

Verónica Gonzales-Stuva*
Pedro Rey-Biel*

REFLEXIONES SOBRE LA COLABORACIÓN ENTRE LAS INSTITUCIONES PÚBLICAS Y LOS INVESTIGADORES EN LA REALIZACIÓN DE INTERVENCIONES CONDUCTUALES

El movimiento hacia políticas basadas en evidencia implica adoptar también la investigación en ciencias del comportamiento, que ha contribuido a este enfoque. Tras aclarar que las intervenciones conductuales no se reducen al enfoque nudge, discutimos ventajas y dificultades de la colaboración entre Administraciones públicas y comunidad investigadora para implementar políticas que consideren cómo las interpretan los ciudadanos. Nos centramos en cuatro requisitos para intervenciones exitosas: una actitud favorable a la experimentación, disponibilidad pública de datos administrativos, diseño evaluable de las intervenciones y evaluación independiente y rigurosa de las políticas.

Reflections on the Collaboration between Public Institutions and Researchers in the Implementation of Behavioral Interventions

The shift toward evidence-based policymaking also entails embracing research from the behavioral sciences, which has significantly contributed to this approach. After clarifying that behavioral interventions go beyond the nudge framework, we discuss the advantages and challenges of collaboration between public administrations and the research community in implementing policies that account for how citizens interpret them. We focus on four key requirements for successful interventions: a favorable attitude toward experimentation, public availability of administrative data, evaluable intervention design, and independent, rigorous policy evaluation.

Palabras clave: economía del comportamiento, políticas basadas en la evidencia, nudges, incentivos, paternalismo libertario, manipulación, experimentación, sesgos de comportamiento.

Keywords: behavioral economics, evidence-based policy, nudges, incentives, libertarian paternalism, manipulation, experimentation, behavioral biases.

JEL: A1, A2, C9, D83, D9.

* ESADE. Universitat Ramon Llull.

Contacto: veronicaisabel.gonzales@esade.edu; pedro.rey@esade.edu

Versión de mayo de 2025.

<https://doi.org/10.32796/ice.2025.940.7941>

1. Introducción

El éxito de las políticas públicas depende crucialmente de la respuesta de los ciudadanos ante ellas. Es decir, de la comprensión, aceptación y grado de cumplimiento que hacen los ciudadanos de las medidas que se les aplican. Frente a un enfoque normativo, que asume que los individuos cumplen una ley o aprovechan los beneficios de una política pública por el mero hecho de promulgarse, existe un planteamiento más realista, pragmático y empiricista que facilita el cumplimiento por parte de la ciudadanía de las intenciones del regulador a través de entender los factores que afectan al comportamiento de los ciudadanos ante una política pública.

De forma similar, la investigación social ha experimentado una revolución empírica, en la que el análisis del comportamiento ha cobrado peso frente a los postulados teóricos sobre el mismo. En concreto, la investigación académica en economía, entendida como el estudio de las decisiones cuando los seres humanos nos enfrentamos a restricciones y a sus consecuencias, ha ampliado su foco de estudio más allá del restrictivo supuesto de que las personas son puramente racionales. La racionalidad en economía asume que los individuos son siempre capaces de evaluar las consecuencias de las alternativas entre las que eligen, que escogen siempre aquello que es mejor para ellos mismos y que no se dejan llevar por sus emociones. Por contra, el reciente auge de la economía del comportamiento tiene en cuenta los problemas que los seres humanos podemos tener para evaluar entre opciones, los posibles errores que nos llevan a no maximizar nuestro bienestar (presente o futuro) y cómo el contexto, las emociones y nuestros sentimientos nublan en ocasiones nuestro comportamiento más analítico. La economía del comportamiento ha basado una parte importante de su éxito en tomar como punto de partida lo que hacen los individuos realmente, en lugar de una idealización «teórica» y simplificada sobre cómo se comportan (o deberían comportarse), para, a partir de ahí, estudiar la toma de decisiones como un proceso complejo y potencialmente subóptimo, en el que

la influencia de los sesgos psicológicos puede ser crucial. La economía del comportamiento no se limita solo a estudiar la toma de decisiones no racionales, sino que diseña medidas, y prueba su efectividad, teniendo en cuenta esos sesgos psicológicos para ayudarnos a decidir y comportarnos «mejor».

Ante un contexto político crecientemente polarizado, donde la ideología impone cada vez más las políticas que se adoptan, se contrapone al movimiento que promueve el uso de evidencia científica como fundamento para el diseño y evaluación de las políticas públicas. Dicho movimiento refuerza la democracia frente al resurgir del populismo informando, a través del conocimiento científico riguroso, sobre qué políticas funcionan y por qué lo hacen. Es en este contexto de potenciar las políticas basadas en la evidencia donde debe entenderse que las buenas políticas públicas deben basarse también en lo que hemos aprendido a través de las ciencias del comportamiento para conseguir mejorar la efectividad de las medidas públicas e incrementar el bienestar de los ciudadanos. Para ello, se debe producir una revolución en el diseño de las políticas públicas que facilite la información y comprensión de las medidas tomadas, que tenga en cuenta las reacciones emocionales de los ciudadanos y su respuesta ante los estímulos psicológicos y los incentivos y que sea responsable del inmenso poder que otorga la capacidad de manipulación sobre el comportamiento que tiene quien diseña y comunica las políticas.

