



## Nombrado Doctor Honoris Causa en el acto del día de la Universidad del curso 92/93

### "Problemas actuales en econometría: una visión personal"

Hace 45 años cuando empecé a estudiar econometría seriamente estaba totalmente convencido de que la aplicación de la metodología matemática y estadística pronto nos permitiría entender mejor los sistemas económicos, y diseñar mejores políticas tanto privadas como públicas. Hoy, ante una economía mundial prácticamente paralizada, y en pleno marasmo en los países de Europa del Este y del mundo subdesarrollado, mi visión resulta claramente optimista. La experiencia nos enseña, incluso en econometría, que cuanto más estudiamos más nos queda por descubrir. Me voy ahora a referir brevemente a los problemas que debemos afrontar a la hora de modelizar una economía de manera útil para entender su funcionamiento y para elegir las políticas gubernamentales.

Existe un problema filosófico que la econometría comparte, en cierta medida, con el resto de las ciencias. Se trata del uso de modelos estocásticos para representar tanto la ignorancia del científico como una cantidad inmensa de detalles triviales. A un nivel más operacional se trataría de permitir a los modelos ajustarse a los datos del mundo real introduciendo errores de varios tipos que, en parte, representan los fallos de los recolectores de datos en su intento de que éstos representen a las variables económicas ideales de nuestras teorías y, en parte, representen a un número importante de factores que no pueden ser recogidos en nuestras ecuaciones sin introducir excesiva complejidad. En cierto sentido el comportamiento humano está determinado por causas no aleatorias. En muy contadas ocasiones se adoptan conscientemente decisiones aleatorias. Pero sabemos tan poco de los detalles de los procesos de decisión individual que tenemos que suponer que algún aspecto de la causalidad es aleatorio. Los modelos econométricos se hacen también difíciles de utilizar por el elevado número de agentes decisorios y de variables que cada uno de ellos controla. Cada modelo, para representar correctamente la verdadera complejidad de una economía, debe contener millones de variables. Incluso hoy en día en que los ordenadores tienen capacidad de miles de megabytes no podemos estimar modelos de esta dimensión. En primer lugar, porque no recolectamos datos con suficiente detalle, y en segundo lugar porque aunque dispongamos de series de datos de cuarenta años mensuales o trimestrales, resulta difícil modelizarlas sin tener en cuenta que el modelo cambia de manera fundamental durante el período. Esto es, el tamaño efectivo de la muestra está limitado, necesariamente, en parte por los cambios institucionales y en parte por los cambios discontinuos en las políticas. Esto nos diferencia de otras ciencias como las de la naturaleza, la meteorología, la astronomía, o la cosmología, en las que los modelos son grandes, con muchas variables, pero la muestra utilizada para estimarlos es aún mayor. En la práctica, los econométricos debemos recurrir a la agregación. Nuestros datos se miden en períodos largos, comparados con los retardos que se producen en la toma de decisiones. De hecho las restricciones impuestas por los datos disponibles, y las dificultades de cómputo que implica utilizar intervalos de tiempo realmente cortos, nos lleva a utilizar casi ineludiblemente datos mensuales o trimestrales. La dificultad de trabajar con innumerables variables micro nos fuerza a utilizar variables macro que en general son agregados más simples o números índices. Estoy utilizando los conceptos micro y macro, por conveniencia, como referencia a las variables originales de un modelo detallado y a las variables agregadas respectivamente. Pero la teoría de la agregación estocástica en modelos dinámicos no lineales impone condiciones muy restrictivas sobre el modelo para que la agregación sea perfecta. Incluso en modelos dinámicos lineales simples, sin restricciones a priori sobre los coeficientes, la agregación genera una estructura de retardos mucho más complicada que las ecuaciones micro correspondientes, y los errores de las ecuaciones tienen una especificación complicada de media móvil. Los modelos que estimamos son meras aproximaciones que sólo podemos esperar que funcionen mientras los datos permanezcan en un conjunto acotado, o la economía se mantenga en el entorno de una determinada senda. Los errores, tal como se presentan en las macroecuaciones, vienen determinados por complicadas funciones de las microvariables, por lo que tienen unas especificaciones estocásticas muy complicadas, cuya estimación sólo nos permite obtener una aproximación.

