

# Las interrelaciones entre investigación y docencia en Economía aplicada<sup>1</sup>

Alfonso Novales

[anovales@ccee.ucm.es](mailto:anovales@ccee.ucm.es)

*Departamento de Economía Cuantitativa. Facultad de Ciencias Económicas y Empresariales,  
Universidad Complutense de Madrid. Campus de Somosaguas. Pozuelo de Alarcón. 28223. Madrid.*

Recibido: 27 de febrero de 2009

Aceptado: 3 de abril de 2009

---

## Resumen

En la primera parte del trabajo se proponen algunas ideas acerca del modo en que las líneas de investigación recientes, junto con la mayor disponibilidad de medios de cálculo en las Facultades, debería cambiar sustancialmente los contenidos y métodos de las materias de Estadística y Econometría. La segunda parte describe algunas deficiencias importantes de la metodología habitual de investigación en Economía aplicada, poniéndolas en relación con conceptos básicos de las materias citadas. Una buena comprensión de estos conceptos en las Facultades evitaría incurrir en tales problemas.

**Palabras clave:** Econometría, metodología investigación, docencia, limitaciones.

**Códigos JEL:** A2, C1.

---

## 1. INTRODUCCIÓN

No puedo sino agradecer al Instituto de Ciencias de la Educación de la Universidad de Zaragoza la invitación a impartir una conferencia sobre un tema tan trascendente, especialmente en los momentos de cambio de planes de estudios que vivimos, como el impacto que la investigación puede y debe tener sobre la docencia en Economía. Pero no puedo olvidar la otra dirección de influencia, que es de hecho la que me viene preocupando y a la que he dedicado buena parte de mis esfuerzos de investigación en los últimos años: la incidencia que los conocimientos adquiridos en la Facultad debieran plasmarse en el modo en que hacemos investigación en Economía.

Aunque sería por supuesto muy interesante debatir acerca de la mutua influencia entre la investigación y la docencia en Economía, en todas sus áreas, pero un debate a tal nivel de generalidad excede de mis posibilidades. Prefiero, por tanto, centrarme en un área, la economía aplicada, y en unas materias, Estadística/Econometría, ganando quizá concreción en el detalle de las propuestas. Debería entenderse, sin embargo, que subyace a mis propuestas concretas un punto de vista más amplio sobre otras materias, como la

Macroeconomía y la Economía Financiera, entre otras, a las que no sería difícil extender propuestas paralelas a las que voy a hacer.

Es claro que la metodología de investigación debe tener una incidencia sobre la enseñanza de las materias incluso básicas, pues no en vano estamos formando investigadores de nivel más o menos alto ¿o cómo consideraríamos a cualquier profesional que prepara habitualmente informes utilizando datos de su empresa, de su economía nacional o de algún sector de la misma, o de entornos económicos más amplios? Y, sin embargo, no creo que a nivel global, estemos haciendo un buen trabajo al respecto, pues a pesar de los importantísimos cambios tecnológicos de las últimas décadas mantenemos unos contenidos definidos para un contexto profesional ya obsoleto, y continuamos utilizando una metodología docente asimismo propia de otra época, que difícilmente puede satisfacer las necesidades profesionales del mundo actual. Y esto puede ser especialmente cierto ahora que las nuevas titulaciones van a tener una mayor brevedad que las licenciaturas y posiblemente un tono más aplicado. No discutamos la idoneidad de algo que, en todo caso, no podemos controlar, aprovechemos más bien la situación que se nos plantea para dar un giro de timón a nuestras actividades docentes.

## **2. PRIMERA PARTE: LA INFLUENCIA DE LA INVESTIGACIÓN SOBRE LA DOCENCIA EN ECONOMETRÍA**

Cubriré la primera parte de mi exposición refiriéndome a la influencia que considero que la investigación debería tener sobre la docencia en la enseñanza de la asignatura de Econometría propia de una licenciatura (o de un futuro título de Grado).

Es claro que a lo largo de los años, la investigación en distintas áreas ha permitido acumular un cuerpo de conocimientos que ha podido hacer cambiar por completo nuestro enfoque acerca de algunas cuestiones conceptuales o de métodos de análisis. A modo de ejemplo, dentro de líneas metodológicas de investigación, basta pensar en la inmensa literatura desarrollada en los últimos años acerca de la estimación de modelos con variables no estacionarias, o la estimación de modelos con paneles de datos o con variables dependientes limitadas. Existen también múltiples ejemplos en otras áreas, como la teoría de la imposición óptima, la influencia de la calidad institucional en el desarrollo económico, o la generación e incidencia de las burbujas inmobiliarias o financieras que deberían tener un reflejo asimismo en programas y métodos docentes.

Tales bloques conceptuales y metodológicos deben incorporarse a la docencia, en algunas ocasiones como temas nuevos de un programa, y en otras como aspectos metodológicos. Pero habría que distinguir entre la docencia en programas de postgrado (Máster/doctorado), y en licenciaturas o grados. Es, desde luego, sencillo, adaptar o actualizar los programas y métodos en cursos de postgrado y, donde estos existen, es claro que la investigación tiene una incidencia muy importante sobre la elección de los contenidos, como creo que debe ser. Lamentablemente, la situación es totalmente la contraria en la docencia en las licenciaturas, donde creo que existe una excesiva inercia en la actualización de programas y métodos. No hay verdaderamente razones administrativas que impidan tal actualización, y creo que ello se debe a una escasa implicación del profesorado, globalmente considerado. Pero vayamos al debate acerca de qué aspectos de la investigación deben influir sobre la docencia en Econometría y el modo en que deben hacerlo.

Utilizando los ejemplos anteriores como vía de exposición, es claro que la correlación espuria, que siempre estuvo presente en los textos, pero a lo que quizá no se prestó suficiente atención, debe ser una cuestión tratada explícitamente, mediante ejemplos *numéricos*, en cualquier curso básico de Estadística o Econometría. Incluiría asimismo una presentación acerca de la filosofía general de los contrastes de raíz unitaria, pero no una presentación exhaustiva de los mismos, y discutiría numéricamente, *mediante simulación y contraste*, las principales propiedades estadísticas de algunas situaciones potencialmente frecuentes en datos reales que pueden afectar al resultado del contraste: parámetros cambiantes, procesos autorregresivos de primer orden con distinto coeficiente positivo, tomando valores entre 0 y 1, procesos con salto en media, procesos con varianza cambiante, procesos no lineales, etc.. Discutiría en un curso de este nivel la noción de cointegración y su contrastación vía Engle-Granger, pero no la teoría de Johansen, y explicaría las implicaciones del teorema de representación de Engle-Granger que conducen al modelo de corrección del error, interpretando su estructura, el papel de sus coeficientes y la hipótesis de ausencia de causalidad unidireccional. Todo esto utilizando al menos un ejercicio preparado con datos reales.