El auge de las políticas basadas en la evidencia depende crucialmente de las colaboraciones de las Administraciones públicas con quienes generan la evidencia científica que informa las políticas¹. Mientras que los organismos públicos aportan datos, opciones de políticas, y conocimiento sobre dificultades de implementación, la comunidad investigadora puede contribuir a impulsar esta colaboración mediante sus conocimientos

¹ Bonargent (2024) reporta que el codiseño y coimplementación de un proyecto de investigación entre investigadores y responsables políticos aumenta la probabilidad de adopción de una política de 17 a 20 puntos porcentuales.

metodológicos para establecer causalidad, su capacidad de análisis, su tiempo, su reputación y su independencia. No obstante, las Administraciones y los investigadores difieren en objetivos, en plazos temporales de ejecución y en la información de la que disponen.

Afortunadamente, la mejora en los sistemas de información ha permitido un crecimiento exponencial de la cantidad de información disponible para los decisores políticos y para los científicos sociales. No obstante, el casi monopolio que tienen los organismos públicos de los registros de datos administrativos genera problemas de acceso a dicha información por parte de la comunidad científica, que es quien puede proporcionar una mayor capacidad de análisis causal sobre la evidencia existente.

El uso de las ciencias del comportamiento será crucial en la necesaria modernización de las políticas públicas. Para poder realizar intervenciones conductuales exitosas que apoyen el desarrollo de políticas basadas en la evidencia, se deberían facilitar los siguientes requisitos: una actitud favorable a la experimentación y a probar qué es lo que funciona y por qué, disponibilidad pública de datos de los diferentes niveles de la Administración, un diseño de intervenciones públicas de forma que permitan realmente ser evaluadas y, finalmente, una evaluación independiente y rigurosa de las políticas públicas desarrolladas.

En este artículo discutimos cómo se pueden utilizar las ciencias del comportamiento para fortalecer la colaboración entre entidades públicas y el mundo académico en todas estas tareas, para mejorar la adopción de políticas basadas en la evidencia sobre cómo se comportan realmente los ciudadanos y cómo se puede influir en dicho comportamiento de manera que redunde en el bienestar de la sociedad en su conjunto.

2. Las Ciencias del Comportamiento van más allá de los *nudges*

La utilización de intervenciones conductuales en la política pública entraña una serie de problemas, que

podríamos resumir en problemas éticos y problemas de efectividad.

Desde el punto de vista de la ética, asumir que las Administraciones públicas saben mejor que los ciudadanos qué es lo mejor para la sociedad en su conjunto, plantea un debate moral profundo sobre los cimientos de nuestro sistema democrático. Una virtud del «paternalismo libertario» esgrimido por Thaler y Sunstein es que ocupa un punto intermedio entre dos maneras de enfocar las políticas públicas. Una es el dirigismo de la izquierda tradicional o del fordismo clásico (puede comprar un modelo T del color que quiera... siempre que sea negro). La otra es el *laissez faire* clásico, que supone que un gobierno que no interviene en absoluto preserva realmente la libertad del individuo y, si acaso, la misión del gobierno es aumentar el abanico de opciones de los consumidores, porque «el mercado es la mejor solución a cualquier problema». El paternalismo libertario equilibra la importancia de permitir que los ciudadanos puedan escoger su comportamiento con la de dirigirles a lo que un legislador, esperemos que «benevolente», cree que es el comportamiento más apropiado. Pero dicho equilibrio no exime de responsabilidad al político sobre la legitimidad de dirigir el comportamiento de la sociedad que le ha elegido. Hasta qué punto el mandato de las urnas supone un contrato del votante aceptando no solo las propuestas de políticas hechas en campaña sino también una aprobación más sutil del ejercicio de influenciar el comportamiento ciudadano, es un debate con implicaciones legales y filosóficas que excede la ambición de este artículo.

Desde el punto de vista de la efectividad, la pregunta no es ya tanto si se deben utilizar intervenciones conductuales, sino cuándo estas funcionan y si son la herramienta más eficiente en términos de coste-beneficio en cada caso para conseguir los objetivos que se pretenden. Como se ha destacado en otros artículos de este mismo número, podemos distinguir cuatro clases de intervenciones tipo *nudge*: i) provisión de información; ii) cambios en el entorno físico; iii) cambios en las opciones ofrecidas «por defecto»; y iv) el uso de las normas

sociales y de la relevancia de las opciones ofrecidas (*salience*). La mayor parte de los ejemplos de intervenciones conductuales que se conocen responde a uno de estos cuatro tipos de *nudge*.

