

DATOS, INCERTIDUMBRE Y LENGUAJE

Por el Académico de Número
Excmo. Sr. D. Alfonso Novales Cinca*

- “*Es fácil mentir con estadísticas, pero es difícil decir la verdad sin ellas*”, Andrejs Dunkels.
- “*In God we trust; all others must bring data*”, W. Edwards Deming.
- “*Without data, you’re just another person with an opinion*”, W. Edwards Deming.

INTRODUCCIÓN

Un reciente estudio de la Universidad de California en San Diego estimaba que nuestro cerebro recibe cada día 34 Gigabytes de información, una cantidad que puede colapsar un ordenador portátil estándar en un mes. El estudio estimaba que, a través de teléfonos móviles, Internet, correo electrónico, radio, periódicos, libros y redes sociales, recibimos unas 105 mil palabras cada día, o, lo que es lo mismo, 109 palabras por minuto cada una de las 16 horas en que se nos supone despiertos. La enorme cantidad de información que generamos y transmitimos es, sin duda, una de las características de nuestro tiempo que nos lleva a pensar que el Big Data, este tema cuyo contenido parece tan incomprensible para el común de los mortales, sea el área de mayor futuro profesional para nuestros jóvenes más brillantes.

No solo hemos creado la necesidad de operar con grandes bases de datos sino que, cada vez con mayor frecuencia, se requiere en comisiones de

* Sesión del día 12 de febrero de 2019.

investigación parlamentarias y sedes judiciales la participación de expertos que puedan aportar evidencia empírica acerca de la cuestión que se está debatiendo o juzgando. La información existe, lo que es más dudoso es que tengamos la capacidad de evaluarla correctamente y, de modo muy importante, si las evaluaciones de los expertos incorporan suficientemente una medida de la incertidumbre de los juicios que derivamos de la información estadística. En definitiva, *“la información, por sí sola, no es conocimiento”*¹.

Las carencias en la interpretación de la información estadística son evidentes en muchos campos; pero las noticias económicas son especialmente abundantes y son difundidas en los noticieros y medios de comunicación más populares. Reciben generalmente evaluaciones de signo opuesto de distintos comentaristas, sin que parezca posible arrojar luz de modo objetivo sobre el significado de un determinado dato, lo que siembra la confusión entre el público generalista y contribuye a devaluar el prestigio social de la tarea profesional de los economistas, sobre la que se hacen bromas del tipo: *“¿Para qué creó Dios a los economistas?”*; respuesta: *“Para que los pronósticos del tiempo nos pareciesen buenos”*, o: *“Los economistas han previsto 9 de las 5 últimas recesiones”*, o: *“Consulta a 3 economistas y tendrás 4 opiniones*. Podríamos añadir el comentario: *“La economía es el único campo en el que dos personas pueden obtener el premio Nobel por decir uno exactamente lo contrario del otro”*. He de decir, sin embargo, que lo último no me parece completamente sorprendente si pensamos que quienes reciben los premios Nobel han trabajado sobre muy diversas cuestiones. El premio se les otorga por sus contribuciones globales o porque desarrollan un determinado área, pero no reciben el premio por una opinión en concreto; en este caso se trataba de la eficiencia de los mercados financieros la cual, por otra parte, es una cuestión que siempre estará abierta a debate, entre otras cosas, por nuestra capacidad para discutir sobre conceptos generales sin concretar las fechas o los mercados a cuya eficiencia nos referimos o, lo que es más importante, la propia definición del concepto sobre el cual se debate.

Esta ponencia está estructurada en tres partes, asociadas a las tres razones que considero que contribuyen a depreciar el trabajo de los analistas económicos: la deficiente calidad de la difusión de noticias económicas en los medios de comunicación, la ausencia de evaluación a posteriori de las predicciones que publicamos los analistas, y la continuada manifestación por nuestra parte de afirmaciones que no estamos en condiciones de hacer. A las tres me referiré sucesivamente.

¹ Se atribuye a Albert Einstein haber afirmado que “la información no es conocimiento”

PARTE 1: LA DIFUSIÓN DE INFORMACIÓN ECONÓMICA

Diariamente recibimos noticias como estos tres titulares publicados el último año: “El IPC cierra el año en el 1,2% y resta poder de compra a pensionistas y funcionarios” (El País, 29/12/2017), “El IPC se modera al 1,1 % en abril por el abaratamiento de viajes y gas” (La Vanguardia, 27/4/2018), “La inflación crece al ritmo más bajo de los últimos 15 meses” [<https://blog.bankinter.com/economia/-/noticia/2018/2/15/analisis-espana-inflacion-enero-2018>]. Son expresiones a las que estamos acostumbrados, se refieren a un aspecto económico que nos concierne a todos, como es el comportamiento de los precios de consumo y si nos pidiesen que explicásemos su significado, todos sabríamos hacerlo. Y, sin embargo, cada uno de estos titulares afirma algo que es necesariamente falso. El IPC no puede cerrar el año en el 1,2%, o moderarse al 1,1% en un determinado mes; es la inflación quien lo hace. Por otra parte, es el índice de precios quien puede crecer al ritmo más bajo de los últimos meses, no la inflación.

Llevo años recogiendo titulares incorrectos de este tipo sin que el problema se haya aliviado lo más mínimo. Los actualizo periódicamente y nunca me falta material del que echar mano, y no únicamente en el tema del comportamiento de los precios. Como todos los índices, el IPC evoluciona con relativa suavidad, mientras que la tasa de inflación es el ritmo al cual varía el IPC, generalmente aumentando, o disminuyendo en algunas ocasiones recientes. Cuando se habla de la inflación, ni siquiera se resuelve la ambigüedad acerca de si se refiere al crecimiento de los precios de un mes sobre el anterior, la inflación mensual, o el crecimiento de los precios respecto del mismo mes del año anterior, la inflación anual. Aunque la inflación anual es, generalmente, claramente superior a la inflación mensual, en estos últimos años en los que los precios han mostrado una evolución muy moderada, ambas tasas han sido similares, generando cierta confusión en la difusión de noticias acerca del comportamiento de los precios. Recordemos que algunas cuestiones socialmente tan relevantes como la actualización de las pensiones han estado vinculadas al comportamiento de los precios de los bienes de consumo por lo que es importante saber si estamos hablando de algo que afecta a este mecanismo de actualización de pensiones o, por el contrario, a un comportamiento mensual que puede no tener trascendencia definitiva.

