tan completas como se pueden hacer dado que se dispone de una ecuación de comportamiento que contiene un componente no explicado: lo no sistemático. Según lo importante que sea este componente, se podrán hacer unas proposiciones que resulten confiables. Esto usualmente se mide con la «correlación múltiple» o su equivalente, la «determinación múltiple», la famosa R^2, que dice qué proporción de la varianza de la variable endógena es explicada por las variables exógenas en conjunto. Lo que debemos reconocer aquí es que conocer la función de densidad de probabilidad de la variable endógena condicional sobre la información disponible (el pasado más el escenario para las variables exógenas para el presente y futuro), es lo que permite hacer las proposiciones probabilísticas. En este caso supusimos normalidad, que le da forma a la distribución y nos dice que $\mu \pm 2\sigma$ define un intervalo de aproximadamente 95%. No es necesario que la distribución sea normal, pero necesitaríamos conocer su forma para hacer los planteamientos probabilísticos equivalentes. Si contamos con series relativamente largas (alrededor de 100 observaciones al menos), suponer normalidad puede no ser muy restrictivo aun cuando el supuesto no se cumpla estrictamente. Ade-

9. \$86,503=0.6*\$98,339+0.7*\$25,000+0.4*\$25,000.
10. \$97,003=0.6*\$98,339+0.7*\$40,000+0.4*\$25,000.

más de conocer la función de probabilidad (normal), también estamos teniendo por conocidos todos los parámetros en el vector de parámetros θ, cuando utilizamos los valores fijos $\alpha = 0.6$, $\gamma_0 = 0.7$, $\gamma_1 = 0.4$ y $\sigma = 500$. Con esto pudimos obtener la media y la desviación estándar necesarias para hacer las proposiciones con la distribución normal.

El problema es que prácticamente nunca conocemos el PGI como lo supusimos en el caso del grupo. Nuestro mejor acercamiento es proponer un modelo como (2)-(3), del que raramente conocemos el valor de los parámetros al nivel que es requerido para poder hacer proposiciones de probabilidad completas como en el párrafo anterior. En términos generales, la teoría económica nos permite acotar un poco el espacio de parámetros, pero no tenemos valores específicos para ellos y, por tanto, este conocimiento resulta incompleto para hacer las proposiciones probabilísticas con respecto al objeto de estudio. Es aquí donde entramos en el terreno de la inferencia estadística, en el que necesitamos incorporar la información empírica que nos permita cuantificar los parámetros para hacer proposiciones probabilísticas. Este proceso consistirá en utilizar datos observados acerca de las variables endógenas y exógenas y buscar en la mejor forma posible «estimar» los valores de los parámetros.

La información empírica no solamente servirá para cuantificar parámetros, sino que también, si hacemos parametrizaciones inteligentes y creativas, nos permitirá evaluar y contrastar aspectos más generales, como la forma de la distribución condicional y la coherencia de la teoría con lo que se está observando en la «realidad» representada por el *PGI*. El planteamiento en (2)-(3) es lo suficientemente general como para que la parametrización nos permita explorar la forma de la distribución y las principales hipótesis derivadas de la teoría económica. En general, los supuestos que podamos expresar en forma paramétrica (valores específicos para algún parámetro o función de parámetros), en nuestros modelos, serán sujetos de contraste empírico.

La parametrización en (3) es central para las hipótesis que se deseen evaluar. Es imprescindible tomar en cuenta que el concepto de «realidad» queda acotado a la propuesta de parametrización para el *PGI* en el modelo para explicar la variable endógena. Esto es, nuestra capacidad de distinguir la realidad está restringido por la amplitud de nuestro modelo. Todos aquellos aspectos de la «realidad» no incluidos no podrán ser evaluados ni contrastados: en el mejor de los casos el encontrar que se rechaza una hipótesis es una evidencia de incoherencia entre la realidad y la «realidad» representada por el modelo. La conceptualización del *PGI* implícita en la parametrización es nuestra forma de poner la información teórica, proveniente del estado de conocimiento acumulado con respecto al objeto de estudio, en el modelo. El concepto es que es este *PGI* el que está generando información empírica en la forma de datos para poder evaluar y contrastar hipótesis.