Sin embargo, un error común consiste en asumir que la aplicación de las ciencias del comportamiento al diseño, implementación y evaluación de las políticas públicas se limita al uso extendido y popularizado de los *nudges*. Los *nudges* son solo uno de los mecanismos psicológicos que pueden inducir cambios de comportamiento en los individuos. Además, y esto es clave en su popularidad, cuando funcionan, los *nudges* producen resultados positivos a un coste muy bajo, o incluso sin coste. El éxito de la aproximación *nudge* radica parcialmente en que las intervenciones resultan (casi) «gratis». Esto tiene como manifestación negativa que el primigenio *Behavioral Insights Team* (BIT) fue usado por el Gobierno británico, en parte, para «blanquear» los recortes presupuestarios en muchas políticas sociales y para no hacer suficientes esfuerzos por atacar realmente las raíces profundas de importantes problemas sociales.

Un problema de la aproximación *nudge* es que no formula una teoría completa del comportamiento humano. Una cosa es hacer constar que existen una serie de errores de decisión predecibles, que se pueden sistematizar vagamente con la idea de los dos sistemas de procesamiento lento y rápido de Kahneman y otra que en cada situación concreta exista una serie de principios que ayuden a predecir qué va a pasar y cómo modificar el comportamiento. Relacionado con ello, la propia evidencia existente puede no ser tan sólida como algunas de las anécdotas sobre *nudges* más conocidas y llamativas parecen sugerir. El éxito de las intervenciones conductuales depende de mil detalles que tienen que ver con el contexto en el que se promueve y desarrolla una política pública.

Para dirigirnos hacia la modernización de la Administración pública y al uso de políticas públicas basadas en la evidencia se debe abrazar de forma más ambiciosa el enfoque de las ciencias del comportamiento

y no limitarse al uso de *nudges*, sino incluir otros instrumentos complementarios y sustitutivos, como los incentivos, que tengan en cuenta la diversidad en los procesos de toma de decisiones de los individuos y el amplio espectro de motivaciones que las generan. Un incentivo supone una razón adicional para comportarse de una determinada manera, y debe apelar a cualquiera de las múltiples motivaciones, no solo monetarias, que influyen en nuestro comportamiento. Los incentivos pueden ser tanto positivos (un premio por un comportamiento deseable) como negativos (una multa por un mal comportamiento). Desafortunadamente, el debate sobre el uso de incentivos suele ceñirse a si funcionan o no, eludiendo entender cómo su diseño y el contexto en el que se aplican afectan a su efectividad.

En Gneezy *et al.* (2011) mostramos cómo los incentivos no siempre funcionan como se espera. En algunos casos, los incentivos pueden desmotivar a las personas al reducir su motivación intrínseca, un fenómeno conocido como «efecto de desmotivación» (*crowding-out* en inglés). Pensemos, por ejemplo, en los padres que dan pequeños premios a sus hijos por sacar buenas notas en el colegio. Debemos preguntarnos dos cosas: i) ¿consiguen los regalos que los niños saquen mejores notas? y ii) si los niños estudian ahora motivados por el incentivo, ¿pueden perder o no llegar a adquirir nunca el sentido del esfuerzo y la responsabilidad? ¿Qué ocurrirá cuando se retiren los premios? La respuesta a estas preguntas es compleja y depende, también, de cómo de importante es que el incentivo cumpla su función o que el individuo se comporte de la manera buscada *por las razones correctas*. Los incentivos también pueden generar comportamientos de corto plazo que no son sostenibles a largo plazo, cuando quizá las restricciones presupuestarias no permiten mantenerlos.

Además, el tamaño del incentivo puede contener valor informativo sobre lo costoso que puede ser un determinado comportamiento, influyendo con ello en que se cumpla o no. El famoso ejemplo del artículo «Una multa es un precio» (*A Fine is a Price*) de Gneezy y Rustichini (2000) demuestra que una multa demasiado baja, unos

pocos euros por llegar tarde a recoger a los hijos a la guardería, puede provocar que los padres consideren que el llegar tarde no es un problema grave, pues la multa es reducida, incentivando el que lleguen aún más tarde. Lo interesante de este ejemplo es que el valor informativo del incentivo se mantiene incluso cuando este se retira: cuando las guarderías se dieron cuenta de su error y retiraron la multa, los padres no pudieron «desaprender» del valor que se había otorgado a llegar tarde, por lo que siguieron llegando tarde al quitar la multa. Como conclusión, si se va a dar un incentivo, mejor dar uno lo suficientemente grande porque, si no, puede incentivar el comportamiento contrario al que se busca.

Estos ejemplos evidencian que el diseño de incentivos es mucho más complejo de lo que parece y debe considerarse el porqué, el cuándo (antes o después de un comportamiento), el cómo (condicional o no a un comportamiento), la cuantía o incluso si debe ser un incentivo de tipo monetario o se pueden usar otros tipos de motivaciones como la reputación, el sentido del deber, la imagen que tenemos de nosotros mismos o el altruismo para conseguir los objetivos de una política pública.