Otra fuente de problemas en la difusión de noticias económicas se refiere al aparente interés de los medios de comunicación por proporcionar titulares frente a un análisis riguroso de la realidad. Una estrategia que, lamentablemente, parece responder a que nuestra sociedad también demanda más el consumo de noticias en la forma de breves titulares y de opiniones emitidas por personajes mediáticos, sin solicitar de quien las emite una justificación de las mismas, ni un análisis argumental, ni una valoración de sus consecuencias. La lista de ejemplos puede ser interminable, pues prácticamente todo lo que se difunde por redes sociales, pero también por prensa y televisión, tiene este

carácter. En ocasiones se establece una comparación con un estándar estadístico de escasa justificación, como cuando se afirma que “El gasto social en España está por debajo de la media europea” (El País 7/12/2017) sin analizar si esto es necesariamente malo, como luego comentaré, o se afirma: “Después de duplicarse durante la década pasada, las emisiones de CO₂ crecen ahora a una tasa anual muy inferior”, pero no se dice si hemos sobrepasado algún umbral de seguridad, o si el volumen de emisiones se ha corregido por algún indicador cíclico, que parece que debería condicionar su interpretación. O la manipulación de la opinión que se hace cuando se afirma: “Se ha reducido en un 40% el tiempo máximo de espera para una determinada operación quirúrgica” ¿Cuántos pacientes se van a ver beneficiados por la reducción en el tiempo *máximo* de espera? ¿Significa algo esta reducción acerca de un menor tiempo de espera para el grueso de los pacientes? Porque bien pudiera ser que se hubieran reducido los tiempos máximos de espera sin incidencia práctica sobre el resto de los tiempos de espera, o incluso podría haber ocurrido junto con un aumento en los tiempos de espera más cortos². O cuando se difundió en la prensa que una investigación del Instituto Karolinska de Suecia mostraba que: “Se duplica el riesgo de desarrollar un neuroma acústico por uso habitual de teléfono móvil”; pero no se comentaba que la incidencia del neuroma acústico se situaba en una de cada 100.000 personas (un 0,001%), por lo que no parece que dicho incremento hubiera de tener la repercusión social que tuvo en su época en contra del uso continuado del teléfono móvil.

Otros equívocos en los que no entraré aquí surgen por no aclarar suficientemente el concepto al cual se refiere una determinada afirmación: si al paro observado en la Encuesta de Población Activa o al paro registrado, al número de ocupados o al de afiliados a la Seguridad Social, a la pobreza absoluta o a la pobreza relativa, a la desigualdad de renta o de riqueza, etc.. Y aún más especialmente, cuando se abusa de lenguaje para debatir sobre si “el actual sistema de pensiones es viable”, cuando la verdadera cuestión no es esa, sino si es posible y razonable detraer de los generadores de renta los recursos que serían precisos para financiar en el futuro a los perceptores de pensiones bajo el actual sistema.

Con estos ejemplos se puede comprender asimismo la facilidad con que partes interesadas pueden difundir información, incluso siendo cierta, para conseguir determinadas reacciones en la sociedad.

Las deficiencias en la difusión de noticias en términos estadísticos se producen en ocasiones por una incorrecta interpretación de conceptos. Describiré cuatro casos especialmente llamativos. Uno se refiere al habitual recurso a

² Imaginemos, aun siendo disparatado, que los quirófanos que se utilizaban para las operaciones de menor espera se acondicionan para ser utilizados exclusivamente en las operaciones que venían requiriendo mayores tiempos de espera.

la media de una variable con una acepción de valor “normal”, “razonable”, “ordinario”, o “aplicable a la mayoría de empresas, familias o países”. Como antes mencioné, incluso se presenta en ocasiones la media como un nivel “óptimo” y “deseable”. La interpretación de la media como valor “representativo” de una variable se deriva de la tradicional utilización, desde los inicios de la Estadística descriptiva, de la distribución normal³ para representar las frecuencias con que se observan los distintos valores numéricos que puede tomar una determinada variable, como el peso, la altura o el coeficiente de inteligencia de un conjunto de personas, la renta de las familias de un país, etc.. La distribución normal es simétrica alrededor de su media, lo que significa que habrá tantas observaciones dentro de una determinada distancia por encima de la media, como por debajo de la media. En particular, si la renta de las familias siguiese una distribución simétrica en una determinada población, habría tantas familias con renta superior a la renta media como personas con renta inferior a la misma, característica que define a la renta mediana. Bajo una distribución normal la renta media no solo coincidiría con la renta mediana sino que sería, además, la renta observada más frecuentemente; si hubiera que dar un valor representativo de toda la población, hay poca duda acerca de que la renta media sería la más indicada.

Pero muchas variables observables en ciencias sociales y, concretamente, muchas variables económicas importantes siguen distribuciones asimétricas: ya sea la renta, el ahorro y el gasto de las familias, los años de escolarización, o el tiempo de espera para una operación quirúrgica. En ellas, no hay igual número de observaciones por encima y por debajo de la media, propiedad que corresponde a la mediana. En distribuciones asimétricas, media y mediana pueden diferir mucho, y ninguna de ellas será, generalmente, el valor más frecuente. Por ejemplo, la renta familiar mediana en Estados Unidos era en 2014 un 25% inferior a la renta media; en tales situaciones, la mediana, y no la media, será más representativa.

En todo caso, incluso en distribuciones simétricas, la renta media no proporciona información alguna acerca de la lejanía con que las familias con rentas altas o bajas se encuentran con respecto de ella. De hecho, la dispersión de valores observados alrededor de la media pueden hacer que indicadores como la media o la mediana sean poco informativos: al cerrar 2018 los medios de comunicación recogían que la rentabilidad anual del inversor en la Bolsa española había sido de -15% porque ese fue el descenso del Ibex en 2018. Pero el Ibex no es sino un promedio especializado de los 35 valores más representativos del mercado cuyas rentabilidades anuales oscilaron entre una caída de -41%

³ La Estadística descriptiva es el conjunto de métodos utilizados para describir las principales características de una variable observable la distribución normal se conoce también como distribución gaussiana en honor del matemático Carl Friedrich Gauss, aunque fue introducida por el matemático francés Abraham de Moivre en su libro *The Doctrine of Chances* en 1738. Algunos historiadores de la ciencia creen que Gauss pudo haber descubierto la distribución normal de modo independiente unos años después.

y una subida del 15%, por lo que la rentabilidad del inversor habrá dependido de las acciones que haya mantenido en su cartera. De modo similar, la tasa de paro nacional a finales de 2018 era de 14,5%, pero varía desde un 9,7% en Cantabria hasta más de un 23% en Extremadura y en las ciudades autónomas de Ceuta y Melilla. Es imposible dar esa información con un indicador central como la media o la mediana; para ello es necesario utilizar alguna de las medidas que se conocen como medidas de dispersión. Por mucho que queramos simplificar las cosas, no podemos proporcionar información acerca de la distribución de frecuencias con que se observan los diferentes valores numéricos de una variable utilizando únicamente un valor representativo de la misma.

Un análisis riguroso de los datos disponibles generalmente requerirá un examen de toda la distribución de valores observados, ya se trate de los tiempos de espera para intervenciones quirúrgicas, la distribución de años de escolarización, o del rango de valores en el que se encuentra la renta de un determinado porcentaje de la población⁴.

La media no tiene ninguna connotación de optimalidad y sin embargo, nos referimos a ella como si fuera el objetivo deseable al comparar con el entorno europeo los niveles de cualquier indicador socioeconómico. Contaminamos más que la media europea, nuestro fracaso escolar es superior a la media europea, o la presión impositiva es en España inferior a la media europea, argumento que se utiliza en favor de elevar la presión fiscal, con objeto de aproximarnos a dicho promedio. No reparamos en que según fuéramos elevando nuestra presión fiscal, la media europea iría elevándose, con lo que otros países que tenían una presión fiscal superior a la media pasarán ahora a estar por debajo de la nueva media y elevarán su presión fiscal. Si todos los países involucrados siguieran este proceso, todos convergerían hacia el mayor valor de todos los observados inicialmente.