Incorporaría asimismo un tratamiento básico de la estimación mínimo-cuadrática de modelos estáticos con paneles de datos, enfocada al tratamiento de posibles efectos individuales no observables. Explicaría el diferente tratamiento preciso en función de la existencia o ausencia de correlación entre efectos individuales y variables explicativas, proponiendo el estimador de efectos fijos (*within*), o el obtenido mediante alguna de las transformaciones *between* o en diferencias, para el primer caso, y el estimador de efectos aleatorios (GLS) para el segundo caso. Nuevamente, todo ello ilustrado con algunos ejemplos preparados con datos reales. Y, por supuesto, una introducción a los modelos de variable dependiente limitada, al nivel de modelos de probabilidad lineal, probit y logit. Es muy posible que cualquier profesional que trabaje con datos se encuentre en situaciones donde deba utilizar alguna de estas herramientas: cointegración, métodos para paneles de datos, o para variables dependientes limitadas, y es una lástima que los alumnos terminen sus estudios de Economía o Administración de Empresas sin conocerlas.

Es evidente las dudas que esto suscita: a) ¿de dónde va a salir el espacio, en el programa de la asignatura, y el tiempo en la docencia para acomodar cuestiones nuevas? y b) ¿están los alumnos de este nivel en condiciones de asimilar tales conceptos? Ambos interrogantes constituyen el centro de la discusión. Adelanto mis respuestas: a) está justificado prescindir de algunos los contenidos actuales como comentaré en seguida, y b) los alumnos pueden asimilar perfectamente un curso de naturaleza aplicada, que incorpore los aspectos citados.

Voy a tratar de concretar ambas respuestas. Daría un giro definitivo, que personalmente llevo tiempo haciendo, hacia una docencia enfocada hacia el análisis, simple pero riguroso, de datos económicos. Creo que el mayor problema al que nos enfrentamos es que la creciente utilización de datos económicos, tanto en actividades profesionales como de investigación, culmina en ocasiones en una interpretación de resultados con lagunas de rigor suficientes como para conducir a conclusiones equivocadas. Soy consciente de la gravedad de la afirmación, pero estoy convencido de ello, y he reunido suficientes ejemplos como para que lo verdaderamente irresponsable fuese silenciarlo. No me refiero a métodos de análisis, sino a la interpretación de resultados obtenidos con base en nuestra comprensión de conceptos básicos, como describiré en la segunda parte de este trabajo.

En el diseño de la docencia, una de las dos sesiones semanales debería llevarse a cabo en un aula informática, y los alumnos deberían resolver cada dos semanas, en pequeños grupos, un determinado ejercicio práctico del tipo propuesto en textos de Econometría, aprovechando que ya es habitual que los archivos de datos sean de acceso público. Debería haber un mucho menor énfasis en probar resultados teóricos, y mucho mayor en entender e interpretar las propiedades teóricas y sus implicaciones prácticas.

Y, como debe ser siempre la norma, en cualquier asignatura, las cuestiones que se discutan en clase deben estar suficientemente motivadas, para lo cual es claro que nos queda hacer algo de trabajo. Porque, no estoy seguro de que proporcionemos respuestas suficientemente claras a preguntas como: a) ¿Qué interés tiene la ausencia de sesgo de un estimador en una ciencia no experimental?, b) ¿Por qué estimamos habitualmente por mínimos cuadrados?, c) ¿Es correcto interpretar individualmente cada coeficiente estimado en un modelo con varias variables explicativas como el efecto de cada una de ellas sobre el fenómeno en estudio?, d) ¿Cuáles son las implicaciones prácticas de la presencia de heterocedasticidad y autocorrelación? ¿En qué situaciones y para qué objetivos concretos son dichas implicaciones realmente preocupantes?

Pensar sobre cuestiones como éstas nos revela rápidamente la inadecuación actual de los programas tradicionales de Econometría. Consideremos la heterocedasticidad. Lo verdaderamente importante es la pérdida de precisión o eficiencia en la estimación, y la necesidad de utilizar una estimación robusta de la matriz de covarianzas del estimador de mínimos cuadrados como, por otra parte, es ya práctica habitual en el trabajo aplicado con datos. Y sin embargo, seguimos presentando en clase la expresión simplificada de dicha matriz como si fuese de aplicabilidad general. Lo que es peor, la verdadera matriz de covarianzas no es el estándar de cálculo en ninguno de los habituales programas estadísticos utilizados con los alumnos.

No es preciso mantener un capítulo del programa para discutir en detalle la contrastación, consecuencias y estimación bajo heterocedasticidad, que puede tratarse como una más de las cuestiones relativas a la estimación del modelo de regresión estándar, como es la colinealidad entre variables explicativas, por citar otro ejemplo de suma importancia. Mi afirmación está motivada a) porque como justificaré en la segunda parte, la consistencia del estimador debe preocuparnos en primera instancia, y su posible eficiencia en segunda instancia, y b) por la enorme dificultad que representa en la práctica tratar de aplicar con éxito en datos reales la transformación requerida por el estimador eficiente, de mínimos cuadrados generalizados. Pero estas cuestiones básicas no siempre se entienden bien, como muestra la lectura frecuente de frases del tipo: “*dada la evidencia de heterocedasticidad, estimamos el modelo anterior por Newey-West*”.

Algo similar diría acerca de la autocorrelación, aunque dándole una mayor relevancia al comentar la estimación de modelos dinámicos, por la inconsistencia del estimador de mínimos cuadrados que surge en este caso, o por las implicaciones que tiene de cara a la predicción. La atención habitualmente dedicada a la discusión sobre heterocedasticidad y autocorrelación, puede disminuirse muy sustancialmente. Y deben discutirse como situaciones que están presentes en todo modelo, salvo que se pruebe lo contrario, y que podemos tratar, aunque no siempre sea conveniente hacerlo, salvo en situaciones de inconsistencia, sino más bien tener en cuenta sus implicaciones a efectos de la matriz de covarianzas de los estimadores.

Centrarnos en la consistencia de los estimadores nos conduce a plantear el modelo de regresión bajo hipótesis del tipo de ausencia de correlación entre término de error y variables explicativas, ya sea en la forma  $E(x_{it}u_t) = 0, i = 1, 2, \dots, k$ , para todo  $t$  (variables explicativas predeterminadas), o  $E(u/X) = 0_T$  (variables explicativas exógenas), según el grado de exigencia que queramos imponer sobre la relación entre variables explicativas y término de error.

Otro capítulo tradicional de los programas de la asignatura es la multicolinealidad. Pero la multicolinealidad exacta no existe salvo en circunstancias especiales, que no cabe siquiera discutir, excepto en el caso de la conocida como “trampa de las variables ficticias” cuestión práctica de evidente interés y, afortunadamente, sencilla. Nuevamente, no es preciso dedicar un tema del programa a esta cuestión. Es muy importante, por el contrario, entender las distintas modelizaciones alternativas que pueden llevarse a cabo en un contexto de variables ficticias, familiarizar al alumno con la interpretación de cada una de ellas, y comprobar con datos cómo, si bien las estimaciones numéricas de los coeficientes varían con la especificación de las variables ficticias, las implicaciones prácticas del modelo son invariantes. Y es muy importante, por el contrario, explicar que la colinealidad es un problema de pérdida de precisión en la estimación y que, como tal, conduce a los sesgos lógicos en la contrastación de hipótesis paramétricas, pero no sobre la capacidad explicativa o predictiva del modelo.

En definitiva, creo más conveniente discutir las implicaciones prácticas y el tratamiento requerido por heterocedasticidad, autocorrelación y colinealidad, que consideraría situaciones siempre presentes en los datos, y dejaría de denominar “problemas”, lo cual no induce sino a una confusión fundamental respecto a las mismas.