Una vez hemos argumentado que las intervenciones basadas en las ciencias del comportamiento, ya sean *nudges* o incentivos, aunque parezcan intuitivas, son complejas de diseñar y, por tanto, deben testarse antes de aplicarse a políticas públicas, nos centramos en los siguientes apartados en discutir los requisitos necesarios para fortalecer su aplicación.

3. Actitud favorable a la experimentación

Los más importantes hallazgos obtenidos por las ciencias del comportamiento provienen de la observación empírica sobre cómo deciden y actúan los seres humanos como participantes de experimentos sociales. Un experimento social consiste en la medición de un determinado comportamiento bajo condiciones controladas que permitan identificar su causa. Los experimentos se basan en la asignación aleatoria de los participantes

a un grupo de tratamiento y control de manera que la comparación en las variables conductuales observadas entre los dos grupos permita identificar si las condiciones diferentes a las que se exponen los dos grupos son la causa del diferente comportamiento entre los dos grupos.

El potencial de un buen experimento social se alcanza cuando permite no solo observar que dos grupos se comportan de manera distinta, sino cuando se puede contrastar la causalidad, es decir, la razón por la que actúan de manera distinta. Para ello, es crucial que las condiciones bajo las que se observa el comportamiento en los grupos de tratamiento y control sean la única diferencia entre los dos grupos. Para lograrlo, la asignación a los grupos debe ser completamente aleatoria y no estar basada en ningún factor endógeno que pudiera afectar al comportamiento en el experimento, como, por ejemplo, la mayor predisposición a querer participar de aquellos individuos asignados al tratamiento, o el que sea más fácil que un determinado colectivo forme parte de dicho grupo. Además, ambos grupos deben ser previamente al experimento lo más similares posibles en todas las características que puedan ser observadas. De esta forma, si dos grupos son idénticos antes del experimento, y la única diferencia entre ellos es que a uno de ellos se le enfrenta a un estímulo que el otro grupo no recibe, cualquier diferencia de comportamiento entre tratamiento y control solo podrá deberse a la única diferencia entre ambos grupos, que es el estímulo aplicado. Con ello, podrá decirse que el estímulo, el tratamiento, es la causa del comportamiento diferencial observado.

En la implementación de las políticas públicas no es común que exista un enfoque experimental. Muchas políticas no se testan antes de aplicarse, y son aún menos los casos en los que, por mucho que el político quiera atribuirse el mérito, pueda realmente identificarse la causa de los efectos observados de la implementación de una política pública. Con frecuencia el político decide qué medidas tomar no basándose en la evidencia, idealmente obtenida bajo las condiciones más controladas posibles. Por el contrario, la intuición,

las demandas de la sociedad, la traslación de lo que otros políticos hacen en otros entornos o, en muchos casos, la ideología son la base del diseño de las políticas públicas. Esto no tiene por qué ser necesariamente negativo. A fin de cuentas, si un político ha sido elegido en lugar de otro democráticamente es porque esperamos que una mayoría de la población a la que representa comparta objetivos y valores con él. No obstante, el que se compartan los propósitos e ideales no implica que las políticas diseñadas para alcanzarlos funcionen o que sean la forma más eficaz de hacerlo. La eficiencia de las políticas públicas podría, sin duda, beneficiarse de un enfoque más experimental que testara en pruebas piloto si realmente las medidas que se proponen van a funcionar.

Aunque en la última década se han hecho avances sustanciales en utilizar pruebas piloto, como, por ejemplo, para programas de renta mínima² o con la creación del Laboratorio de Inclusión Social del Ministerio de Inclusión, Seguridad Social y Migraciones, estamos muy lejos de que el enfoque experimental sea práctica común en el diseño de las políticas públicas. Uno de los impedimentos es la falta de formación en la metodología experimental de nuestros responsables políticos y de los funcionarios a cargo de implementar sus políticas. No se puede hacer aquello que se desconoce. La colaboración entre los investigadores experimentales y el sector público es una de las vías más prometedoras para cerrar esta brecha de conocimiento³. Es más, la colaboración con las Administraciones públicas en pruebas piloto de evaluación de políticas públicas se vería favorecida si dicha formación se produjera *antes* de diseñar un programa conjunto para ser evaluado.

² Ver, por ejemplo, <https://ajuntament.barcelona.cat/serveissocials/es/canal/projecte-pilot-b-mincome-combinant-una-renda-minima-garantida-i-politiques-socials-actives-en> y <https://www.inclusion.gob.es/web/policy-lab/laboratorio>

³ Por ejemplo, Mehmood *et al.* (2024) muestran que entrenar a los responsables en técnicas cuantitativas y econométricas incrementa la probabilidad de que asignen fondos a políticas para las que existe evidencia proveniente de ensayos aleatorios controlados (*randomized control trials*, RCTs), demostrando una mayor apreciación por la evidencia causal.