¿Qué importancia tiene esto para nuestra forma usual de estimar y testar? Si estimamos como si nuestro macromodelo estuviese verdaderamente generando las macrovariables, por ejemplo cuando modelizamos tratando a las variables agregadas como si fuesen bienes homogéneos, o tratando los índices de precios como si fuesen los precios correspondientes, o tratando cada bien como si hubiese sido producido por empresas perfectamente competitivas e idénticas...etc ¿podemos suponer que los estimadores maximoverosímiles de las macroecuaciones nos generarán ecuaciones de predicción que sean casi óptimas en algún sentido? La contestación, de sentido común, es que si nuestro

supuesto de que la agregación se lleva a cabo sobre agentes decisorios y bienes homogéneos es aproximadamente correcto, entonces podemos esperar que nuestras predicciones sean casi correctas. Pero en la práctica los agentes decisorios se diferencian en sus funciones de reacción y en sus retardos. Parece que si no imponemos demasiadas restricciones podemos obtener modelos que predigan razonablemente bien, simplemente en base al supuesto de que después de llevar a cabo las transformaciones adecuadas nuestras variables económicas pueden ser tratadas como si fuesen variables aleatorias estacionarias; y por lo tanto, que podemos utilizar técnicas de predicción óptimas basadas exclusivamente en la teoría estadística de la predicción en series temporales. Esta es la conclusión a la que llegan Liu T.C (1960) y Sims. C.A. (1980). En estas circunstancias no podemos esperar que nuestros modelos predigan bien si las variables se escapan del rango que cubrieron en el período de estimación muestral, y si predecimos haciendo uso de un modelo parcial y se produce un cambio en las ecuaciones que se requieren para cerrar el sistema. Por ejemplo, un cambio de gobierno o de política pueden aumentar el error de predicción de las ecuaciones que describen el sector privado de la economía. Indudablemente existe también alguna duda sobre la eficacia de utilizar tests del tipo de Neyman-Pearson cuando no creemos que nuestros modelos sean nada más que simples aproximaciones. Las dificultades son similares a la hora de aplicar métodos basados en la teoría de la decisión o métodos Bayesianos a la elección de la especificación del modelo.

Una vez que hemos contemplado la posibilidad de adoptar otro criterio alternativo que tenga en cuenta explícitamente la dificultad de creer en la validez exacta de cualquiera de nuestros modelos estocásticos, resulta claro que su elección sería enormemente arbitraria, de manera que siguiendo la tradición Popperiana parece que un procedimiento razonablemente robusto para problemas sencillos sería estimar y testar como si el modelo más general, que está incluido en el conjunto de modelos que estamos considerando, fuese un modelo estocástico válido, garantizando que los procesos estocásticos que generan los errores son suficientemente generales y sin restricciones. Para modelos grandes con muchos parámetros esto puede no funcionar muy bien, pero hay que reconocer que es muy difícil encontrar una alternativa. Yo encuentro poco atractivo el enfoque Bayesiano o el de la teoría de la decisión, ya que en general exigen especificar probabilidades apriorísticas para un conjunto grande de modelos alternativos, computar probabilidades a posteriori para nuestra muestra, y generar una predicción combinada basada en una función de pérdidas para los errores de predicción que es bastante arbitraria.

Ya he mencionado que la naturaleza de las variables omitidas cuando se tratan como errores aleatorios, y el efecto que tiene la agregación en el tiempo y de las variables, hacen necesario tratar los errores aleatorios como si estuviesen autocorrelacionados, y estimar y testar procesos estocásticos bastante complicados para explicar estos errores. Pero un problema más difícil es el tratamiento de la tendencia. La consideración de las tendencias como paseos aleatorios hace caso omiso del hecho de que para la mayoría de las tendencias económicas es posible determinar su signo a priori, y de que el mecanismo que lo garantiza es muy probablemente no lineal. Por ejemplo, un sector de la economía en el que el hecho de que la tendencia sea un paseo aleatorio es subjetivamente más probable es el del nivel de precios y el de las variables relacionadas que se miden en valores corrientes. En este sector tanto las instituciones públicas como las privadas toleran una tasa de inflación positiva y finita, y los costes tienen tendencia a ser rígidos a la baja. En sentido contrario actúa el progreso técnico, pero sus efectos primarios sólo son importantes en algunos sectores de la industria, e incluso en estos sectores pueden dar lugar a un aumento en los salarios, para que los trabajadores acepten la nueva tecnología, en vez de reducir los precios. Desde luego el progreso técnico es importante para determinar el nivel de producción, y sólo ocurrirá si aumenta la producción y conduce a una tendencia de producción creciente. Por lo tanto, debemos esperar que al menos algún componente de la tendencia sea no estocástico.

En los últimos diez años se ha llevado a cabo un trabajo verdaderamente interesante testando estas diferentes aproximaciones a la tendencia, y resulta claro que el poder de los tests depende mucho de las propiedades dinámicas de las variables económicas, pero que frecuentemente se requiere una gran muestra para poder discriminar entre las hipótesis de que la tendencia es lineal y no estocástica, o de que la tendencia es de tipo paseo aleatorio. Desde luego puede ser sensato suponer que los dos tipos de tendencia pueden estar presentes. Cuando la muestra es excesivamente pequeña para discriminar, algunos resultados de Monte Carlo para modelos simples indican que los sesgos en los coeficientes estimados de las variables económicas son razonablemente pequeños, comparados con sus errores estándar, cualquiera que sea el tratamiento que demos a la tendencia (aunque las proyecciones a largo plazo de los dos modelos sean diferentes).