La mejor manera y, casi, la única manera de definir un cambio metodológico docente como el que propongo es la exposición de material docente preparado al efecto.<sup>2</sup> Es lo que procede hacer, y no tiene mayor interés abundar en las cuestiones planteadas en el contexto de estas páginas, si bien considero que los puntos que he suscitado representan un cambio mayor en el contenido y, sobre todo, en la orientación docente.

### **3. SEGUNDA PARTE: LA UTILIZACIÓN DE LOS PRINCIPIOS ESTADÍSTICOS BÁSICOS EN LA INVESTIGACIÓN EMPÍRICA EN ECONOMÍA**

Pasemos ahora a considerar la relación en el otro sentido, la influencia que debiera tener la docencia en Estadística y Econometría sobre la investigación aplicada en Economía. Entiendo que puede parecer sorprendente pero, como dije antes, creo que los investigadores deberíamos recordar mucho más a menudo muchos de los conceptos básicos que aprendimos en los cursos de Estadística o Econometría. Quizá lo que sucede es que dichos conceptos deberían enseñarse de modo mucho más explícito en tales cursos. Es, en ocasiones, verdaderamente llamativo que investigadores que imparten docencia en estas materias, puedan incurrir en determinadas carencias en su trabajo de investigación aplicado.

Hay tres cuestiones particularmente importantes en cuanto a condicionar el alcance y las pretensiones que puede tener el trabajo empírico:

- A pesar de la insistencia que se pone sobre el tema en las consideraciones introductorias de muchos textos y cursos de Econometría, pronto olvidamos acerca del carácter no experimental de la (ciencia) económica. La imposibilidad de experimentar: a) impide demostrar proposiciones mediante el análisis de la única muestra disponible, b) impide inferir relaciones de causalidad a partir de las correlaciones observadas, c) nos hace trabajar con inputs (variables explicativas) no controlables, generalmente colineales y, como consecuencia, d) condiciona de forma muy específica la interpretación de los coeficientes estimados en un modelo.
- Una de las razones más habituales para el trabajo con datos en Economía consiste en el interés por contrastar en última instancia una o varias hipótesis paramétricas. En algunos casos, con ello pretendemos discriminar entre teorías alternativas, lamentablemente estimulados por los comentarios vertidos en la Introducción a muchos textos de Econometría. ¿Pero podemos encontrar evidencia concluyente en los datos (la única muestra disponible) acerca de un determinado modelo o de que una determinada teoría sea correcta? ¿Son justificadas las afirmaciones del tipo: “*We have shown, beyond any reasonable doubt, that...*”, “*We have found conclusive evidence on the relevance of X to explain the behavior of Y*”, que con frecuencia leemos en la sección de conclusiones de artículos académicos publicados en las mejores revistas internacionales?
- El contraste más habitual en el trabajo aplicado en Economía es un contraste de significación estadística de un coeficiente, mediante el que pretendemos concluir acerca de la capacidad explicativa (¿o contenido informativo, quizás?) de una variable del modelo sobre la variable dependiente. ¿Está justificada la identificación de estos dos conceptos: significación estadística y capacidad explicativa?

Son cuestiones que se refieren a básicos aspectos que debieran estar resueltos a partir de los contenidos que habitualmente se discuten en las asignaturas del área. Claro está, siempre que la docencia preste una atención central a la interpretación y motivación de los temas que cubre y que, como bien expresa el refrán: *las ramas no nos impidan ver el bosque*. Son asimismo cuestiones de suma importancia, cada una de las cuales genera una clase de deficiencias en el trabajo empírico en Economía, entre ellas:

- Ignorar la precisión en la estimación, y su relación con la potencia de los contrastes de hipótesis paramétricas.
- Identificar la decisión de no rechazar la hipótesis nula con haber probado que dicha hipótesis es cierta, como consecuencia del total olvido de la función de potencia en el análisis empírico en Economía
- Interpretar coeficientes individuales en modelos con múltiples variables explicativas, como efectos de cada una de ellas sobre la variable dependiente,
- Identificar, por confusión, dos propiedades tan distintas como la significación estadística de un *coeficiente* estimado, y el contenido informativo en la *variable* que le acompaña, que afectan a entes de distinta naturaleza.
- Mantener en el tratamiento de datos la pretensión de que estos nos permiten demostrar teorías, o probar una teoría como falsa o cierta, en lugar de buscar en ellos su posible consistencia (o corroboración) con una determinada creencia conceptual a priori. ¿Podemos demostrar proposiciones con los datos? ¿Qué es un buen modelo? ¿Cómo queremos comparar modelos alternativos? ¿Queremos decir que entre  $k$  modelos alternativos,  $k-1$  modelos son igualmente falsos?
- Dedicar una escasa atención a posibles variaciones paramétricas

Repasemos con la brevedad que este formato nos permite, algunas de estas cuestiones.

### 3.1. RELACIÓN ENTRE PRECISIÓN Y POTENCIA

Al estimar la ecuación  $y_t = \alpha + \beta x_t + u_t$ , en la que estamos interesados en contrastar la hipótesis nula:  $H_0 : \beta = 1$  frente a la alternativa  $H_0 : \beta < 1$ , consideremos un coeficiente estimado:  $\hat{\beta} = 0,50(0,40)$ . La hipótesis nula no se rechaza y, suponiendo Normalidad, el valor-p del contraste es 0,106. La misma conclusión se alcanzaría con un estimación  $\hat{\beta} = 0,50(1,0)$ , que tendría un p-valor de 0,308. Pero a los efectos que nos llevaron a plantear la posibilidad de una elasticidad unitaria en el modelo, que sugiere que fluctuaciones en  $x_t$  se transmiten en igual cuantía a  $y_t$ , ¿es admisible una estimación  $\hat{\beta} = 0,50$ ? No lo parece.

No debería resultarnos sorprendente que si decidiéramos cambiar por completo nuestra creencia, y pasar de la duda sobre el valor unitario de la elasticidad, a cuestionarnos por completo la capacidad explicativa de la variable  $x_t$ , tampoco rechazaríamos la hipótesis nula, y concluiríamos que las variaciones en  $x_t$  no contribuyen a explicar los cambios en  $y_t$ . Así de fácil es obtener conclusiones equivocadas si aplicamos de manera mecánica métodos estadísticos cuyo diseño es impecable, sin pensar en la motivación que nos lleva al análisis. En este ejemplo, podríamos concluir que  $x_t$  es un predictor insesgado de  $y_t$ , es decir, que la constante del modelo es igual a cero y la pendiente del mismo igual a 1. Pero también podríamos mantener que  $x_t$  no contiene información sobre  $y_t$ , si bien es de esperar que no sea el mismo investigador quien deduce ambas conclusiones. Y, lo que es quizá más llamativo, es que hay razones de peso para pensar que ambas conclusiones son equivocadas. Es más que probable que  $x_t$  contenga información útil sobre  $y_t$ , si bien no siendo un factor de enorme relevancia, y también que su elasticidad posiblemente sea inferior a la unidad.