Muchas de las dificultades de estas colaboraciones se producen porque los responsables públicos no han sido suficientemente instruidos en las ventajas de la asignación completamente aleatoria a tratamiento y control, en tener una hipótesis, lo que se suele llamar una «teoría del cambio» sobre cómo la política alcanzará sus objetivos y en la importancia de definir *ex-ante* las variables con las que se va a medir el impacto de una política.

Esto no quiere decir que los investigadores no necesiten también ser instruidos sobre la realidad de cómo se deciden e implementan las políticas públicas, y que no deban aprender sobre la dificultad de mantener sus condiciones de laboratorio ideales cuando se mide el impacto de las políticas. Además, los investigadores podemos contribuir dando más importancia a la divulgación y a las actividades de transferencia, aprovechando ventanas de atención de temas candentes para buscar mayor apoyo a ciertas investigaciones, entendiendo mejor los intereses de los políticos y de la sociedad civil y encontrando un lenguaje común para comunicarnos, no escondiéndonos en los tecnicismos.

Se trata de un proceso de formación bidireccional en el que ambas partes se beneficiarían de acercar posturas en cuanto a los métodos y los objetivos, pero partiendo de una comprensión mutua de los problemas de cada área y de un lenguaje común. Un ejemplo de que esta colaboración puede llegar a ser exitosa son los trabajadores de instituciones como J-PAL, cuyos investigadores afiliados trabajan directamente con sus gobiernos⁴.

4. Disponibilidad pública de datos de los diferentes niveles de la Administración

Para fomentar la colaboración entre la comunidad científica y los responsables políticos es imprescindible que ambos tengan acceso a los datos, tanto a los que permiten conocer las características de los

⁴ J-PAL. Abdul Latif Jameel Poverty Action Lab. <https://www.povertyactionlab.org/es>

afectados por las políticas públicas como los que permiten medir el impacto de las mismas. Como ya mencionamos en Almunia y Rey-Biel (2021), artículo publicado en *Cuadernos Económicos de ICE* y del que este apartado es un resumen, en muchos casos, son las Administraciones públicas las que disponen del monopolio de los datos y se muestran reticentes a compartirlos con los investigadores.

La recolección y procesamiento de datos son actividades esenciales que las Administraciones públicas deben realizar para cumplir muchas de sus funciones básicas. Por ejemplo, las Administraciones públicas necesitan registros de la actividad económica de los contribuyentes para recaudar impuestos, registros de afiliación a la Seguridad Social para gestionar los sistemas de previsión social, o información sobre los contagiados (y sus contactos), ingresados y fallecidos para afrontar una pandemia. No es casualidad que el término «estadística» provenga originalmente del latín *statisticum*, «relacionado con el Estado».

La Administración no puede, ni debe tener el monopolio del análisis de los datos que posee. No puede, en primer lugar, porque no es su función principal. La prioridad de la Administración es gestionar las políticas públicas, no realizar un análisis profundo de los datos de los que dispone. En segundo lugar, porque los recursos públicos disponibles para este fin son reducidos. Aunque la Administración cuenta con algunos organismos especializados en producir información estadística, como el Instituto Nacional de Estadística (INE) o el Centro de Investigaciones Sociológicas (CIS), los recursos económicos de los que disponen para evaluar políticas públicas son limitados. Esta limitada capacidad de análisis de los datos administrativos supone un despilfarro de los recursos destinados a recolectarlos y una ineficiencia respecto al enorme potencial que el uso de esos datos podría tener para el desarrollo y la evaluación de las políticas públicas. Además, no es deseable que la Administración tenga el monopolio del análisis de estos datos, porque podrían surgir conflictos de interés al hacerlo. Por ello, la creación de una entidad

independiente como la Autoridad Independiente de Responsabilidad Fiscal (AIReF) ha significado un paso adelante en la ruptura de los potenciales intereses partidistas en el análisis de los datos.

Permitir que la comunidad científica tenga acceso, en los términos que se consideren oportunos, a registros administrativos resultaría en un enorme incremento de la capacidad conjunta de la sociedad para analizarlos. Expandiría nuestro conocimiento sobre la efectividad de las políticas públicas, aumentaría la transparencia de la actuación de las Administraciones y, con ello, desarrollaría la calidad de nuestra democracia. Como nos gusta decir a los economistas, «los incentivos están alineados»: la Administración se beneficiaría porque ampliaría su limitada capacidad de análisis de los datos, mientras que la comunidad investigadora tendría la oportunidad de explotar esta ingente fuente de información para realizar investigación académica de calidad. Los nuevos conocimientos obtenidos se podrían utilizar para diseñar políticas públicas basadas en la evidencia, algo que desafortunadamente es poco habitual en nuestro país. Afortunadamente, no partimos de cero. En España existen ya algunas bases de datos administrativos disponibles para la investigación, como la Muestra Continua de Vidas Laborales (MCVL) o el Panel de Declarantes del IRPF, pero los ejemplos aún son escasos.