El siguiente paso es atender a la información empírica: los datos. Los datos nos permitirán estimar los parámetros, contrastar y evaluar las hipótesis de interés, y pronosticar y analizar escenarios. Para poder construir un modelo econométrico se necesitan observaciones sobre las variables, esto es, $Y=\left(Y_1,Y_2,\ldots,Y_T\right)'$ para la variable endógena, y $X=\left(X_1,X_2,\ldots,X_T\right)'$ para las variables exógenas, donde T es el tamaño de la muestra. Siguiendo el trabajo estadístico de Fisher y sus seguidores, toda la información en los datos Y, X está contenida en la *función de verosimilitud*, dada por

$$\mathscr{L}\left(\theta|Y,X\right)\propto D\left(Y,X;\theta\right)=D\left(Y_T;\theta\right)D\left(Y_{T-1};\theta\right)\ldots D\left(Y_1;\theta\right), \qquad \text{-(4)}$$

donde $D\left(Y_t;\theta\right)$ es la distribución de probabilidad (*e.g.* la distribución normal) para hacer proposiciones acerca de la variable Y_t y el símbolo $\propto$ denota «proporcional a». Aquí es importante hacer varias aclaraciones. Así como $D\left(Y_t;\theta\right)$ con θ conocido es suficiente (todo lo que necesitamos) para hacer proposiciones acerca de Y_t, podemos pensar en el reverso: $D\left(Y_t;\theta\right)$ con Y_t dado (ya observado y traducido en un valor) es suficiente para hacer proposiciones acerca del vector de parámetros θ. La proporcionalidad denotada por $\propto$ solamente introduce una constante multiplicativa y es necesaria formalmente para poder tratar a la función de verosimilitud como si fuese una función de densidad de probabilidad para los parámetros. El concepto de suficiencia aquí es un concepto técnico estadístico, que nos dice que TODA la información que la observación sobre Y_t contiene acerca de θ está en $D\left(Y_t;\theta\right)$: no hay que buscar en otro lado. Por ejemplo, si tenemos una observación que sabemos proviene de una variable con distribución normal, digamos con el valor 7.8, es evidente que esta observación nos sugiere el valor para la media de 7.8. Pensar cualquier otra cosa sería absurdo. Con una sola observación, sin embargo, no podemos derivar información para la desviación estándar: no hemos visto variar la variable. Si obtenemos una segunda observación independiente de la primera, digamos 9.2, ¿dónde colocaríamos ahora nuestra estimación de la media? Como la normal es simétrica y se concentra alrededor de su media (que por ello es además su moda y mediana), esperaríamos que las observaciones se distribuyan alrededor de dicha media de forma igual: esto nos lleva a colocar nuestra nueva estimación de la media justo en medio de las dos observaciones, en su promedio, es decir, 8.5. Podemos derivar ahora un primer estimador para la desviación estándar, que ha de entenderse como la distancia a ambos puntos desde la media, 0.7, de manera que busquemos que, como en la distribución normal, el 67% de las observaciones estén en un rango de 2 desviaciones estándar, una a cada lado de la media; 95% de las observaciones estén en un rango de 4 desviaciones estándar, 2 a cada lado de la media; y 99% de las observaciones estén en un rango de 6 desviaciones estándar, 3 a cada lado de la media. Claro que con apenas 2 observaciones no podemos hacer cumplir estas tres condiciones pues solamente tenemos todo (100%), nada (0%) o la mitad (50%). Pero en la medida que vamos teniendo 3, 4, 5 y más observaciones, entonces podemos ir conformando mejor estas condiciones y ello nos llevaría a ir juntando la función de verosimilitud de Fisher en (4), que contiene TODA la información que tiene la muestra (del tamaño que sea) con respecto a los parámetros. En el caso de la normal, si siguiéramos con el argumento que íbamos desarrollando, llegaríamos a que la distribución más verosímil, de haber generado las observaciones que hemos visto, tendría su media en el promedio de las observaciones y su varianza en el promedio de los cuadrados de las desviaciones de las observaciones alrededor del promedio.