El error en este caso proviene de no constatar que la situación de estimación es de una reducida precisión, como manifiesta una desviación típica casi igual al coeficiente estimado. La pérdida de precisión conduce a bandas de confianza excesivamente amplias a los niveles de significación habituales, por lo que no se rechazaría casi ninguna hipótesis nula. En esta situación los contrastes de hipótesis pierden potencia, sea cual sea la hipótesis nula. La pérdida de potencia no está relacionada con la posible hipótesis nula que se contrasta, sino con la pérdida de precisión en la estimación que puede derivarse de una reducida cantidad y/o calidad de la información muestral,<sup>3</sup> posible inestabilidad paramétrica o de la forma funcional de la relación estimada a través de la muestra, un tamaño importante del término de error, entre otros factores. Simplemente, la reducida precisión debería llevarnos a no contrastar hipótesis acerca del verdadero valor numérico del parámetro.

Y, por supuesto, como muestran estos sencillos ejemplos, *no rechazar una hipótesis nula no es*, en modo alguno, *equivalente a haber probado su validez*, a pesar de la frecuencia con que el investigador manifiesta haber encontrado evidencia favorable a su creencia. Este error conceptual se deriva precisamente de la total ignorancia en que hemos dejado la función de potencia de un contraste de hipótesis. La razón básica por la cual ambos conceptos no son equivalentes es que la falta de precisión en la estimación puede conducir, por las razones arriba apuntadas, a no rechazar la hipótesis nula aun a pesar de que la coherencia entre la evidencia muestral y dicha hipótesis sea muy débil.

### 3.2. CONTRASTES UNILATERALES

Con frecuencia se olvida que son precisas dos condiciones para rechazar una hipótesis nula: a) que la evidencia muestral sea contraria a la hipótesis nula, y b) a la vez que es favorable a la hipótesis alternativa. En contrastes bilaterales ambas condiciones son equivalentes, lo que hace que tendamos a pensar con excesiva frecuencia que la hipótesis nula se rechazará siempre que se cumpla la primera condición. Sin embargo, este no es el caso en contrastes unilaterales, en los que ambas condiciones no son equivalentes y deben satisfacerse ambas para que los tests de hipótesis Neyman-Pearson conduzcan al rechazo de la hipótesis nula. Ignorar este hecho conduce en ocasiones a una interpretación errónea de los resultados positivos (no rechazo de la hipótesis nula) en este tipo de contrastes.

Conviene añadir asimismo, que los contrastes unilaterales deberían ser bastante más frecuentes de lo que son en el trabajo aplicado en Economía, pues existe una gran cantidad de situaciones en las que en caso de decidir acerca del incumplimiento de la hipótesis nula, el investigador estaría en condiciones de afirmar que el verdadero valor numérico del parámetro desconocido se halla a un determinado lado de dicha hipótesis nula.

Supongamos que para el contraste de la hipótesis nula:  $H_0: \beta = 1$  frente a la alternativa  $H_0: \beta < 1$ , disponemos de un coeficiente estimado:  $\hat{\beta} = 1,15(0,30)$ . No es preciso realizar el contraste, pues no rechazaremos la hipótesis nula, debido a que la evidencia muestral, resumida en el estimador disponible, *no es favorable a la hipótesis alternativa*. Tampoco rechazaríamos la hipótesis nula si la estimación hubiera sido:  $\hat{\beta} = 1,50(0,20)$ , un caso en que sería dudoso creer que el verdadero valor del parámetro desconocido es  $\beta = 1$ . Ignorar esta particularidad de los contrastes unilaterales, junto con identificar, nuevamente de manera errónea, el no rechazar la hipótesis nula con haber probado su veracidad, nos llevaría a concluir que la hipótesis nula  $H_0: \beta = 1$  es correcta. La única decisión razonable en estos casos es no llevar a cabo el contraste, por supuesto, y plantear por qué se ha obtenido una estimación numérica que puede discrepar tanto de la hipótesis nula como de la hipótesis alternativa. Hay tres posibilidades: a) concluimos, de modo informal, que la desviación respecto de la hipótesis nula es suficientemente pequeña como para ser interpretada como error muestral o estadístico, b) creemos que la muestra carece de calidad como para que la estimación numérica resulte relevante, o c) debemos considerar detenidamente la creencia en una hipótesis alternativa unilateral.

### 3.3. SIGNIFICACIÓN ESTADÍSTICA Y RELEVANCIA ECONÓMICA

El contraste de hipótesis más habitual en Economía consiste en la significación individual de un coeficiente. Este tipo de contrastes está sujeto a los comentarios anteriores acerca de los casos en que el contraste sea unilateral, así como acerca de la contrastación de hipótesis en contextos de reducida precisión en la estimación. Pero existe un problema adicional de bastante importancia, consistente en la injustificada identificación de la ausencia de significación estadística de un coeficiente, por un lado, con la ausencia de contenido informativo en la variable que le acompaña, por otro.

Volvamos a los comentarios acerca de la relación entre pérdida de precisión en la estimación y pérdida de potencia. El estadístico  $t$  para el contraste de la hipótesis nula de significación es el cociente entre el valor numérico estimado para el coeficiente, y la desviación típica estimada para el mismo. En contextos de baja precisión en la estimación, la desviación típica estimada será elevada, y el estadístico  $t$  tenderá a ser reducido, con independencia del valor numérico estimado para el coeficiente. Es un ejemplo más de las consecuencias de la pérdida general de potencia. En el caso particular de contrastes de ausencia de significación, tenderemos a no rechazar con excesiva frecuencia la hipótesis nula, concluyendo que el coeficiente no es significativamente distinto de cero. El problema es que enlancemos la lógica del argumento hasta concluir, injustificadamente, que la variable asociada no contiene información acerca del fenómeno que se está tratando de explicar.

Es perfectamente posible que una variable sea un determinante relevante del fenómeno en estudio, pero la muestra no permita estimar con suficiente precisión el impacto de dicha variable. Cuando esto sucede, la conjunción de ambos hechos, la reducida precisión, y una interpretación incorrecta de la ausencia de significación estadística, nos puede conducir fácilmente a concluir la irrelevancia de la variable explicativa. Varias razones pueden conducir a la pérdida de precisión en estimación, pero quizá la más importante sea la variabilidad temporal (o la variabilidad a lo largo de la muestra de corte transversal) de la relación entre ambas variables.

También puede producirse la situación inversa: Quizá por contar con un tamaño muestral muy elevado, o porque la variable asociada presente una alta variabilidad muestral, la precisión en la estimación puede ser muy elevada, estimándose por tanto una desviación típica muy reducida para el impacto que se está tratando de cuantificar. Es perfectamente posible que el estadístico  $t$  asociado a un coeficiente numérico muy reducido sea, sin embargo, suficientemente elevado (digamos que mayor que 2,0). Este es un caso en que rechazaríamos la hipótesis nula de ausencia de significación estadística; la identificación conceptual a que nos venimos refiriendo nos llevaría a concluir que la variable asociada tiene una capacidad explicativa importante, cuando lo contrario puede ser cierto. El coeficiente podría ser tan reducido, que incluso un cambio en la variable explicativa, en una cuantía importante para lo observado en la muestra, pudiese tener un efecto despreciable sobre la variable dependiente, nuevamente en términos de su variabilidad muestral.

En definitiva, el concepto de significación estadística se refiere a la estimación numérica de los coeficientes estimados en modelos econométricos. Los comentarios anteriores cuestionan, sin embargo, la relevancia económica de dicho concepto. No existe el concepto de significación estadística de una variable, y no tiene sentido afirmar que *una variable explicativa sea o no estadísticamente significativa*.