Los datos administrativos presentan numerosas ventajas respecto a los datos provenientes de otras fuentes, como las encuestas, para la evaluación de las políticas públicas. Al cubrir a toda la población, se consigue un tamaño muestral mucho mayor, corrigiendo posibles sesgos de selección que a menudo afectan a las encuestas. Los datos administrativos tienen una tasa de no respuesta mucho menor, lo cual permite disponer de bases longitudinales (panel) más extensas en el tiempo y con menor desgaste (lo que se conoce en inglés como *attrition*). Además, tienen menos errores de medición, aunque estos nunca se eliminan por completo. Por último, los registros administrativos se pueden cruzar con otras fuentes de datos utilizando identificadores individuales o de hogar. Esto es esencial para poder realizar análisis

más profundos sobre causalidad y evaluación de políticas públicas. Los datos administrativos también presentan algunas limitaciones que se deben tener en cuenta (Slemrod, 2016). **En primer lugar**, los datos se recogen atendiendo a objetivos de gestión, lo que afecta al formato, estructura y alcance de los mismos. Por ello, el número de variables disponibles en cada registro suele ser limitado. Por ejemplo, los datos fiscales no suelen tener información socioeconómica de los contribuyentes, más allá de la edad y el sexo. Esta es otra de las razones por las que la posibilidad de cruzar datos de distintos registros es esencial para explotar su verdadero potencial. **En segundo lugar**, otra desventaja de estas bases de datos es que, en la actualidad y salvo recientes convenios firmados por instituciones como el Banco de España y el INE, no se pueden cruzar con otros registros administrativos. **En tercer lugar**, una limitación adicional para el análisis es la segmentación geográfica, pues muchos de estos registros son locales y siguen existiendo dificultades tanto para compartirlos como para recogerlos en formatos estandarizados. No obstante, casos recientes muy notorios, como la caótica gestión de la información durante la pandemia de COVID-19, abren una ventana de oportunidad para conseguir un consenso de todos los sectores de la sociedad sobre que este cambio de cultura es no solo imprescindible, sino beneficioso para la sociedad en su conjunto.

Por último, la colaboración entre los científicos sociales y la Administración pública podría potenciarse aún más si, además de permitir el acceso a datos administrativos y el cruce entre bases de datos, se diera el paso adicional de poder ligar estos datos a los de los mismos participantes en experimentos sociales. Al igual que en los dos temas anteriores, este es un asunto en el que en España vamos comparativamente con cierto retraso frente a los países de nuestro entorno. Cole *et al.* (2020) muestran ejemplos de convenios que se han establecido entre organizaciones y gobiernos para manejar datos y poder ligar información en experimentos. Estos ejemplos pueden servir de guía para los pasos a seguir en nuestro país.

En este sentido, fuera de nuestras fronteras destacan iniciativas como el *Behavioral Insights Team* o las agencias del tipo *What Works* (¿Qué Funciona?) del Reino Unido. La creación de agencias que colaboren con investigadores para la realización y evaluación de políticas basadas en la evidencia puede ser uno de los muchos pasos necesarios hacia la modernización de la Administración pública en España. Afortunadamente, muchos de estos pasos (disponibilidad de datos públicos, creación de una agencia de evaluación, apertura a la colaboración científica para el desarrollo de intervenciones que permitan basar las políticas en la evidencia), están alineados, crean sinergias y pueden marcar un verdadero cambio en la forma de tomar decisiones políticas en nuestro país. Una señal esperanzadora en este sentido es que el Real Decreto-ley por el que se aprobó el Ingreso Mínimo Vital (IMV) incluye en su artículo 30 el compromiso de evaluar los efectos de esta política en concreto. Esperemos que esta evaluación no solo se lleve a cabo bajo los estándares metodológicos más exigentes, sino que sienta el precedente para la evaluación de muchas otras políticas públicas.

5. Diseño de intervenciones públicas de forma que permitan realmente ser evaluadas

La disponibilidad de mejores datos también permitiría diseñar medidas políticas bajo criterios que permitieran su evaluación, creando una colaboración activa, y no solo pasiva, entre investigadores y Administraciones.

De igual forma que la experimentación social informa sobre el modo en el que se comportan los ciudadanos, el enfoque experimental permite también realizar experimentos aplicando directamente políticas públicas. Los conocidos como experimentos aleatorios controlados (*randomized control trials*, RCTs), trasladan la metodología de asignación aleatoria a grupos de tratamiento, que son expuestos a una política pública, y a un grupo de control, de forma que puedan valorarse los efectos de dicha política, en muchos casos en un piloto, antes de escalarla a una población más amplia.