Además ¿por qué queremos estar en la media? ¿Quién ha dictaminado que la presión fiscal media europea actual es óptima? Y lo que seguramente es más importante: ¿son comparables las presiones fiscales de los diferentes países europeos? ¿son los países europeos suficientemente homogéneos como para tener la misma presión fiscal? Y lo mismo podría decirse de la mayoría de los indicadores para los que se establecen este tipo de comparaciones. Se introduce de este modo un comportamiento gregario que carece de justificación rigurosa, pero que puede tener serias consecuencias para la sociedad⁵.

⁴ Precisamente, una mayor sensibilidad en las últimas décadas por el análisis de toda la distribución de frecuencias, junto con la disponibilidad de datos, ha impulsado el desarrollo de análisis empíricos sobre desigualdad en la distribución de renta y la distribución de riqueza, un área de la mayor relevancia para el análisis del bienestar de una sociedad.

⁵ Es similar al que se produce entre las distintas agencias que contribuyen a un panel de predicciones, quienes, tentados por comparar sus previsiones con las de los restantes contribuyentes, tienden a aproximarlas a la media. Un promedio que será, en gran parte arbitrario.

El segundo caso de incorrecta interpretación de datos se refiere a la utilización de las tasas de variación, muy habitual en la difusión de datos y noticias económicas; por ejemplo, si se afirma que “la tasa de inflación registrada en febrero, del 0,2%, duplicó la registrada en igual mes del año anterior”. En primer lugar, porque duplicar una tasa mensual tan reducida significa únicamente que en ambos meses de febrero, la inflación fue reducida, y no cabe dar a la duplicación del aumento mensual de los precios de consumo una connotación peyorativa. En segundo lugar, y esto es algo sorprendente, porque las tasas de inflación se publican con un solo decimal, lo que es insuficiente en tasas potencialmente reducidas. Al hacerlo así, podríamos haber tenido tener en los dos meses de febrero citados, tasas de inflación de 0,14% y 0,16%, que son prácticamente indistinguibles pero que, por efecto del redondeo, se convertirían en tasas de 0,1% y 0,2%, una doble de la otra. Algo similar sucede con noticias de resultados empresariales, cuando se dice: “El beneficio de la empresa se duplicó el año pasado”, cuando ambos están situados en porcentajes insignificantes. Otra noticia confusa es del tipo: “la rentabilidad de la Bolsa española se ha duplicado hasta el 4%”, cuando se produce, por ejemplo, en una situación con tipos de interés nominales del 6%, por lo que mediante una inversión en activos con mucho menor riesgo que la inversión en Bolsa podría haberse obtenido una rentabilidad comparable.

Un problema al analizar tasas de variación de ratios surge porque, a menudo, es más importante el nivel del ratio que su variación. Si, por ejemplo, leemos que durante determinado periodo de tiempo “la tasa de morosidad ha aumentado un 50% hasta situarse en el 2,6%”, la primera conclusión que hemos de sacar es que estamos en una situación saneada, con una morosidad reducida. Evidentemente, deberemos preocuparnos acerca de si dicha evolución forma parte de una tendencia que puede llevarnos en un futuro cercano a una situación de riesgo insostenible. La segunda cuestión tiene que ver con dar la evaluación en términos de tasas porcentuales de crecimiento. Pensemos en lo extraño que resulta afirmar que “la implementación de la política monetaria ha estimulado un reducción del 50% en los tipos de interés hasta el 1%”.

Las tasas de variación de saldos presentan una dificultad adicional cuando se trata de saldos, como el déficit comercial o el déficit por cuenta corriente, o el beneficio empresarial, que pueden pasar de ser positivos a negativos ¿Cómo se calcula entonces la tasa de variación? Si un saldo pasa de valer 50 un año a ser -10 el año siguiente ¿cuál ha sido su tasa de variación? ¿Y si pasa de valer 3 a valer -3 el año siguiente?

Por último, se olvida con generalidad que una tasa de variación debe asignarse al punto central del periodo de tiempo que se ha utilizado en su cálculo. Así, si comparamos un dato del mes de febrero con el dato de febrero del año anterior, la tasa de variación resultante habría que asignarla al mes de agosto previo. Nos daría el ritmo al que la variable en cuestión estaba creciendo en dicho

mes, no el ritmo al que está creciendo en febrero. Si queremos estimar el ritmo de crecimiento actual, deberemos obtener previsiones para la variable hasta el mes de agosto próximo y comparar la previsión para dicho mes con el dato observado en agosto del año anterior (más sobre previsiones más adelante).

El tercer problema de interpretación suele surgir en algunas de las magnitudes macroeconómicas más relevantes, que tienen naturaleza de saldos, es decir, son la diferencia entre dos variables. Así, el paro es la diferencia entre la población activa y el número de ocupados; el déficit comercial, la diferencia entre ingresos por ventas y pagos por compras en el exterior; el déficit por cuenta corriente la diferencia entre ingresos y gastos de las distintas balanzas contables con el exterior; el déficit de caja es la diferencia entre ingresos y gastos del Estado; y los beneficios empresariales son la diferencia entre ingresos y gastos de la empresa. Un aumento en el paro se presenta en ocasiones como causado por una destrucción de empleo de igual calibre. Sin embargo, es claro que el paro puede aumentar incluso en un periodo en que se crea empleo, si en dicho periodo se produce un aumento de población activa superior a la creación neta de puestos de trabajo. Por el contrario, el paro puede reducirse incluso en un periodo en que se destruye empleo en términos netos, si se produce un descenso todavía mayor en la población activa, como en ocasiones sucede en periodos de escasa creación de trabajo, por el denominado efecto “desánimo” de los trabajadores. Es claro que, dependiendo de si el aumento del paro se produce en una u otra situación, su interpretación debiera ser muy diferente. De modo similar, se suele presentar inapropiadamente un descenso de beneficios como una elevación de costes, y un aumento del déficit de caja del Estado como un aumento de los gastos cuya corrección requiere un recorte quizá drástico de los mismos; un aumento del déficit comercial como una caída de las exportaciones, y un descenso en dicho déficit como un incremento de nuestra competitividad exterior que nos está permitiendo exportar más, etc

Esta confusión nunca debiera producirse. Las variables que tienen naturaleza de saldos o ratios no debieran analizarse por sí solas, sino a través de sus dos componentes, y el saldo o ratio resultante únicamente como consecuencia de ambos componentes, especialmente cuando la naturaleza y los determinantes de ambos pueden ser muy diferentes. La población activa puede estar afectada por la situación económica, pero principalmente por elementos sociológicos y demográficos y aspectos institucionales, calendarios escolares, períodos vacacionales en los que se contrata con carácter ocasional, etc.. Por el contrario, el número de ocupados en el sector privado responde a decisiones empresariales vinculadas a las expectativas de negocio, fundamentalmente. Las previsiones de paro debieran obtenerse como diferencia entre las previsiones de población activa y ocupados; la diferencia que a posteriori se observe entre el paro previsto y el paro realizado puede provenir de desviaciones en la población activa o en el empleo con respecto a las previsiones que habíamos realizado, o en ambas variables, y es instructivo realizar este análisis, más que

analizar únicamente la desviación entre el número de parados observado y la previsión con que se contaba⁶.