Una variable puede contener o no información acerca de las fluctuaciones que experimenta en la muestra la variable dependiente, con independencia de la significación estadística de la estimación del coeficiente asociado.

### 3.4. ¿HAY RELACIÓN ENTRE EL VALOR DEL ESTADÍSTICO T DE STUDENT Y LA RELEVANCIA ECONÓMICA DE LA VARIABLE?

Los procedimientos para evaluar la capacidad explicativa de una variable no pueden reducirse a la contrastación de la ausencia de significación estadística de un coeficiente.

En algunos casos, aunque no estoy seguro de que se haga conscientemente, el investigador procede a pronunciarse acerca de la relevancia cuantitativa de una variable explicativa. Generalmente esto sucede una vez que se ha comprobado la significación estadística de su coeficiente. Pero entonces surge otro problema importante, dada la frecuencia con que se utilizan estos procedimientos en el trabajo aplicado, y es que tampoco puede concluirse acerca de la relevancia de una variable explicativa examinando la magnitud del coeficiente estimado, ni de su estadístico t de Student. Es bastante habitual afirmar que si un coeficiente estimado es de 5 unidades, por ejemplo, la variable asociada es bastante relevante. En definitiva, una variación de una unidad en la variable explicativa implicaría una variación de 5 unidades (!) en la variable dependiente. Evidentemente, también existe la tendencia a concluir que con un coeficiente estimado de 0,05, incluso si el estadístico t de Student es superior a 2,0, la variable asociada no es “*muy relevante*”. Ninguna de estas dos afirmaciones está, por sí sola, justificada, sin examinar las propiedades muestrales de ambas variables.

No conviene olvidar un principio general de inferencia estadística a partir de la información muestral, como es la necesidad de limitar el análisis de inferencia a situaciones observadas en la muestra. Es decir, si el rango muestral observado de una variable explicativa es de 50 a 100, digamos, no conviene inferir del modelo estimado qué efecto tendría sobre la variable explicativa un incremento en la variable explicativa (supongamos que controlable por la autoridad económica o por el responsable del análisis) de su valor actual a un valor de 1.000. Del mismo modo que no conviene utilizar un modelo estimado en presencia de un tipo impositivo sobre el consumo que ha tomado diversos valores a lo largo de la muestra (digamos que entre 10% y 25%), para evaluar el impacto que sobre el consumo tendría la desaparición de dicho impuesto, o para elevarlo sensiblemente por encima de su máximo valor muestral observado.

En este sentido, podemos calcular distintos indicadores de variabilidad muestral de una variable explicativa: su rango de variación o soporte de la distribución de valores muestrales de dicha variable, como indicador de variabilidad total o, si no queremos incluir valores extremos, su rango intercuartílico, o el rango entre los percentiles 10% y 90%, o su desviación típica muestral. La desviación típica es el indicador apropiado de la variación que cabe observar entre dos observaciones muestrales consecutivas de una variable estacionaria (aunque quizá debemos corregir por autocorrelación). Ambas medidas son complementarias, proporcionando información bien distinta y en ambos casos bien importante.

Si queremos evaluar la relevancia cuantitativa, o económica, en nuestro caso, de una variable explicativa, deberíamos corregir, es decir, multiplicar, su coeficiente estimado por

alguno de los indicadores mencionados, sea un rango o la desviación típica de dicha variable explicativa. El producto de ambos puede ponerse en relación, como porcentaje del rango comparable o de la desviación típica de la variable dependiente.

En un modelo con una única variable, estaríamos estimando de este modo la capacidad de la variable explicativa, según recorre su rango de variación, de reproducir el rango de variación observado del fenómeno en estudio. Es perfectamente posible que un elevado coeficiente, junto con un reducido rango de variación de la variable explicativa, sea incapaz de reproducir el rango de variación de la variable dependiente. Bien es cierto que este es un caso en que el coeficiente de determinación será generalmente reducido y, precisamente, una elevada *t* de Student, junto con un reducido  $R^2$  es evidencia sospechosa acerca del estadístico *t*. La información proporcionada por la desviación típica es diferente y, por construcción, más afín a la utilizada en el cálculo del estimador de mínimos cuadrados. Un diseño experimental donde la variable dependiente se genera a partir de una variable explicativa y un término de error, siendo la relación de varianzas muy favorable al término de error, va a ser un caso en que el producto del coeficiente estimado por la desviación típica de la variable explicativa puede ser notablemente inferior a la desviación típica de la variable dependiente, incluso con un valor numérico del estadístico *t* de Student que denote significación estadística del coeficiente estimado a los niveles de significación habituales.

### 3.5. ECONOMETRÍA DE SIGNOS Y DE ASTERISCOS

Afirmaciones habituales del tipo: “*El coeficiente estimado tiene signo contrario al esperado, pero no es estadísticamente significativo*” o, su dual: “*El coeficiente tiene el signo acorde a la teoría subyacente, pero no es significativamente diferente de cero*”, pero constituyen observaciones interesantes, aunque carecen del sentido que habitualmente reciben. Lo que sucede es que su interés depende de la relevancia cuantitativa o económica de la variable asociada, dado el valor numérico estimado para el coeficiente. Por tanto, en las frases anteriores sería preciso sustituir la referencia a la significación estadística del coeficiente o su ausencia de significación, por la posible relevancia cuantitativa de la variable asociada.

Y, en línea con estos comentarios, la tan habitual práctica de señalar con un determinado número de asteriscos los coeficientes estimados en distintos modelos, según su valor *p*, no sólo no carece de interés, sino que se presta y, muy posiblemente se debe, al error conceptual que he descrito en el epígrafe anterior. Que el estimador del coeficiente de una variable explicativa tenga un *p*-valor de 0,01 o inferior no significa en modo alguno que la variable asociada sea más relevante que otra cuyo coeficiente asociado tienen un estimador con *p*-valor de 0,04 ó 0,08. La práctica de asignar un número de asteriscos en función del *p*-valor de un coeficiente estimado, y utilizarlos en la interpretación del modelo para evaluar la contribución de las distintas variables a la explicación del proceso en estudio es errónea por al menos tres razones: a) la injustificada identificación entre significación y contenido informativo, b) la ambigüedad inherente al valor numérico del estadístico *t*, c) la colinealidad entre variables explicativas a la que enseguida me referiré.

### 3.6. COMPARACIÓN DE CAPACIDAD EXPLICATIVA ENTRE VARIABLES, O ENTRE SUBMUESTRAS

Aún más llamativa, y más frecuente, es la utilización del estadístico  $t$  de Student para evaluar el contenido informativo (habitualmente referido como la capacidad explicativa de una variable). Tal uso se basa en la errónea identificación conceptual entre significación estadística y relevancia cuantitativa a la que ya he hecho referencia. Pero es sorprendente la extendida creencia acerca de una posible relación monótona entre valor absoluto del estadístico  $t$  y capacidad explicativa de la variable asociada, como se refleja en afirmaciones del tipo: “con un estadístico  $t$  superior a 8,0, *la variable X muestra una enorme capacidad explicativa*”. En ocasiones, los estadísticos  $t$  obtenidos al estimar el modelo en dos submuestras, seleccionadas con el objeto de contrastar la homogeneidad paramétrica, o para evaluar los efectos de una determinada intervención, son comparados, llegando a afirmar: “*nótese cómo la variable X se ha hecho mucho más significativa en la segunda submuestra que en la primera*”.