Los RCT, al igual que otros tipos de experimentos, permiten la evaluación causal de una política, de manera que se pueden perfilar los detalles de la misma en distintos tratamientos para medir la contribución de cada uno de los factores que la componen al objetivo social buscado.

El trabajo de los premiados en 2019 con el Premio Nobel de Economía, Esther Duflo y Abhijit Banerjee (junto con Michael Kremer), fundadores de J-PAL, consiste precisamente en realizar los RCT en países en desarrollo para identificar las mejores soluciones a la pobreza. Estos autores destacan que para poder diseñar intervenciones que puedan ser evaluadas, hay que tener en cuenta cada detalle del diseño y la implementación. Los evaluadores necesitan tener un rol activo en el diseño que les permita identificar riesgos, identificar preguntas relevantes. Muchas políticas bien intencionadas fallan en la práctica por detalles que parecen irrelevantes a primera vista. No es suficiente que los investigadores ayudemos en su diseño. Para que las medidas sean efectivas, debemos también implicarnos en la implementación de sus pequeños detalles, aunque esto conlleve, en ocasiones, dejar atrás el cómodo marco de la teoría económica para adentrarnos en un terreno mucho más «fangoso», de ahí su metáfora del «economista como fontanero».

La experimentación no solo consiste en probar distintas cosas y adaptar el diseño de las medidas cuando no se obtienen los mejores resultados. No estamos hablando de medidas improvisadas, sino que estos ensayos controlados son guiados por la teoría económica y, en muchos casos, inspirados por la psicología y la economía del comportamiento, para contar con una intuición previa sobre qué funcionará y qué no. A fin de cuentas, los experimentos se hacen con participantes reales que sufren las consecuencias de las medidas tomadas incluso durante la fase de experimentación y que, además, pueden plantearse si es justo haber sido asignado al grupo de tratamiento de control, por razones puramente de aleatoriedad estadística. Por tanto, no sería ni ético ni práctico probar nuevas medidas,

aunque sea a pequeña escala, sin la intuición rigurosa que puede aportar la teoría ni sin la evidencia empírica previa, que en muchos casos proviene de experimentos controlados más abstractos hechos en un laboratorio.

Cada contexto tiene su especificidad y entender bien lo que diferencia dos entornos puede ser crucial para entender por qué una política puede haber funcionado en el pasado en uno, pero quizá no funcione en otro. Por ello es importante que, si los investigadores queremos seguir participando en el debate público, nos bajemos de la atalaya académica y admitamos con modestia que necesitamos entender mucho mejor las restricciones institucionales y las motivaciones individuales concretas de los sujetos a los que afectan las medidas que proponemos. Duflo ofrece algunos ejemplos de detalles cruciales que pueden determinar que una medida funcione o no (*piping issues*): el nivel de transparencia informativa sobre distintos programas de alimentos o el flujo de fondos entre distintos niveles de la Administración en el caso de políticas descentralizadas. Pero también se ocupa de detalles aún más específicos con soluciones sorprendentes, como en el caso famoso del programa de vacunación de niños en India (Banerjee *et al.*, 2010), en el que comprobó que al incluir un pequeño incentivo no monetario para quienes se vacunaran, unas pocas lentejas, aumentaba exponencialmente las tasas de vacunación. El programa de vacunación era exactamente el mismo, se dieran o no las lentejas, pero solo consiguieron que la gente vacunara a sus hijos cuando lo acompañaban, con un coste mínimo, de unas pocas lentejas.

Los experimentos aleatorios controlados facilitan la evaluación posterior de los efectos de una política pública. No obstante, no siempre es posible realizarlos, ni en muchos casos son necesarios. Por ejemplo, el desarrollo de técnicas de evaluación cuasiexperimentales, atribuida, por ejemplo, a los premiados con el Premio Nobel de Economía de 2021, Joshua Angrist, David Card y Guido Imbens, aprovecha los llamados «experimentos naturales» para evaluar los efectos de una política que no haya sido diseñada con un experimento en

mente, pero que replican bastante las condiciones de un experimento.