Siguiendo el argumento, la función de verosimilitud es la función de los parámetros, dados los datos, que toma la forma matemática de la distribución conjunta de los datos dados los parámetros (salvo por una constante multiplicativa que resulta irrelevante, de ahí el uso del símbolo de proporcionalidad $\propto$). Esto quiere decir que podemos utilizar, para cada uno de los parámetros, la función de verosimilitud para hacer proposiciones de confianza, exactamente de la misma manera que hacemos proposiciones probabilísticas para una variable aleatoria cuando conocemos su distribución de probabilidad. La diferencia entre las proposiciones de confianza y las proposiciones probabilísticas es que estas últimas las hacemos para variables aleatorias, como Y_t, mientras que las proposiciones acerca de parámetros que no son aleatorios, no son sino constantes DESCONOCIDAS. Resaltamos lo desconocido de los parámetros porque nos hace tratar a nuestro conocimiento acerca de ellos como si fuese una variable aleatoria y hacer proposiciones probabilísticas que numéricamente se construyen igual que las proposiciones probabilísticas referentes a variables aleatorias, pero que al referirse a parámetros las llamamos proposiciones de confianza, reflejando que estamos usando nuestro conocimiento muestral a través de la verosimilitud para expresar nuestro conocimiento acerca de ellos. Una ventaja importante de la función de verosimilitud es que, aun cuando las variables involucradas y observadas no provengan de la distribución normal, la función de verosimilitud para los parámetros se va pareciendo cada vez más a la normal y si tenemos suficientes observaciones, podemos utilizar la distribución normal para hacer nuestras proposiciones de confianza. Nos basta, entonces, conocer un *estimador* $\hat{\theta}$ (el valor numérico derivado para estimar el parámetro, como el promedio para la media) y su *error estándar* $EE(\hat{\theta})$ (el indicador de su dispersión en función de los datos) para hacer proposiciones de confianza para cada parámetro ¿en lo? individual. Las propuestas de confianza tomarán simplemente la forma $\hat{\theta} \pm c\ EE(\hat{\theta})$, donde la constante c dependerá del tamaño del intervalo que queramos: 1 para confianza de 2/3 (67%), 2 para confianza de 95%, 3 para confianza de 99%, etc. Si queremos hacer propuestas conjuntas, solamente requerimos las covarianzas adicionalmente.

Para la construcción de la densidad condicional en (4), es necesario incorporar un supuesto de *permanencia estructural*, que implica la misma estructura parametral para todos los periodos: el vector $\hat{\theta}$ es el mismo a través de la muestra. Si esto no se diera, cada observación nos hablaría de distintos parámetros (los de su periodo) y no obtendríamos más información al tener más observaciones: observar sería desinformativo porque traería nuevos parámetros que conocer con cada observación. Y también es necesario introducir ¿conocimiento? con respecto a la relación probabilística entre observaciones, a lo que se suele llamar *no autocorrelación*, que requiere que el condicionamiento de los momentos condicionales esté reconociendo la estructura dinámica del fenómeno, de forma tal que ε_t, efectivamente, sea puramente aleatorio o no sistemático. La estructura de condicionalidad en el tiempo sobre el pasado genera de manera natural que la distribución conjunta de todas las observaciones en (4) (la «muestra»), se obtenga multiplicando secuencialmente las distribuciones de cada observación[11]. Permanencia estructural es un supuesto de estabilidad en el tiempo, y no autocorrelación es un supuesto de que la dinámica incorporada en el modelo es lo suficientemente rica como para representar adecuadamente al *PGI*.

En una ciencia experimental, los supuestos que hemos mencionado se incorporan al experimento que generará los datos, el *PGI*, y se observa la realidad para estimar parámetros y probar hipótesis. Al incorporarse hipótesis al estado de conocimiento y generarse nuevas hipótesis se define un nuevo experimento, se vuelve a observar y con-

11. Esta afirmación sigue de una aplicación secuencial de la relación entre la probabilidad conjunta y la probabilidad condicional $P(A,B) = P(A|B)P(B)$, conforme la información se va haciendo disponible en el tiempo y se va convirtiendo en dada: $D(Y_t, Y_{t-1}) = D(Y_t|Y_{t-1})D(Y_{t-1})$.