Una tercera situación en que se produce una deficiente interpretación de la información surge cuando esta se presenta en términos de probabilidades, a pesar de que, a diferencia de los análisis empíricos, la Teoría de la Probabilidad tiene leyes bien establecidas y no está sujeta a interpretaciones discutibles. Supongamos que se está considerando la aplicación generalizada de una prueba de detección de un tipo de células cancerígenas, con la intención de poder proporcionar un tratamiento en una fase inicial de la enfermedad, habiendo constatado que la presencia de dichas células cancerígenas afecta a uno de cada mil habitantes (un 0,1%), y sabiendo que la prueba da un resultado positivo, correcto, en un 99% de casos en presencia de células cancerígenas, y un 99% de resultados negativos, asimismo correctos, cuando las células cancerígenas no están presentes. Ambos porcentajes son elevados, por lo que parece que la prueba es muy eficaz. Cuando se pregunta a personas del ámbito médico acerca de la probabilidad de que una persona que da un resultado positivo en la prueba tenga células cancerígenas, la respuesta más habitual es: 99%. Tal respuesta se debe a la dificultad de distinguir entre la probabilidad de tener la enfermedad, condicional en que la prueba haya dado positivo, y la probabilidad de que la prueba arroje un resultado positivo, condicional en que exista la enfermedad (¡Qué lástima la escasa importancia que se da a los estudios de lógica en la enseñanza primaria y secundaria!). Un 99% es una sobreestimación excesiva de la presencia de la enfermedad; de hecho, el verdadero resultado es muy llamativo: con las cifras anteriores, la probabilidad de que una persona que arroja un resultado positivo en la prueba tenga realmente células cancerígenas es tan solo del 9%, mientras que con probabilidad 91%, no tiene cáncer⁷. Por tanto, la comunicación del resultado y la posterior aplicación del tratamiento serían improcedentes en 9 de cada 10 casos⁸.

Una mayor eficacia de la prueba en personas sanas incrementa la probabilidad de que un resultado positivo corresponda realmente a una persona enferma, pero necesitaríamos casi una total ausencia de error en la aplicación

⁶ Algo similar puede decirse acerca de ratios como la tasa de paro, que es el cociente entre el número de parados y el número de activos, y que solo deberían interpretarse a través de la evolución seguida por las dos variables de las cuales se obtienen.

⁷ El resultado es tan sorprendente que puede parecer increíble. Se debe al teorema de Bayes, pero también puede obtenerse la solución mediante un sencillo cálculo. Supongamos que el test se aplica a 200.000 personas. Dada la incidencia de la enfermedad, 200 tendrán células tumorales y 199.800 no las tendrán. Entre las primeras, 198 (200 por 0,99) darán resultado positivo en el test, y 2 darán resultado negativo. Entre las personas sanas, 197.802 (199.800 por 0,99) darán resultado negativo, y 1998 darán resultado positivo. Por tanto, un total de 2.196 personas darán resultado positivo, de las cuales solo 198 tendrán realmente la enfermedad, un porcentaje de $198/2.196=9,01\%$.

⁸ Quienes dan positivo no tienen ni mucho menos seguridad de tener cáncer, si bien resulta ahora más probable que antes de tomar el test (9% frente a 0,1%). El test da bastante tranquilidad a quienes dan negativo, que tienen una probabilidad prácticamente nula de tener cáncer (0,00001 o 0,001%), muy inferior a la que tenían antes de pasar el test, que era de 0,1%.

de la prueba a personas sanas para generar suficiente confianza en la prueba médica. Por otra parte, la eficacia en personas enfermas apenas afecta a los resultados. La reducida probabilidad de que un resultado positivo se deba a la presencia de células cancerígenas se debe fundamentalmente a la escasa incidencia de la enfermedad en la población analizada. De hecho si la incidencia no fuese de una entre mil, sino de una entre cien personas, la probabilidad de que una persona que da positivo tuviera células tumorales sería mucho más elevada, 50%. Pero la incidencia del tumor no es controlable, y tan solo podríamos intentar mejorar la eficacia del test en ausencia del tumor, con el limitado resultado ya mencionado. Seguramente la conclusión razonable es que con una incidencia tan escasa de la enfermedad, no procede aplicar este test dado que los costes del mismo y quizá de la aplicación de un tratamiento, junto con los costes sociales y personales de creerse enfermo erróneamente pueden ser enormes.

El denominado procedimiento de *credit scoring* es utilizado por las entidades de crédito para decidir si atender a una determinada solicitud; se trata de un sistema estadístico para estimar la probabilidad de que se produzca un posible impago en caso de ser concedido el crédito solicitado. Supongamos que la entidad dispone de un método que identifica correctamente el 80% de los casos de morosidad y el 80% de los casos en que no se producen dificultades en el servicio del préstamo, cuando se aplica a una muestra histórica de solicitantes, cuyos créditos ya han vencido. Supongamos asimismo que el porcentaje de solicitudes de baja calidad, que darán lugar a retrasos en los pagos, se estima en un 20%. Estas cifras parecen bastante razonables, pero implican que la mitad de los solicitantes que son detectados por el procedimiento de *credit scoring* como dudosos cuando solicitan su crédito, en realidad no lo son. El procedimiento conduce, por tanto, a perder un apreciable volumen de concesiones de crédito de buena calidad.

Si el modelo de *credit scoring* identifica correctamente el 90% de los casos de morosidad y el 90% de los casos en que no se producen dificultades en el servicio del préstamo, y el porcentaje de solicitudes de baja calidad, que darán lugar a retrasos en los pagos, se estima en un 10%, entonces la mitad de los solicitantes que son detectados por el procedimiento de *credit scoring* como dudosos cuando solicitan su crédito, de nuevo no lo serían. Pero si, en esta situación, el procedimiento de *credit scoring* identificase correctamente a los buenos clientes en un 80% de los casos, con un 20% de falsas identificaciones como potenciales clientes morosos, entonces, tan solo una tercera parte de los solicitantes identificados como potencialmente morosos lo serían. También podríamos aplicar las probabilidades descritas a una situación de detección de fraude fiscal, en el que el método de detección en vigor se ajustase a las probabilidades mencionadas y un 10% de las declaraciones incorporasen algún tipo de fraude. Tendríamos que aproximadamente la mitad de las declaraciones llamadas a ser revisadas no tendrían, en realidad, ningún defecto de cumpli-

miento de las normas tributarias. Como podemos ver, es importante no dejarse llevar de la primera impresión causada por probabilidades como las mencionadas, y efectuar el cálculo completo, pues a veces la situación real difiere bastante de la primera impresión recibida.

En un tono más lúdico, imaginemos que Rafael Nadal gana el primer set en un torneo de tenis al mejor de tres sets, y el locutor hace una afirmación del tipo: “en el 80% de los partidos en que Nadal gana el primer set, termina ganando el partido”. En realidad, este tipo de comentarios se escuchan con frecuencia en las retransmisiones televisadas de estos partidos. Pues bien, es muy sencillo comprobar que si los dos tenistas son de la misma calidad y si suponemos que el resultado de un set no influye en el del siguiente, la probabilidad de que quien gana el primer set gane un partido a tres sets es de $\frac{3}{4}$. Si quien gana el primer set es el tenista de más nivel, dicha probabilidad será aún más elevada. Por tanto, el comentario del locutor no aporta ninguna información y de hecho, quizá la probabilidad observada que nos anuncia sea inferior a la que deberíamos esperar. Por cierto, bajo los supuestos mencionados, la probabilidad de que el tenista que gana el primer set gane un partido a 5 sets es de 68,7%, y si gana los dos primeros sets, dicha probabilidad se eleva a 87,5%.