Ya hemos comentado en detalle las serias dificultades a las que está sujeta la utilización del estadístico  $t$  de un coeficiente para evaluar la capacidad explicativa de una variable. La comparación de los valores numéricos del estadístico  $t$  de un mismo coeficiente en dos submuestras puede tener que ver tanto con el contenido informativo relativo que en ambas submuestras tiene la variable asociada, como con la precisión con que dicho coeficiente se estima en cada una de ellas.

Si difícil justificación tiene tal práctica, mucho menor la tiene la comparación de los estadísticos  $t$  asociados a los coeficientes de dos variables diferentes, cuando, por comparación de los valores numéricos de sus respectivos estadísticos  $t$ , se afirma: “*Entre los posibles determinantes de Y considerados, X es sin duda el más relevante*”. Lamentablemente, este tipo de análisis es muy habitual.

### 3.7. ¿QUÉ INFORMACIÓN PROPORCIONA UN MODELO CON VARIAS VARIABLES EXPLICATIVAS?

Tras estimar un modelo que pretende explicar el comportamiento de una determinada variable, es habitual analizar los coeficientes estimados individualmente, para evaluar la capacidad explicativa de cada una de las variables. Esta práctica está sujeta, por un lado, a la crítica anterior acerca de la utilización de los coeficientes estimados por sí solos, o a la comparabilidad de estadísticos  $t$  de distintos coeficientes. Está, sujeta asimismo a las dificultades que surgen de la probable colinealidad entre variables explicativas a la que me referiré más adelante.

Un coeficiente individual puede ayudarnos a examinar el efecto que sobre la variable en estudio tendría una variación de una desviación típica en una de las variables explicativas, o si ésta recorriera su rango de variación. Pero ello se hace bajo el habitual supuesto “*ceteris paribus*”, que plantea la eventual variación en una única variable explicativa, mientras las restantes permanecen constantes. Suceso este muy poco verosímil en caso de correlaciones no despreciables entre las variables explicativas. Incluso si efectuamos el cálculo cuidadosamente, obtendríamos una respuesta válida, sólo que a una pregunta poco interesante, por referirse a un suceso inverosímil.

Aunque es posible hacer el cálculo oportuno, *un modelo con varias variables explicativas no está diseñado para evaluar el efecto total que una de las variables explicativas tiene sobre la variable dependiente*. Cada uno de sus coeficientes mide el impacto parcial que una variación producida únicamente en la variable que le acompaña, tendría sobre la variable dependiente. Dicho de otro modo, el contenido informativo de una variable X sobre el fenómeno en estudio es siempre un *concepto condicional* en la información proporcionada por el resto de las variables explicativas ya incluidas en el modelo.

En un modelo lineal con múltiples inputs (variables explicativas) el coeficiente de dicha variable nos proporcionará el *efecto parcial*, ya comentado, válido únicamente bajo el supuesto “*ceteris paribus*”. Pero dicho modelo podría utilizarse para estimar el *efecto total* que sobre la variable dependiente tendría una variación en una de las variables explicativas. Si sometemos a las restantes variables explicativas a los cambios que cabría esperar que experimentasen cuando la variable explicativa que nos ocupa varía en una unidad (o en la cuantía supuesta), podríamos agregar los efectos que individualmente se generen sobre la variable dependiente. En el caso de dos variables explicativas, la triangularización o descomposición de Cholesky de la matriz de covarianzas de ambas variables explicativas permite realizar este cálculo con bastante sencillez.

El resultado que se obtiene de esta evaluación indirecta que agrega diversos efectos individuales, es prácticamente idéntico al que se obtiene del modelo de regresión simple, en respuesta a una variación como la planteada para la primera variable explicativa, por lo que podemos ahorrarnos el esfuerzo de cálculo que acabo de describir. Lo que nos lleva a una conclusión bastante sorprendente desde el punto de vista de la orientación habitual de los cursos de Econometría:

- El modelo de regresión simple nos proporciona una estimación *insesgada* del impacto global o total que una variable explicativa tiene sobre la variable dependiente objeto de estudio. Nos proporciona asimismo una estimación *sesgada* del efecto individual o parcial de dicha variable explicativa, lo que explica la denominación de “*sesgo de variables omitidas*” que se utiliza en los cursos habituales.
- Supuesto que no hayamos omitido ninguna variable relevante, un modelo con varias variables explicativas proporciona una estimación posiblemente *insesgada* del efecto parcial que sobre la variable dependiente tiene una variación en cada una de las variables explicativas. Por el contrario, salvo que se instrumenten los cálculos indirectos que he descrito, el modelo con múltiples determinantes proporciona una estimación *sesgada* del efecto global que sobre la variable dependiente tiene una variación en una determinada variable explicativa.

### 3.8. RESUMEN EXCESIVO DE LA INFORMACIÓN MUESTRAL

No puede considerarse muy eficiente reducir la información muestral a unos pocos estadísticos t, F, R<sup>2</sup> o valores del logaritmo de la función de verosimilitud. La evaluación de un modelo, o su comparación con modelos alternativos, debe basarse en utilizar de manera descriptiva la información muestral.

Al reducir excesivamente la información muestral, puede llegar a afirmarse, por ejemplo, que dos modelos explican lo mismo, porque tienen un R<sup>2</sup> similar, o un valor muy próximo del logaritmo de la función de verosimilitud, cuando cada uno de ellos explica

bien una parte de la muestra y mal la otra. Su capacidad explicativa podría, por tanto, ser completamente distinta.

En ocasiones, dos modelos pueden explicar de modo muy similar la mayor parte de la muestra, y diferir en un porcentaje reducido de observaciones muestrales, quizá correspondientes a un corto episodio temporal o a los valores máximos (o mínimos) observados de la variable dependiente. Un estadístico global de ajuste no revelará tal diferencia que podría ser, por otra parte, de gran interés para el investigador. Para desvelar este tipo de información, es preciso utilizar descriptivos del ajuste de cada modelo, que permitan contemplar toda la información muestral, es decir, la capacidad de ajuste del modelo para cada observación muestral. Es claro que en muestra grandes (por ejemplo, de más de 200 observaciones) tal examen exhaustivo puede ser difícil, por ejemplo, en términos gráficos. Es preciso en tales casos examinar el ajuste en submuestras, que serán cronológicas en el caso de datos temporales. En el caso de datos transversales, conviene establecer distintas clasificaciones, según los rangos de valores de la variable dependiente o de alguna de las variables explicativas, y examinar el ajuste en cada submuestra. En este caso, un mal ajuste para valores grandes o para valores pequeños de una variable explicativa puede revelar una dependencia no lineal con respecto a esta variable, que puede ser fácilmente incorporada a una nueva y mejora especificación del modelo.

### 3.9. UN ENFOQUE ALTERNATIVO PARA CONTRASTAR EL CONTENIDO INFORMATIVO DE UNA VARIABLE SOBRE OTRA

La discusión anterior ha puntualizado la imposibilidad de contrastar formalmente, con carácter absoluto, la capacidad explicativa de una variable. En un modelo con múltiples determinantes de la variable dependiente, es posible contrastar *si uno de dichos determinantes aporta información sobre el fenómeno en estudio, que no esté ya contenida en el resto de los determinantes incluidos en el modelo.*

Estamos nuevamente apelando al hecho de que en tal modelo, los efectos atribuidos a cada variable son efectos parciales. Sólo en ausencia de correlación entre variables explicativas, en un contexto lineal, coincidirían los efectos parciales con los efectos totales, como es bien sabido. Otro modo de constatar este resultado es que, en el cálculo indirecto que antes apuntábamos, la ausencia de correlación entre variables explicativas haría innecesario dicho cálculo.