trastar, y así sucesivamente hasta lograr un conocimiento adecuado. Lo que resulta un poco incómodo en Economía con respecto al enfoque clásico de la inferencia es que la posibilidad de muestrear en forma repetida es prácticamente imposible, y con series temporales es imposible: ¡no hay manera de hacer regresar el tiempo para volver a observar el comportamiento de las variables en el mismo periodo! Con esto, trabajar con estimadores y estadísticos cuyas propiedades e interpretación dependen de esta posibilidad resulta un tanto absurdo. Tener valores estimados (numéricos) y estarlos interpretando como estimadores (variables aleatorias) y entonces volver a hacer inferencias tales como pruebas de hipótesis, re-estimaciones puntuales y regionales, etc., se puede convertir en un ejercicio difícil de comprender y que trata de darle una apariencia científica a un proceso de pesca con los datos hasta acomodarlos a cualquier verdad que uno quiera defender. La práctica puede llegar a ser aplastante: para contrastar las nuevas hipótesis se tiene que utilizar la misma evidencia sobre las que esas mismas hipótesis se han construido. La honestidad del investigador juega un papel crucial en la veracidad de los supuestos y, lamentablemente, no siempre conoce uno a los investigadores que reportan sus resultados en las diversas publicaciones profesionales como para avalar su integridad. Una posible excepción, sin ser experimental, es la de los mercados financieros en los que es posible generar información con alta frecuencia, que puede entonces ser segmentada en submuestras que se utilicen secuencialmente.

4. EL PROCESO DE CONOCIMIENTO

Es indispensable reconocer que, en vez de aprender de nuevos conjuntos de datos, se aprende de nuevos conjuntos de reflexiones que representan, también, nueva información que se incorpora al modelo pero que, al no ser observables objetivas, deben explicitarse en cada paso en que se avanza en la construcción del modelo. Poner la inferencia econométrica en un contexto más modesto, en la que no se renuncia a la búsqueda de la verdad, pero se interpretan de forma más humilde los resultados obtenidos de la conjunción de teoría y datos, parece un principio difícil de cuestionar. Para ello, trataremos de establecer nuestra metodología como un proceso de aprendizaje, en el que vamos aprendiendo y obteniendo nueva información, y estos elementos se van agregando al estado del conocimiento que van representando nuestros modelos. Desde luego, el incorporar el conocimiento bien establecido ya por la profesión es un principio válido para dar forma a la parametrización del *PGI* en el inicio de la investigación, y se incorporan también los supuestos de (o) existencia de una relación de equilibrio (cointegración), (i) correcta especificación, (ii) exogeneidad, (iii) no multicolinealidad, (iv) linealidad, (v) homoscedasticidad, (vi) normalidad, (vii) homogeneidad o permanencia estructural y (viii) no autocorrelación o independencia estocástica.

La inferencia bayesiana y la aplicación de su enfoque de conocimiento a la Economía pueden ayudarnos a plantear el proceso de aprendizaje. En primer lugar, esta escuela no se sustrae de la especificación del modelo para el *PGI* en (2)-(3) para darle forma a la hipótesis mantenida. Tampoco se sustrae del planteamiento de que la función de verosimilitud en (4) contiene toda la información que los datos pueden ofrecer acerca de los parámetros de la hipótesis propuesta y mantenida por el *PGI*. Lo que la escuela bayesiana hace de manera convincente, es el planteamiento de cómo ir integrando fuentes de información en un proceso de aprendizaje continuo. Para lograr esto, simplemente se utiliza en forma reiterada la Regla de Bayes, también llamada de la probabilidad inversa,

$$\Pr\left(A \middle| B\right) = \frac{\Pr\left(A, B\right)}{\Pr(B)} = \frac{\Pr\left(B \middle| A\right)\Pr\left(A\right)}{\Pr(B)} \propto \Pr\left(A\right)\Pr\left(B \middle| A\right),$$

haciendo reconocimiento de que, si el evento *B* está dado, entonces Pr(*B*) es una constante con respecto a *A*. Nótese que, por el contrario, $\Pr\left(B \middle| A\right)$ depende de A y por ello sigue conteniendo elementos relevantes para obtener $\Pr\left(A \middle| B\right)$, otros que una constante.