El desarrollo del modelo de corrección de errores ofrece un compromiso interesante en el que algunas variables son estacionarias, independiente de una tendencia no estacionaria de orden bajo, y donde otras variables son estacionarias en primeras diferencias con tendencias que son paseos aleatorios. Voy a ilustrar esto refiriéndome brevemente a un modelo de inflación en términos de precios y salarios que construí a principios de los años sesenta. El modelo, en su versión más sencilla, contenía dos variables estacionarias, el salario real y el nivel de desempleo, y una tercera variable que era la primera diferencia de la tasa de variación de los precios monetarios. El desempleo sólo surgía en las ecuaciones como una variable retardada, y la ecuación que lo explicaba no se estimó. Los precios se explicaban por una simple ecuación homogénea de primer orden que incluía el precio de las importaciones como variable exógena. La ecuación de salarios era del tipo de corrección de errores, de marea que la primera diferencia en los salarios monetarios estaba relacionada con la diferencia entre los salarios reales corrientes y su nivel de equilibrio. La idea básica de esta ecuación era que los trabajadores tenían memoria y que su nivel de renta deseado estaba determinado por su experiencia reciente en los salarios reales. Si, por cualquier razón, su salario real caía, por ejemplo porque los precios estaban creciendo por encima de lo esperado, entonces tenían motivo para pretender un aumento mayor en sus salarios. Recordar los salarios reales recientes puede pensarse que tenga sentido en un plazo bastante largo; ya que es bastante frecuente que después de mantener bajos durante bastante tiempo algunos salarios los trabajadores exijan aumentos importantes basándose en que los salarios no han crecido a la velocidad que lo hicieron los precios al por menor. De hecho, este tipo de ecuación funcionó bastante bien durante un período bastante largo de la economía

británica.

En la medida en que esta pueda ser una explicación válida del funcionamiento del mercado laboral se refuerzan las ventajas de una economía en la que el salario real y la productividad crecen en paralelo, versus una en la que la renta real cae. En la economía a la que nos referimos no sólo el crecimiento en la renta conllevaría un ahorro creciente mediante retardos en el consumo, sino que los costes laborales serían menores en la medida en que los salarios reales reaccionasen con retraso. Esto contribuye a explicar el problema al que se enfrenta una economía excesivamente madura, como la británica, a la hora de generar un nivel de inversión adecuado, o para modernizar sus instituciones, como los sistemas educativos y de transporte, o el problema de la sobreexplotación de algunos recursos no renovables.

Un desarrollo deseable que hará menos necesario el uso de modelos agregados, y que de hecho hará posible el uso de modelos mucho más detallados y sencillamente testables y estimables, es el continuo crecimiento en el tamaño y la velocidad de los ordenadores; y en particular, el uso de la computación en paralelo. Cada vez se computarizan más transacciones y se dispone de más información sobre intercambios individuales a nivel micro. Disponemos cada vez de mayor número de encuestas muestrales sobre las expectativas de los hogares y las empresas. Al mismo tiempo vivimos tiempos interesantes para la economía mundial en el que las variables están tomando valores extremos en regiones inestables del espacio de variables, y esto debe significar que si la oferta creciente de datos económicos puede ser explotada adecuadamente obtendremos una comprensión mucho más profunda de la complejidad de las relaciones micro, posiblemente no lineales, con estructuras individuales de retardos. Podremos tratar las variables de política pública como endógenas, y en períodos más cortos de tiempo, siendo una fuente de mejora del poder predictivo de los modelos. De hecho, en los últimos años la comprensión de la estructura peculiar de los datos panel y en particular del problema de las variables endógenas/exógenas, de las estructuras peculiares de los errores, o de la presencia de outliers en estos modelos, ha mejorado significativamente, de manera que al acumular datos y construir modelos cada vez más complicados lograremos una comprensión mejor y una mayor capacidad de predicción.

Una fuente adicional de confianza es la alta calidad de los investigadores jóvenes en el campo de la econometría. En los años 70 parecía difícil atraer a jóvenes con buena formación estadística a este campo. Pero ahora resulta obvio cuando uno lee revistas como "Econometric Theory" o "The Journal of Econometrics" o "Econometrica" que se está produciendo trabajo original e interesante, de manera que según descubrimos nuevos problemas, surgen más investigadores de calidad para resolverlos. La consecuencia natural es que la calidad de los modelos aplicados también mejora. Los investigadores aplicados cada vez estarán menos inclinados a tratar la estimación como una estrategia de ajustar ecuaciones para encontrar fallos en sus modelos, y serán más conscientes de los problemas de "minería de datos" que surgen en la especificación de los modelos.

Existe una gran esperanza sobre el futuro de la econometría en el mundo, y del beneficio que aportará a la economía. En particular, en España existen nuevos trabajos en curso de gran interés. Estoy especialmente orgulloso de recibir un título honorífico de la Universidad Carlos III en la que tanto la estadística como la econometría están tan bien representadas. Gracias.