PARTE 2: AUSENCIA DE EVALUACIÓN

La mala calidad en la difusión de la información confunde a quien la recibe, y contribuye a crear la impresión de que es prácticamente imposible tener una idea clara de lo que ocurre en las economías reales. Puede generar también la impresión de que tal confusión es bien recibida por los economistas, con objeto de ocultar nuestra incapacidad para entender los fenómenos sobre los que deberíamos proporcionar un análisis riguroso de las cuestiones que preocupan a la sociedad. En dos aspectos es esto especialmente visible: por un lado, en las opiniones acerca de los mercados financieros. Al comienzo de cada año, así como a la vuelta de las vacaciones de verano, los medios de comunicación suelen dedicar espacio a recoger las opiniones de gestores de bolsa acerca de lo que cabe esperar en el transcurso del año. Siempre me ha asombrado su disponibilidad a prestarse a ello, pero lo cierto es que cuando se recogen tales opiniones, el resultado es un abanico de escenarios de una enorme disparidad, lo cual ya es en sí mismo fuente de confusión para el lector ¿Cómo es posible que supuestos buenos profesionales tengan opinión tan dispar acerca de lo que cabe esperar durante el año? Tampoco es evidente por qué hayan de recabarse opiniones en relación con un escenario concreto como el final de año ¿no es relevante lo que sucede antes o después? Lo peor, sin embargo, es que no recuerdo haber visto que, pasado ese tiempo, quienes dieron opinión hayan proporcionado un análisis a posteriori acerca de por qué se cumplió o por qué no se cumplió su escenario. Claro que hacerlo requeriría elaborar un

análisis de los elementos que condicionan la evolución de los mercados y esto requiere un esfuerzo mayor, así como disponer de cierto criterio para explicar los mercados, y también de la capacidad de analizar rigurosamente algunos datos históricos.

La evaluación a posteriori también se echa en falta en la predicción de las grandes variables macroeconómicas, como la tasa de crecimiento de la economía, la creación de empleo, la evolución del paro, el déficit público, la inflación, etc.. Un buen número de instituciones emiten informes al respecto, pero raramente se detienen a evaluar a posteriori el grado de cumplimiento de las mismas, así como las causas que hayan podido generar la desviación entre la predicción emitida y el dato observado posteriormente. En ausencia de tal diagnóstico, la proliferación de noticias de este tipo es tan enorme que nadie puede recordar las últimas predicciones publicadas, ni formar una opinión al respecto, asistiendo a este espectáculo como algo que aun concerniéndonos a todos, nos resulta ajeno por completo.

Parece que no se entiende que las predicciones económicas van a ser refutadas por la realidad, y que tan importante como la predicción es el seguimiento de la predicción. Las predicciones deben hacerse con el objeto de conocer mejor el fenómeno que se analiza, sea éste el paro, la inflación, el empleo, el saldo comercial con el exterior o, el PIB. Lo mismo sucede cuando se evalúa cuál pueda ser el impacto de una determinada política económica sobre el objetivo que se persigue, como luego describiré. Se quiere lograr dicho objetivo del modo más aproximado, pero se sabe que no se logrará exactamente, y conviene investigar las causas del desajuste que pueda observarse con objeto de afinar mejor el diseño de la política económica para próximas ocasiones. En consecuencia, analizar las razones por las cuales las previsiones se desvían de los datos observados con posterioridad es muy instructivo para mejorar el modelo explicativo del fenómeno que analizamos. Sin este ejercicio perdemos una importantísima fuente de información y no deja de resultar sorprendente que no se aporte tal información a los destinatarios de los informes de predicción. Además, sin este tipo de evaluación, cualquiera puede opinar acerca del futuro de la economía sin perder credibilidad, porque nadie nunca va a volver a considerar dichas opiniones.

El hecho es que, desafortunadamente, tampoco parece haberse implantado en la comunidad científica un procedimiento riguroso para evaluar un nuevo dato económico. Quizá lo que parece más natural es lo que raras veces se hace: todo dato nuevo debería evaluarse en relación con la predicción que anteriormente se haya preparado. Al comparar un nuevo dato con la predicción disponible, un aumento en el número de parados puede considerarse “bueno” si la predicción con la que contábamos anticipaba un aumento todavía superior al observado; un descenso reducido del paro constituye un dato “bueno” si la predicción sugería un aumento del mismo. La razonable estrategia de comparar

con una previsión obtenida con anterioridad hace que distintos analistas puedan discrepar en la evaluación del nuevo dato por haber obtenido predicciones relativamente dispares a partir de sus modelos, y es en estos términos en los que debería establecerse la discusión, y no en debatir sobre si el nuevo dato es “bueno” o “malo” sin hacer referencia a ninguna predicción, lo cual es difícilmente aceptable, si bien todavía lo más frecuente.

Una vez hecha esta evaluación, procederá analizar si el nuevo dato aporta información nueva; esta es quizá la cuestión más importante. Para ello, se debe incorporar el nuevo dato a los datos previamente disponibles y actualizar las predicciones a medio plazo; por ejemplo, a fin de año, o a fin del año siguiente, si estamos próximos al final del año actual. Especialmente cuando se trata de variables como la inflación, el empleo o el crecimiento económico, cuyos objetivos a medio plazo pueden estar recogidos en la ley de Presupuestos o en un Plan de crecimiento posiblemente establecido con anterioridad, y cuyo cumplimiento o incumplimiento es importante supervisar.

Además, no deberíamos prestar una excesiva atención a los datos que se publican frecuentemente, sino más bien a la tendencia que siguen los indicadores a que se refieren, alertando de los cambios que se anticipen en dichas tendencias, y tratando de explicar sus causas y sus implicaciones.

PARTE 3: ¿ESTÁN JUSTIFICADAS LAS AFIRMACIONES QUE HACEMOS?

Las herramientas de la economía son los modelos económicos y el análisis de datos. Un modelo establece un conjunto de relaciones de diverso carácter entre variables, recogiendo un determinado punto de vista teórico acerca del modo en que los agentes económicos toman sus decisiones, así como del diseño adecuado que debe adoptar la política económica en sus distintas áreas. Algunas relaciones pretenden explicar el comportamiento de una variable a partir de sus posibles determinantes; así, una relación puede explicar el número de ocupados en España en función de los salarios y de la actividad productiva; o explicar la evolución de los precios de consumo en función del crecimiento de la cantidad de dinero en circulación y de la renta. Una parte esencial de tales relaciones es el impacto unitario que cambios en una variable tienen sobre otras variables; por ejemplo, el efecto que sobre el empleo del colectivo afectado tiene un incremento de un 1% en el salario mínimo. Otras relaciones son identidades que reflejan la definición de una determinada variable (por ejemplo, la igualdad entre el PIB y la suma de sus componentes: consumo, inversión, sector exterior, o aspectos tecnológicos, como la capacidad de una economía para generar bienes a partir de los factores productivos. La estructura de un modelo puede ser compleja, dada la elevada interacción entre variables: los precios dependen de la demanda de consumo, pero también ésta dependerá de los precios de dichos bienes; el número de ocupados depende

de la actividad productiva, pero también ésta dependerá del número de ocupados. La necesidad de recoger en un modelo tales interacciones incorporando supuestos acerca de la estructura de la economía y del comportamiento de los agentes económicos es una característica del análisis económico, a diferencia de las ciencias experimentales, donde las relaciones de causa y efecto son mucho mejor conocidas, por estar controladas.