Adicionalmente, y de vital importancia al argumento, la escuela bayesiana reconoce el papel de datos y parámetros en cuanto a la aleatoriedad desde el punto de vista del conocimiento, de tal manera que los primeros son constantes dadas (como los datos realmente son) y los parámetros son interpretados como «el conocimiento acerca de los parámetros» más que los parámetros mismos, y este conocimiento es claramente aleatorio por estar sujeto a incertidumbre, en tanto no se conozcan perfectamente. Obviamente, cuando ya se conoce el valor de algún parámetro o parámetros, ya no hay que hacer inferencia sobre ellos: el proceso de aprendizaje concluyó exitosamente.

Nuestro interés es plantear el estado de conocimiento acerca de los parámetros θ dada la información obtenida hasta el momento, es decir, $D\left(\theta \middle| Y, X\right)$, lo que utiliza el argumento de la aleatoriedad de θ y la constancia de los datos Y, X. Si para ello aplicamos la Regla de Bayes tenemos que

$$D\left(\theta\middle|Y,X\right)=\frac{D\left(\theta,Y\middle|X\right)}{D\left(Y\middle|X\right)}=\frac{D(\theta)D\left(Y\middle|X;\theta\right)}{D\left(Y\middle|X\right)}\propto D(\theta)\,\mathscr{L}\left(\theta\middle|Y,X\right) \quad \text{-(5)}$$

donde la condicionalidad de la variable endógena *Y* sobre las variables exógenas *X* se reconoce en todo momento, y en la Regla de Bayes se condiciona también sobre la información empírica acerca de la variable endógena misma en el vector *Y*. Se utiliza además que $\mathscr{L}\left(\theta\middle|Y,X\right)\propto D\left(Y\middle|X;\theta\right)$, y se utiliza el conocimiento previo sobre θ contenido en la distribución *a priori* $D(\theta)$. Esta distribución *a priori* puede estar representando el estado de conocimiento de otros investigadores previo a iniciar la investigación en curso, o el conocimiento adquirido por el propio investigador de otras fuentes, incluyendo la posibilidad de la ignorancia. La literatura en la escuela bayesiana acerca de cómo plantear el conocimiento *a priori* $D(\theta)$ cuando el conocimiento es «difuso» es sumamente extensa y poco relevante para nuestro objetivo, por lo que no entraremos en ella. Lo importante de (5) es que nos plantea claramente un proceso que integra el conocimiento previo con la información muestral. Si un nuevo conjunto de información empírica se hace disponible, es claro que la verosimilitud de dos conjuntos de datos (no traslapados) es el producto de verosimilitudes

$$\mathscr{L}\left(\theta\middle|Y,X\right)=\mathscr{L}\left(\theta\middle|Y^{(1)},X^{(1)}\right)\mathscr{L}\left(\theta\middle|Y^{(2)},X^{(2)}\right),$$

donde $Y=(Y^{(1)\prime},Y^{(2)\prime})'$, $X=(X^{(1)\prime},X^{(2)\prime})'$, con lo que el conocimiento se puede ir formando secuencialmente, conformando un verdadero proceso de aprendizaje

$$D(\theta),$$

$$D\left(\theta\middle|Y^{(1)},X^{(1)}\right)\propto D(\theta)\mathscr{L}\left(\theta\middle|Y^{(1)},X^{(1)}\right),$$

$$D\left(\theta\middle|Y,X\right)\propto D\left(\theta\middle|Y^{(1)},X^{(1)}\right)\mathscr{L}\left(\theta\middle|Y^{(2)},X^{(2)}\right), \text{ etc.}$$

Igualmente, el conocimiento nuevo podría provenir de fuentes no muestrales, pudiéndose integrar en la misma forma que un conjunto de información muestral al proceso de aprendizaje, simplemente volviendo a aplicar (5) para integrar la nueva información con el estado de conocimiento existente, es decir,

$$D\left(\theta\middle|Y,X,\mathscr{I}\right)\propto D\left(\theta\middle|\mathscr{I}\right)D\left(\theta\middle|Y,X\right), \quad \text{-(6)}$$

donde $\mathscr{I}$ representa el conjunto de información nueva, sea muestral o no. Si bien el planteamiento en (6) parece una extensión trivial de la regla de conocimiento bayesiana, lo cierto es que en la realidad el caso no es tan simple: es necesario que la nueva información en $D(\theta|\mathscr{I})$ sea independiente de la información en $D(\theta|Y,X)$, en particular, de los datos. Esto, claramente, excluye la posibilidad de incorporar como información extraña lo que se aprende de los datos mismos: debe haber una fuente independiente de dicha información y no se valida, desde luego, el proceso de «pesca» con los datos. Con el estado de conocimiento en (4), (5) o (6) podemos hacer todo tipo de proposiciones de confianza para los parámetros.