En todo caso, lo antes comentado se refiere a contabilizar numéricamente el impacto global o total de una variable explicativa sobre la variable dependiente, pero no se presta a un contraste formal de dicha hipótesis. Sería factible hacerlo a través de un contraste tipo F de una particular combinación lineal de coeficientes, la definida por la descomposición de Cholesky, pero estaríamos incurriendo de nuevo en la interrelación entre precisión en la estimación y potencia en el contraste a que ya nos hemos referido con anterioridad.

Consideremos como hipótesis nula  $H_0$ : *La variable Z no añade información sobre Y a la información ya contenida en las variables que constituyen el vector X.* Evidentemente, tal hipótesis dista de afirmar que la variable Z no contiene información sobre Y, pero es una hipótesis contrastable. Consideremos ahora la siguiente definición: *“La variable Z no aporta información sobre Y a la información ya contenida en el vector X si y sólo si los residuos del modelo que explica Y*

*mediante X no difieren significativamente de los residuos del modelo que explica Y mediante (X,Z)."* El lector apreciará la similitud con definiciones del concepto de causalidad. Posiblemente sería más riguroso si estableciéramos la definición en términos de la comparación de esperanzas condicionales:  $E(Y / [X, Z]) = E(Y / X)$ , pero podemos dejarlo en los términos coloquiales que hemos propuesto.

Una crítica razonable de esta propuesta podría aducir tres elementos: a) ¿qué entendemos por residuos significativamente diferentes?, b) ¿por qué tiene interés contrastar la significación estadística de la diferencia entre dos conjuntos de residuos y no lo tiene contrastar la significación estadística de un coeficiente?, c) llevar a cabo un contraste de igualdad de residuos, utilizando para ello algún estadístico global ¿no nos va a hacer incurrir en un excesivo resumen de la información muestral?

Antes de proseguir con esta discusión, quizá valga la pena hacer algunos comentarios:<sup>4</sup>

- La lógica de análisis propuesta puede utilizarse para contrastar, en términos absolutos, el contenido informativo de una variable X sobre el fenómeno en estudio, representado por la variable Y. En tal caso, el resto de las variables estaría constituido por el término constante, y los residuos de este modelo restringido serían los datos de Y en diferencias respecto de su media muestral. La comparación de los residuos del modelo que incluye X como variable explicativa, con los datos de Y en desviaciones respecto de su media constituiría el contraste de contenido informativo absoluto.
- Nuevamente el mismo enfoque puede utilizarse para comparar la capacidad explicativa de dos variables, W y Z. Si es una comparación sobre su contenido informativo en términos absolutos o total, basta ver qué conjunto de residuos, en ambas regresiones simples, incluyendo alternativamente a W o a Z como variables explicativas, difieren en mayor medida de los datos de Y en desviaciones respecto de su media. Si la afirmación es condicional en la información proporcionada por un vector de variables X, se trataría de comparar los residuos del modelo que utiliza X como variables explicativas, con los residuos de las regresiones de Y sobre (X,W) por un lado, y de Y sobre (X,Z), por otro.
- Una alternativa consistiría en construir<sup>5</sup> tanto las variables  $W_Z$ , componente de W no correlacionado con Z, como  $Z_W$ , componente de Z no correlacionado con W, y comparar el ajuste del modelo que explica Y mediante Z con el ajuste que se obtiene al añadir  $W_Z$ , y comparar el ajuste del modelo que explica Y mediante W con el ajuste obtenido al añadir  $Z_W$ . Generalmente, estas comparaciones son suficientes para aclarar qué variable es más relevante.<sup>6</sup>

### 3.10. EL USO RAZONABLE DE INSTRUMENTOS ESTADÍSTICOS BÁSICOS

Pero queda un problema por resolver: ¿es posible contrastar la igualdad de residuos sin incurrir en las mismas contradicciones ya apuntadas para los contrastes de significación? Puede ser posible, puesto que se trata de dos situaciones muy diferentes. Estamos proponiendo ahora contrastar la igualdad de dos conjuntos, no la igualdad de dos estimaciones paramétricas. La precisión en la estimación quedará recogida en la magnitud

de los residuos, y no juega un papel en la comparación entre los residuos proporcionados por los dos modelos alternativos.

La dificultad ahora es técnica, por cuanto que no contamos con una teoría muy apropiada para contrastar la igualdad de conjuntos de residuos. Viene a la memoria la clase de contrastes estadísticos no paramétricos que tan útiles resultan en ciencias sociales. Son contrastes que permiten contrastar si dos muestras proceden de la misma distribución, por ejemplo. Tienen la virtud de no estar basados en ningún supuesto sobre la distribución de probabilidad del término aleatorio del modelo que relaciona  $Y$  con sus determinantes, lo cual es extremadamente conveniente, dada la reducida información con que contamos sobre ello. Tienen, en su mayoría, valores críticos válidos en muestras finitas, por mucho que en la mayoría de los casos haya que interpolar entre los tamaños muestrales tabulados y, en ocasiones, entre los niveles de significación. Además, la mayoría de ellos acepta aproximaciones por distribuciones continuas conocidas (Normal, en muchos casos), para tamaños muestrales incluso moderados.

Lamentablemente, este enfoque no nos resulta útil. Necesitamos comparar los residuos ordenados que surgen de ambos modelos, porque también queremos compararlos con cada observación muestral disponible. No nos basta con que procedan de igual distribución. De hecho, esto es poco relevante para nuestro análisis. Podrían proceder de una misma distribución (Normal, por ejemplo), pero de modo que un modelo genera residuos negativos donde el otro los genera positivos, y viceversa, y no querríamos decir que, en tal caso, ambos modelos ajustan los datos de igual manera.

El segundo aspecto crítico en el que no queremos incurrir es la excesiva reducción de la información muestral a que antes nos referimos, que surge cuando colapsamos dicha información en el valor numérico de un reducido conjunto de estadísticos que miden distintos aspectos del ajuste global del modelo. Por tanto, es necesario examinar los residuos de los modelos alternativos en toda la extensión muestral, quizá utilizando clasificaciones muestrales temporales o de otra índole, como ya se ha apuntado con anterioridad.