Es bien conocida la dificultad de contrastar empíricamente la existencia de relaciones causales entre variables, tema sobre el que existe una enorme literatura en distintas áreas científicas, y sobre el que nos ilustran Jaime Terceiro (“Causalidad en las Ciencias Sociales”, 2015) y Juan Arana (“Los sótanos del universo”, 2012, Biblioteca Nueva, ed. Siglo XXI), entre otros. Pues bien, un modelo impone relaciones causales, unidireccionales o bidireccionales, o la ausencia de las mismas, basándose en una determinada visión conceptual del fenómeno en estudio, y deriva sus implicaciones de modo condicional en tales supuestos.

Un modelo bien especificado tiene solución, la cual nos permite explicar la evolución de un grupo de variables, denominadas endógenas, que reflejarán generalmente algún aspecto del comportamiento de los agentes económicos: empresas y familias; son las variables cuya evolución queremos explicar con el modelo. Las variables exógenas son aquellas que utilizamos como factores explicativos y el modelo las toma como determinadas desde fuera del mismo; este sería el caso del precio del barril del petróleo, en el caso de un modelo de la economía española. Los instrumentos de política económica, como los tipos impositivos, determinados tipos de interés, o el crecimiento monetario, aparecerán en un modelo como variables exógenas, mientras que el objetivo buscado, sea la tasa de inflación, la creación de empleo o el crecimiento del PIB, aparecerán como variables endógenas⁹. Esta clasificación de variables guarda relación con la estructura básica de un problema de política económica, que requiere un objetivo y unos instrumentos, los cuales deben estar bajo el control de la autoridad económica.

La política económica responde generalmente a un problema de optimización; en dicho problema se pretende optimizar una función objetivo actuando sobre unos instrumentos y estando sujeto a unas determinadas restricciones, que recogen el modo en que la intervención sobre los instrumentos se transmite al objetivo buscado. El modelo económico es, precisamente, la especificación de dichas restricciones. La solución de dicho problema determina el nivel que debe tomar el instrumento para alcanzar el valor óptimo del objetivo, y proporciona asimismo el nivel óptimo que es posible alcanzar del objetivo deseado, una vez puesta en práctica dicha política. En realidad, el proceso lógico de buscar el valor óptimo de una función objetivo, que podemos querer

⁹ También puede tratarse de una política de precios por parte de una empresa con poder de mercado.

maximizar o minimizar, según su naturaleza, estando sujetos a unas restricciones, y actuando sobre unos instrumentos bajo nuestro control subyace a la toma de decisiones que continuamente tomamos en nuestra vida diaria. En otras ocasiones, la autoridad económica busca un determinado valor numérico del objetivo, más que un valor óptimo; por ejemplo, una tasa de inflación del 2%. En tales casos, no se trata de resolver un problema de optimización, sino de resolver el propio modelo económico, forzando en el mismo el valor deseado de la variable objetivo y encontrando los valores numéricos de los instrumentos que permiten alcanzarlo. Generalmente, puede haber más de una solución, es decir, puede haber distintas combinaciones de los instrumentos que permitan alcanzar el valor deseado de la variable objetivo; ello dependerá del número de objetivos y de instrumentos, como fue estudiado por J. Tinbergen (*On the Theory of Economic Policy*, North Holland, 1952). Debe haber al menos tantos instrumentos como objetivos; cuando el número de instrumentos es superior al número de objetivos, puede haber múltiples diseños de política económica que logren el mismo conjunto de objetivos.

Así, un ejercicio de la mayor importancia consiste en anticipar los efectos de la puesta en práctica de una determinada política económica; y esto requiere disponer de una estimación numérica del impacto unitario de las variables exógenas sobre las endógenas. Por tanto, los modelos económicos son de naturaleza cualitativa, pero requieren un tratamiento cuantitativo. A la autoridad monetaria no le basta con saber que elevando los tipos de interés puede conseguir reducir la inflación; necesita saber en cuánto ha de elevar los tipos de interés si quiere reducir la inflación en un punto porcentual, por ejemplo. Porque si la elevación de tipos de interés es insuficiente, habrá perturbado los mercados y habrá generado incertidumbre, sin lograr su objetivo; si la subida de tipos es excesiva, puede lograr sobradamente el objetivo perseguido, pero también crear una recesión. El éxito de la política económica dependerá de: a) que se tenga un buen control sobre los instrumentos, b) que el modelo utilizado en su diseño sea adecuado (básicamente, que recoja correctamente las influencias importantes de los instrumentos sobre el objetivo), c) que las estimaciones de los coeficientes de impacto de las variables exógenas sobre las endógenas sean cercanas a sus verdaderos valores, que son desconocidos, y d) que el contexto no varíe excesivamente durante la aplicación de la política, respecto del contexto que generó los datos que se han utilizado en la estimación del modelo y en el diseño de la política.

Dos analistas pueden diferir en sus opiniones acerca del efecto de una determinada medida de política económica simplemente porque mantengan en su cabeza modelos de la economía suficientemente distintos. Pero no tiene mucho sentido discutir sobre las diferencias de opinión acerca de un tema determinado, como habitualmente se hace en debates públicos, sin hacer referencia a los modelos que subyacen a cada posición. Tal debate no puede tener un final mínimamente interesante. Una discusión verdaderamente rigurosa acerca

de un tema debe centrarse sobre cuáles son los supuestos razonables, cuál es la estructura de relaciones entre variables, cuál es la magnitud de los impactos; en definitiva, sobre cuál es el modelo relevante. Sólo tiene sentido discutir sobre una determinada cuestión de política económica a través de los modelos considerados por cada analista, y evaluar dichos modelos a la luz de un seguimiento reiterado del cumplimiento de sus previsiones.

Siendo esta la metodología del análisis económico ¿estamos en condiciones de hacer estas afirmaciones?: “*Un incremento del 10% en el precio del petróleo generaría una reducción de medio punto porcentual del PIB español*”, o como escuchamos o leemos frecuentemente: “*El Fondo Monetario Internacional (o el Banco Mundial, o la OCDE) prevé un crecimiento del PIB en España del 2,0% para 2019*”. Para saberlo, hemos de conocer qué tipo de información nos proporciona un modelo económico, una vez estimado a partir de los datos disponibles. Para ello, tan importante es evaluar su capacidad explicativa sobre las variables endógenas, como caracterizar las propiedades del componente de dichas variables que no ha sido explicado por el modelo, lo cual es un problema puramente estadístico. A partir de esta información, un modelo puede utilizarse con fines predictivos, para lo cual se simula mediante técnicas de Monte Carlo, que consisten en generar posibles trayectorias temporales para las variables endógenas hasta el horizonte temporal al cual se quiera predecir.

Generando un número elevado de trayectorias, con un coste computacional generalmente reducido, tendremos una distribución de frecuencias para los valores numéricos de cada una de las variables endógenas del modelo (PIB, inflación, empleo), en el horizonte deseado. Se trata, por tanto, de obtener la predicción de la distribución de frecuencias que seguirá la variable objeto de estudio, en el horizonte temporal deseado para la predicción. Dicha distribución es la que nos permitirá obtener respuesta a cualquier suceso que se nos pueda plantear sobre dicha variable, en términos probabilísticos, como vamos a ver. Una vez obtenida dicha distribución, su mediana sería comparable con una predicción puntual, obtenida del modo habitual en los informes económicos. Tal comparación nos muestra la enorme simplificación que representa la predicción puntual, que sustituye toda la distribución de frecuencias por un único valor numérico. Bien es cierto que en el enfoque habitual también existe la predicción por intervalos, que proporciona un rango de valores numéricos dentro del cual se encontrará el valor futuro de la variable, intervalo que suele construirse al 90%, 95% o 99%; pero disponer de la distribución nos permite responder a estas preguntas y también a muchas otras. Este método de análisis por simulación, todavía escasamente utilizado en economía, es habitual en ciencias experimentales, meteorología, climatología, medio ambiente, biología.