5. LA ESTRATEGIA DE CONOCIMIENTO

La estrategia de conocimiento utiliza los elementos anteriores para ir construyendo propiamente un modelo econométrico. El punto de partida es dar forma al modelo en (3) para el *PGI* en (1) con la selección de variables, especificación de los momentos e incorporación de supuestos en la parametrización. En la especificación del modelo se debe, en todo momento, mantener la reflexión con respecto a que las variables en el modelo tengan contrapartes observables cuya información empírica esté o vaya a estar disponible al investigador. Esta información empírica se incorpora para formar la función de verosimilitud en (4).

El siguiente paso es evaluar la especificación del *PGI*, en cuanto a la credibilidad de los supuestos incorporados para darle la forma estadística inicial. A esta etapa la llamamos la *Evaluación Econométrica*, y busca ver si los supuestos econométricos que se incorporaron, por sencillez en muchos casos, para dar forma al *PGI* no son demasiado restrictivos. Para hacer esta evaluación se sigue una *estrategia inductiva* o de *parsimonia*, que va de lo particular (imponer el supuesto) a lo general (no imponerlo).

Reescribimos el modelo en (3) como

$$Y_t = Z_t \beta + \varepsilon_t, \qquad \text{-(7)}$$

donde la variable $Z_t = (Y_{t-1}, X_t, X_{t-1})$ incluye tanto a las variables exógenas como a las predeterminadas (en el pasado) y el vector $\beta = (\alpha, \gamma_0, \gamma_1)$ contiene todos los parámetros de la media condicional. Si el modelo está correctamente especificado, entonces la relación en (7), debería estar descomponiendo a la variable Y_t (en su nivel) en un componente sistemático $\mu_t = Z_t\beta$ y uno no sistemático ε_t. En la medida en que este último componente contenga elementos sistemáticos, el modelo será incorrecto. Cuando esto suceda, habrá que corregir el modelo incorporando el aprendizaje derivado de la detección de componentes sistemáticos aún no explicados por el modelo, pues no están en el componente sistemático o explicado $\mu_t = Z_t\beta$. La corrección se da «agrandando» el modelo para que incorpore el nuevo elemento sistemático detectado, yendo de lo particular a lo general. Para detectar elementos sistemáticos no incorporados en $\mu_t = Z_t\beta$, tomamos el modelo

$$Y_t = Z_t\beta + e_t + \varepsilon_t = Z_t\beta + u_t,$$

donde $u_t = e_t + \varepsilon_t$ denota a un proceso de error o innovación que todavía puede tener algo de sistemático, en este caso representado por e_t, que no es sino el error de modelística dado por la diferencia entre el componente sistemático del *PGI* y el del modelo,

$$e_t = \mu_t - \mu_t(\beta) = \mu_t - Z_t\beta.$$

Así, cuando $e_t = 0$, tenemos que $u_t = \varepsilon_t$ y el modelo es correcto. La base del diagnóstico será proponer alternativas de forma que pueda tomar el error e_t, es decir, $e_t = W_t\gamma$ para algún conjunto de variables en el conjunto de información disponible, de suerte que si $\gamma = 0$ se sigue que $e_t = 0$. Podemos estimar el modelo ampliado

$$Y_t = Z_t\beta + e_t + \varepsilon_t = Z_t\beta + W_t\gamma + \varepsilon_t$$

y evaluar si $\gamma = 0$, según el cero esté contenido en la propuesta de confianza del nivel que requiramos. Si concluimos que $\gamma \neq 0$, el diagnóstico indica error de especificación y habría que proceder a re-especificar el modelo. Hay diagnósticos constructivos, que proporcionan información para esta re-especificación, y hay diagnósticos destructivos, que simplemente nos dicen que el modelo tiene problemas pero no nos da muchas pistas para arreglarlo. Mientras más alternativas consideremos para e_t que resulten en $e_t = 0$, más confianza tendremos en nuestro modelo. Escogeremos diferentes variables para evaluar correcta especificación, exogeneidad, linealidad, permanencia estructural o no autocorrelación.