Utilizar estadísticos globales, es decir, calculados con toda la muestra puede ser una manera útil de resumir la información muestral, siempre que no basemos nuestras conclusiones exclusivamente en su utilización, sino que prestemos atención al ajuste de los modelos competidores a lo largo de toda la muestra. De hecho, puede ser interesante pensar en tales comparaciones globales. Por ejemplo, supongamos que tenemos dos modelos alternativos, uno con el vector  $X$  como variables explicativas, y otro añadiendo la variable  $Z$ , es decir, con  $(X,Z)$ . Para tener una primera idea, lamentablemente sólo global, de la similitud entre el ajuste proporcionado por ambos modelos, podemos calcular el coeficiente de correlación simple entre ambos conjuntos de residuos. Un coeficiente de 0,995 indicaría un ajuste muy similar, y sólo cabría analizar si la discrepancia entre ambos conjuntos de residuos está concentrada en un determinado rango de observaciones, o de valores de alguna de las variables del modelo, o representa una pequeña divergencia relativamente distribuida a lo largo de la muestra.<sup>7</sup>

Es sencillo construir distintos ejercicios de simulación al respecto; por supuesto, que la cantidad de situaciones que pueden presentarse es enorme, pero todo apunta a que una correlación lineal inferior a 0,95 indica que el ajuste proporcionado por ambos modelos no es idéntico. El umbral de 0,95 sólo puede considerarse indicativo, y generalmente será preferible ser conservador, utilizando un umbral incluso más elevado, puesto que es peor

cometer un error tipo II: *No rechazar la hipótesis nula: ambos modelos ajustan igual, cuando no es cierto*, que cometer un error tipo I: *Rechazar la igualdad de ajuste de ambos modelos cuando es cierta*. Si ambos modelos ajustan igual, como en el segundo caso, poco importa que rechacemos uno a favor del otro. En cambio, si ambos modelos (posiblemente una versión restringida uno del otro) ajustan los datos en distinta medida, no queremos quedarnos con el modelo restringido. En todo caso, conviene reiterar que la decisión última sobre los modelos no debe basarse en un criterio global de este tipo que proporciona, sin embargo, una información útil que debe tenerse en cuenta.

#### 4. CONCLUSIONES

Dejemos en este punto la exposición de deficiencias metodológicas en el trabajo empírico en Economía. El objeto de este trabajo era puntualizar la opinión del autor, siempre sujeta a debate, sobre tales posibles limitaciones. Pero mi intención ha sido no sólo plantear posibles deficiencias, sino lamentar que se refieran a aspectos que son objeto de discusión en los cursos habituales de Estadística y Econometría y, por tanto, no debieran suceder. En este sentido, he intentado comentar cada una de ellas desde el punto de vista de la teoría estadística y econométrica elemental y, como el lector habrá apreciado, sin recurrir a ningún aspecto teórico de dificultad.

Además de las cuestiones que he tratado, considero importante, como mencioné antes, el tratamiento de posibles variaciones paramétricas, o incluso funcionales, a lo largo de la muestra. Y me resulta especialmente preocupante la falta de rigor que me parece apreciar en la utilización, cada vez más generalizada, de métodos numéricos de estimación de modelos no lineales.<sup>8</sup> Pero cada uno de estos aspectos podría ser objeto de un largo debate.

Tampoco he tratado directamente la cuestión relativa al alcance del trabajo empírico en Economía, aun siendo el sustrato fundamental de todos los comentarios que he vertido. En respeto al lector que haya llegado hasta este punto he de manifestar, al menos, mi total creencia de que el carácter no experimental de la Economía impide de manera fundamental demostrar proposiciones con datos, y la actitud del investigador debiera ser siempre la de aproximarse a la evidencia muestral con la intención de evaluar si la información que contiene permite corroborar sus creencias a priori, si no lo hace o si, como por supuesto puede suceder, sugiere que alguna de tales creencias apriorísticas puede ser errónea. Nunca, sin embargo, podremos concluir que, efectivamente, dicha creencia es falsa, al ser rechazada por la información muestral, o que es cierta, por no haber sido rechazada.

La información muestral disponible debe utilizarse para, de modo formal o informal, actualizar el énfasis con que creemos que una determinada proposición o ciencia sea cierta. En consecuencia, siguiendo claramente un enfoque de tipo Bayesiano, la respuesta al interrogante que antes suscitó sería que, en presencia de  $k$  modelos alternativos, el investigador debe explicitar la verosimilitud relativa de cada uno de ellos, y actualizar dichas verosimilitudes mediante el uso de la información muestral. Ningún modelo es completamente cierto ni totalmente falso, y la información muestral es relevante, junto con las creencias previas al análisis de los datos, para evaluar la confianza del investigador en la validez de cada uno de ellos.

En definitiva, son muchas y muy fundamentales las limitaciones que en ocasiones impregnan la metodología de investigación aplicada en Economía, y no están relacionadas con la utilización de modelos simples de regresión, pues son cuestiones más básicas, que afloran en la interpretación de cualquier modelo estimado. Lo que es verdaderamente llamativo es que son aspectos que no surgirían con un buen nivel de comprensión de las materias de Estadística y Econometría de nuestras Licenciaturas o Grados.

## Notas

<sup>1</sup> Trabajo preparado para la conferencia del autor en la Facultad de Ciencias Económicas y Empresariales Universidad de Zaragoza, el 25 de febrero de 2009, invitado por el Instituto de Ciencias de la Educación.

<sup>2</sup> Desde hace unos años vengo preparando material docente a lo largo de estas líneas. Dicho material que, en muchos casos toma la forma de un caso práctico alrededor de un ejemplo de libro de texto, está disponible para cualquier lector interesado que lo solicite.

<sup>3</sup> No me refiero al número de datos. Precisamente, la utilización de medidas de cantidad y calidad de la información muestral es todavía una gran laguna en Economía aplicada.

<sup>4</sup> El lector interesado puede solicitar del autor documentación sobre ejemplos de aplicación de este enfoque.

<sup>5</sup> Mediante proyecciones lineales, por ejemplo.

<sup>6</sup> Si bien una posible respuesta pudiera ser que ambas tienen una capacidad explicativa o un contenido informativo similar.

<sup>7</sup> Otros enfoques son posibles, como el uso del test de signos: Si una variable  $Z$  no añade información a un vector  $X$  acerca del fenómeno en estudio, los residuos del modelo ampliado serán unas veces superiores y otras veces inferiores, en valor absoluto a los proporcionados por el modelo ampliado. Por tanto, el número de períodos en los que el residuo del modelo ampliado es inferior (¿quizá en valor absoluto?) sigue una distribución binomial  $B(T; 1/2)$  siendo  $T$  el tamaño muestral, y podemos llevar a cabo el contraste de la hipótesis nula: “ $Z$  no añade información sobre  $Y$  a la proporcionada por el vector  $X$ ” a través de dicha distribución, o de su aproximación  $N(T/2; T/4)$ .

<sup>8</sup> Que nos lleva en ocasiones a publicar como estimación numérica de un modelo no lineal el conjunto de valores numéricos donde el algoritmo numérico se ha detenido, sin que controlemos sus aspectos básicos: elección de condiciones iniciales, longitud de paso, utilización de primeras o segundas derivadas, y elección de niveles de tolerancia sobre los distintos criterios de parada preestablecidos

## Agradecimientos

El autor agradece la financiación recibida de los proyectos PROMETEO/2008/206, de la Comunidad Valenciana, y SEJ2006-14354, del Ministerio de Educación.

## Abstract

The first part of the paper contains some suggestions as to how results obtained in current research, together with the greater availability of computational means in Economics schools, should help to change substantially the contents and teaching methods of Statistics and Econometrics courses. The second part of the paper appeals to Basic Statistical concepts to point out some significant deficiencies in the standard use of Statistics in applied Economics research. A good understanding of these issues in undergraduate courses would avoid running into these problems.

**Key-words:** Econometrics, research methods, teaching, limitations.

**JEL Codes:** A2, C1.