Supongamos que hemos estimado un modelo con una única relación que explica la evolución temporal del PIB español utilizando como factores explicativos el precio del barril de petróleo, el tipo de cambio euro-dólar, la pre-

sión fiscal, etc¹⁰. Si no se quiere aportar información extramuestral, comenzaríamos estimando modelos predictivos para cada uno de los determinantes del PIB, con los cuales podríamos generar múltiples sendas para cada una de dichas variables. En este proceso es muy importante que las sendas generadas conserven las relaciones de dependencia entre sí que dichas variables han mantenido durante el periodo muestral. Estas son *predicciones incondicionales*. En segundo lugar, simularíamos trayectorias para el componente del PIB no explicado por el modelo utilizando sus características estadísticas, que previamente habremos estimado¹¹. Cada una de las trayectorias obtenidas para las variables exógenas y para el componente no explicado, llevadas al propio modelo, nos permiten deducir trayectorias para el PIB. Examinando los valores numéricos resultantes del PIB en el horizonte deseado, podremos pronunciarnos en términos probabilísticos acerca de cualquier evento que se considere relevante, utilizando una interpretación puramente frecuentista de la probabilidad: así, si hemos generado mil trayectorias hasta final de año, y en doscientas cincuenta de ellas el PIB ha crecido en durante más de un 2%, diremos que, la probabilidad de que el PIB crezca más de un 2% es del 25%. Pero podríamos responder a muchas otras cuestiones, como: ¿Cuál es la probabilidad de que la economía española crezca más de un 2%? ¿Cuál es la probabilidad de que crezca entre un 2,0% y un 2,5%?, ¿Y la probabilidad de que crezca más que el año anterior? o ¿En qué rango de tasas de crecimiento tenemos una confianza del 90%?

En su realización este ejercicio ilustra, de manera natural, que la dispersión de valores numéricos que el PIB alcanza a través de las distintas trayectorias simuladas aumenta y, con ello, la predicción está sujeta a mayor incertidumbre, si el horizonte temporal para la predicción es largo, o si la capacidad explicativa del modelo es reducida. También ilustra claramente que algunas variables son más difíciles de predecir que otras; por ejemplo, el rango en el que puedo predecir que se moverá el Ibex35 en 2019 con una probabilidad del 80% es tan amplio, que difícilmente puede juzgarse como informativo¹².

Una posibilidad alternativa, bastante más interesante, consiste en construir una *predicción por escenarios* o *predicción condicional*. En ella, las sendas futuras para las variables explicativas no se obtienen de modelos adicionales,

¹⁰ El modelo puede incluir también como variables explicativas valores en trimestres anteriores de algunos de estos determinantes, dado que sus efectos pueden no transmitirse al crecimiento económico inmediatamente, en el mismo trimestre. Asimismo, puede incluir entre las variables explicativas algún valor pasado de la propia variable endógena, cuya evolución se quiere explicar. Tendríamos en este último caso un modelo dinámico, cuya estimación por mínimos cuadrados no sería estadísticamente consistente si el componente no explicado por el modelo tiene autocorrelación, es decir, si está correlacionado consigo mismo a través del tiempo.

¹¹ O utilizando los valores numéricos que para este componente no explicado hayamos obtenido en el proceso de estimación del modelo a lo largo del periodo muestral, en técnicas que se conocen como bootstrapping.

¹² En todo caso, nunca debe utilizarse un modelo para predecir la evolución de variables endógenas cuando las variables exógenas toman valores fuera del rango observado en la muestra.

sino a partir de escenarios establecidos por el analista. Así, un *escenario* pudiera ser: “el precio del petróleo va a aumentar un 25% durante 2019, el tipo de cambio euro-dólar permanecerá estable y la presión fiscal se elevará un 5%”. Esto tiene especial sentido cuando se trata de predecir los efectos que sobre el objetivo buscado pudiera tener una determinada política económica pues, en ese caso, la autoridad económica correspondiente controla los instrumentos sobre los que va a actuar y, por tanto, puede considerar determinadas trayectorias para las mismas. Una vez que el escenario se haya reflejado en sendas para los instrumentos de política a lo largo del horizonte de previsión, dichas sendas, unidas a las obtenidas por simulación del componente del PIB no explicado por el modelo, permiten genera la distribución de frecuencias del PIB al horizonte deseado del mismo modo que antes he descrito.

Evidentemente, sería poco apropiado realizar este tipo de ejercicio para un solo escenario pues sugeriría que tenemos una enorme confianza en el mismo lo cual será, generalmente, difícilmente justificable. Precisamente, parte del valor añadido más enriquecedor del analista económico consiste en establecer justificadamente escenarios razonables, además de pronunciarse acerca de la verosimilitud relativa de cada uno de ellos. Tal verosimilitud se incorpora fácilmente en el procedimiento de simulación bajo los distintos escenarios considerados, obteniendo una distribución de frecuencias del PIB al horizonte deseado que incorpora la incertidumbre del analista acerca de los escenarios que haya propuesto sobre sus determinantes.

La incertidumbre acerca del verdadero valor numérico de los coeficientes de impacto puede incorporarse asimismo sin gran dificultad al proceso de simulación¹³.

En un nivel adicional, el analista reconocerá la posibilidad de contar con varios modelos alternativos, a cada uno de los cuales debe asignar una verosimilitud relativa, la cual puede asimismo ser incorporado al procedimiento de simulación sin dificultad. Existen procedimientos bien establecidos en estadística, especialmente en estadística bayesiana, para actualizar la verosimilitud asignada a cada modelo en función de los resultados predictivos que va generando. Lamentablemente, rara vez se tiene esta consideración en los ejercicios predictivos habituales en economía, sino que se busca un buen modelo, tan complejo como sea necesario, que permita justificar su uso exclusivo. No estoy seguro de que sea esta una buena estrategia, pudiendo ser más útil establecer varios modelos relativamente simples; la experiencia no muestra que un modelo muy sofisticado vaya a predecir mucho mejor.

¹³ Extrayendo aleatoriamente, para cada trayectoria, valores numéricos para dichos parámetros a partir de la distribución conjunta que para los mismos se haya caracterizado en el proceso de estimación del modelo.

De este modo se puede cuantificar la incertidumbre existente en el análisis predictivo acerca de cuatro aspectos: los valores futuros de los determinantes del fenómeno en estudio, el modelo que los relaciona entre sí y, dentro de cada modelo, los valores numéricos de los coeficientes de impacto; por último, la incertidumbre acerca de la evolución futura del componente no explicado por el modelo. Me he referido únicamente a predicciones expresadas en términos cuantitativos, cuyo grado de cumplimiento puede ser evaluado de un modo concreto. Es frecuente que diferentes observadores emitan previsiones no concretadas en elementos cuantitativos, del tipo de las emitidas por una oficina del Banco Mundial: “El próximo año el crecimiento global será más robusto”, o “La continuada mejoría en gobernanza y en renta en gran parte de África generará confianza y dinamismo en el continente”, que no consideraré en mi ponencia. En estos casos tendría que añadirle incertidumbre acerca no solo de la evaluación de la predicción una vez se conozca el resultado real del suceso previsto, sino la incertidumbre acerca de lo que la predicción realmente dice, lo cual no es tratable.