Esto define un ciclo de modelística que nos remite a reconsiderar la especificación teórico-empírica del modelo, y del que debemos salir una vez que se ha considerado una gama amplia de posibilidades de errores de especificación en los distintos momentos condicionales (una amplia gama de variables W_t), en los que el cero resulte incluido en las proposiciones de confianza acerca del parámetro γ.

Una vez que el investigador está satisfecho con la evaluación econométrica de su relación, sigue la *Evaluación Económica*, que busca ver si las hipótesis económicas en estudio tienen sentido y son coherentes con la realidad observada en los datos y enmarcada en la especificación del *PGI*. Para hacer esta evaluación se sigue una *estrategia reductiva*, que va de lo general (no imponer el supuesto) a lo particular (imponerlo). Cada vez que se incorpora una hipótesis de este estilo, se evalúa su verosimilitud a través de proposiciones de confianza en el contexto del modelo. Esta fase termina cuando se han agotado las hipótesis por considerar, habiendo incorporado aquellas que resultaron en estadísticos de confianza adecuados y que el juicio o información del investigador consideró adecuado para el objeto de conocimiento. La consideración de cada grupo de hipótesis puede a su vez detonar la consideración de nuevas hipótesis o bien retroalimentar el conocimiento de forma tal que lleve a un ciclo de reconsideración del modelo y que puede volver a requerir de estimación, evaluación econométrica, etc. La independencia de la información nueva es crucial.

BIBLIOGRAFÍA

Asteriou, D. (2006), *Applied Econometrics*, Palgrave Macmillan, New York.

Charemza, W. W. y Deadman, D. F. (1997), *New Directions in Econometric Practice*, 2.ª ed., Edward Elgar, Northampton.

Davidson, R. y MacKinnon, J. G. (2004), *Econometric Theory and Methods*, Oxford University Press, New York/Oxford.

Edwards, E. (1971), *Likelihood*, Cambridge University Press, Cambridge.

Engle, R. F. (1982), «Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation»: *Econometrica*, 50, 987-1007.

Engel, R. F. y Granger, C. W. (1987), «Co-integration and error correction: representation, estimation and testing»: *Econometrica*, 55, 251-276.

Harvey, A. (1990), *The Econometric Analysis of Time Series*, MIT Press, Cambridge (Mass.).

Hendry, D. F. (1980), «Econometrics – alchemy or science»: *Economica*, 47, 387-406.

Intrilligator, M. D., Bodkin, R. G. y Hsiao, H. (1996), *Econometric Models, Techniques and Applications*, 2.ª ed., Prentice Hall, New Jersey.

Leamer, E. (1978), *Specification Searches: ad hoc inferences with non experimental data*, John Wiley, New York.

Maddala, G. S. (1992), *Introduction to Econometrics*, 2.ª ed., Macmillan, New York.

Pagan, A. R. y Hall, A. D. (1983), «Diagnostics tests as residual analysis (with discussion)»: *Econometric Reviews*, 2, 159-254.

Patterson, K. (2000), *An Introduction to Applied Econometrics*, St. Martin's Press, New York.

Sabau, H. (1998), *Econometric Inference with Heteroskedastic Models*, tesis doctoral, Australian National University, Canberra.

Sabau, H. (2007), *Análisis Econométrico Dinámico*, borrador, notas para libro, Universidad Iberoamericana, México.

Sabau, H. y Ruprah, I. J. S. (1984), «Modelos Econométricos para la Evaluación de la Política Económica: una perspectiva metodológica», en *Economía Mexicana, Serie Temática: Modelo Macroeconómico*, 9-23, CIDE, México.

Spanos, A. (1986), *Statistical Fondations of Econometric Modeling*, Cambridge University Press, Cambridge.

Wooldridge, J. M. (2000), *Introductory Econometrics: A Modern Approach*, South-Western College, Michigan.

Wooldridge, J. M. (2001), *Introducción a la Econometría: un Enfoque Moderno*, Thomson Learning, México.

Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*, John Wiley and Sons, New York.