Curiosamente, es en el mundo financiero donde se ha popularizado en los últimos años, por mandato del comité de Basilea de medición de riesgos, la consideración explícita del denominado *riesgo de modelo*, consistente precisamente en establecer los modelos alternativos que se consideren razonables, para ser utilizados conjuntamente, bajo ponderaciones adecuadas, así como teniendo en cuenta la incertidumbre acerca del valor numérico de los coeficientes de impacto, en el análisis de riesgos de una entidad financiera. Tal análisis de riesgos es en gran medida un ejercicio de predicción; especialmente, un ejercicio de simulación al que pueden asociarse todos los comentarios que he vertido acerca de la elaboración de predicciones.

Para ilustrar estas ideas he realizado para esta ponencia predicciones del PIB a lo largo de 2019 mediante un sencillo modelo que utiliza únicamente información histórica del PIB, sin incorporar ningún determinante externo que, lógicamente, podría aportar información útil y mejorar las predicciones resultantes. La mediana de las trayectorias generadas apunta a un crecimiento del 2,07%, que podemos redondear, bajo la habitual práctica de usar un solo decimal, a 2,1%, levemente menos optimista que la predicción del Banco de España, que es de 2,2%, aunque a mi juicio no significativamente diferentes desde el punto de vista estadístico. Estimo que la probabilidad de que el crecimiento económico sea inferior en 2019 al de 2018 es de 62%; la probabilidad de crecer más de un 2% es de 52%. Con la información actualmente disponible la economía española crecerá entre un 1,88% y un 2,25% con una probabilidad de 90%. En este ejercicio he incorporado la incertidumbre acerca del modelo, utilizando dos modelos diferentes, ambos con la misma estructura econométrica, pero uno estimado únicamente con datos hasta el tercer trimestre de 2018, ya que el cuarto trimestre tiene aún una naturaleza provisional y otro, estimado incluyendo también el dato de dicho trimestre. También he incorporado la incertidumbre acerca del componente no explicado por el modelo, pero no la incertidumbre paramé-

trica. Por otra parte, en este modelo sencillo no procede tratar la incertidumbre acerca de los escenarios futuros de las variables explicativas.

Este tipo de análisis es muy diferente del que ha sido habitual hasta ahora y, sorprendentemente, no requiere técnicas estadísticas ni computacionales complejas. Afirmaciones categóricas del tipo: “La economía crecerá un 2,2% en 2019” sugieren que su autor “conoce” o tiene una confianza muy sólida en dicha tasa de crecimiento, lo cual es contrario a la naturaleza del análisis económico empírico. Ni las características de los datos ni la naturaleza de los modelos económicos justifican tal tipo de afirmaciones, que tan sólo deberían hacerse en términos probabilísticos como los que he descrito. Ni tampoco procede considerar como relevante una noticia como: “*El Banco Mundial reduce en una décima su previsión de crecimiento para España en 2019*”. Tales afirmaciones surgen al ignorar las cuatro fuentes de incertidumbre que he descrito y solo pueden entenderse como un excesivo resumen de la rica información que el analista podría haber generado por simulación de sus modelos. Al presentar tal resumen, omitimos una información que es crucial para evaluar la relevancia de la afirmación, cual es el grado de incertidumbre en dicha predicción así como la verosimilitud que asignamos a otras posibles tasas de crecimiento, y generamos una presunción de conocimiento que inevitablemente se verá más tarde refutada por la realidad, contribuyendo con ello a cuestionar de nuevo la tarea de los analistas económicos.

RESUMEN

Los estudiosos de las ciencias sociales tenemos un importante reto por delante. Los ciudadanos necesitan estudios rigurosos que puedan ayudarles a formar opinión sobre temas económicos y sociales de la mayor importancia, pero lo que reciben de manera continuada es un auténtico aluvión de informaciones sobre distintos aspectos que afectan a su vida diaria, como las pensiones o el paro, así como sobre otras variables cuya relevancia no se les explica, como la tasa de crecimiento de la economía mundial. Lamentablemente, raras veces está elaborada dicha información con suficiente rigor, a menudo incluso sin definir con precisión los conceptos utilizados; el problema está especialmente presente en Economía, pero afecta a todas las ciencias sociales. Parte del problema es que se transmite información con excesiva frecuencia, lo que hace que sea habitual que la evaluación que se hace de un determinado aspecto, como el mercado bursátil, la inflación, o el propio crecimiento económico tengan distinto signo en meses sucesivos.

Frecuentemente, quienes emiten su opinión, ni siquiera son profesionales acreditados de las materias sobre las cuales opinan, lo cual sucede con especial frecuencia entre quienes desempeñan tareas en la política y en los medios de comunicación. He expuesto comentarios críticos a la labor de estos últimos,

quienes deberían mejorar significativamente sus niveles de rigor al proporcionar información estadística, pero este colectivo no es responsable único del problema. Lo es en gran medida nuestra sociedad, que demanda información simple y rápida, sin deseos de profundizar en argumentaciones ni reflexiones profundas. Pero los primeros responsables somos los investigadores sociales, quienes estamos excesivamente ausentes, especialmente desde el mundo de la universidad, enfrascados en la persecución de objetivos específicos del mundo académico, definidos de modo endogámico, cuya rentabilidad social es a veces discutible. Los incentivos introducidos en la Universidad han servido frecuentemente para distraer la actividad de los investigadores hacia la elaboración de modelos sofisticados, cuyo beneficio es discutible, con poco énfasis en proporcionar a la sociedad elementos para formar opiniones rigurosas y bien fundamentadas acerca de los temas que le preocupan: pensiones, paro, empleo, rentabilidad de la educación, etc., o en analizar la calidad de los inputs que utilizamos en nuestro trabajo. Los investigadores somos quienes inexcusablemente hemos de buscar el máximo rigor en el uso de los conceptos y métodos estadísticos en nuestros análisis y en proporcionar una información rigurosa, lo que seguramente requiere una verbalización de conceptos y resultados muy diferente de la que actualmente utilizamos. Además, hemos de saber resolver la tensión entre la complejidad (que no dificultad) que casi inevitablemente va asociada con el rigor, y la necesidad de hacer que las conclusiones de nuestros estudios puedan ser transmitidas al gran público, para quienes hemos de ser comprensibles.

Las herramientas de que disponemos los economistas: modelos y análisis de datos, generan implicaciones de carácter probabilístico que siempre deberíamos evaluar a posteriori, no solo por deferencia a quienes las escucharon en su momento, sino por la utilidad de la información que proporciona el análisis detallado de las discrepancias entre previsiones y datos observados. Asimismo, es imprescindible proporcionar, junto con las conclusiones de nuestros estudios, especialmente en los análisis predictivos y en la evaluación de políticas económicas, una evaluación de la incertidumbre asociada a dichos resultados. Al no hacerlo así, facilitamos la trivialización del discurso económico y las críticas, en buena parte merecidas, a nuestro trabajo.