Imaginemos, por ejemplo, que queremos evaluar el impacto de un programa de becas a estudiantes en su probabilidad de graduarse o en sus ingresos futuros. Un enfoque experimental supondría asignar aleatoriamente becas a un grupo de tratamiento y no darlas a los estudiantes de un grupo de control. Sin embargo, esto puede crearnos, en ocasiones, problemas éticos o puede que estemos, precisamente, interesados en los efectos de la beca sobre un grupo concreto de estudiantes desfavorecidos. En estos casos, podemos utilizar la cuasialeatorización. Por ejemplo, y utilizando un requisito de nota mínima para obtener una beca, podemos tratar como grupo de control, y darles la beca, a los alumnos que consigan una nota justo superior al requisito de nota para obtener la beca, comparándolos con los alumnos que justo se han quedado a unas décimas de dicha nota y no han obtenido la beca. Ambos grupos serán muy similares, puesto que el sacar unas décimas más de nota en una prueba apenas nos dice algo sobre características diferenciales de ambos grupos. Esta técnica de análisis se denomina «regresión de discontinuidad» y se usa cada vez más para evaluar políticas. Otra posibilidad sería aprovechar si, por alguna razón exógena como el calendario de aplicación de una política, algunas regiones hubieran dado ya las becas, mientras que otras no. Podríamos usar datos de estudiantes con ingresos similares y resultados escolares en diferentes regiones como tratamientos y control. En este caso, no obstante, el grupo de control no es perfecto, porque los estudiantes de dos regiones pueden diferir en muchos otros aspectos, relacionados con la región en la que viven, con lo que las diferencias en ingresos futuros podrían también estar causadas por estos factores. En estas situaciones se puede usar el método de «diferencias en diferencias», consistente en comparar los ingresos de los alumnos con beca con el grupo de control en otras regiones, y para descartar que las diferencias se deban a otros factores, podríamos comparar también a estudiantes de las mismas características en las dos regiones, pero antes de

que se pusiera en marcha el programa de becas. En esta segunda comparación, las diferencias entre los dos grupos no se deben a las becas, lo que nos permitiría «eliminar» los efectos de otros factores que difieren entre las dos regiones.

6. Evaluación independiente y rigurosa de las políticas públicas

La evaluación de las políticas públicas por parte de agencias independientes es crucial por varias razones que tienen que ver con la objetividad, la transparencia y la mejora continua de las políticas implementadas.

Las agencias independientes, al no estar directamente vinculadas a los intereses de los gobiernos o de los actores políticos, pueden proporcionar una evaluación más objetiva. Esto es vital para garantizar que los resultados sean confiables y reflejen la efectividad real de las políticas, no solo lo que interesa al gestor público. Por ejemplo, una debilidad del *Behavioral Insights Team* británico cuando se creó como oficina dependiente del primer ministro, fue que no tenía obligación de ofrecer consejos independientes. Esto llevó, en el contexto de la pandemia del coronavirus, a embarazosas declaraciones del director de la agencia, David Halpern, y a la percepción de que el fallido intento del Gobierno británico de no confinar a su población estaba basado en mala ciencia conductual.

Además, cuando las evaluaciones son realizadas por agencias externas e independientes, los resultados suelen percibirse como más creíbles por parte de la ciudadanía. La transparencia en la evaluación de las políticas fortalece la confianza en las instituciones. Las agencias independientes pueden identificar de manera más imparcial los puntos débiles de las políticas públicas, lo que facilita la corrección de errores y la mejora de los programas. Sin la presión política de «aprobaciones» inmediatas, estas agencias pueden centrarse en ofrecer recomendaciones claras para optimizar los programas y mejorar su impacto social. Cuando las evaluaciones se realizan internamente dentro de las mismas instituciones

que implementan las políticas, puede existir el riesgo de conflictos de interés, ya que las agencias encargadas de implementar las políticas podrían sentirse comprometidas a defender sus propios proyectos o decisiones. Las agencias independientes eliminan este riesgo y permiten una revisión más imparcial.

Las agencias independientes que evalúan políticas públicas también juegan un papel fundamental en promover la rendición de cuentas. Su existencia asegura que el gobierno y otras instituciones rindan cuentas sobre el uso de los recursos públicos y la efectividad de las políticas implementadas, lo que resulta en una mayor responsabilidad ante la ciudadanía.

La colaboración de la comunidad científica es crucial en la creación y mantenimiento de estas agencias independientes. De hecho, estas agencias se nutren cada vez más de personal cuya formación científica permite superar la falta de formación técnica en métodos de evaluación del personal administrativo más tradicional de las Administraciones públicas. El rigor que aporta el enfoque científico y la menor intencionalidad política de quienes han sido expresamente formados para estudiar las causas y los efectos de un fenómeno, deben convertirse en un activo esencial en el reciente fenómeno de creación de agencias de evaluación, al que España poco a poco se está incorporando.

7. Conclusiones

La colaboración entre los científicos del comportamiento y las Administraciones públicas es fundamental para diseñar y mejorar políticas públicas más efectivas y centradas en las necesidades reales de la sociedad. Los científicos del comportamiento, con su comprensión de cómo las personas toman decisiones, procesan la información y responden a incentivos, pueden ofrecer a los responsables políticos herramientas basadas en la evidencia para abordar problemas complejos de manera más eficiente. Al integrar principios de la psicología, la economía del comportamiento y otras disciplinas relacionadas, se pueden crear intervenciones que no solo

sean viables, sino también sostenibles desde la perspectiva del comportamiento de los ciudadanos. Esta colaboración también permite a las Administraciones públicas mejorar la eficacia de sus programas, aumentar la aceptación de nuevas políticas y promover cambios de comportamiento positivos en la población. Además, al aplicar los conocimientos del comportamiento de manera práctica, las políticas pueden ser adaptadas a contextos locales y a las diversas realidades sociales, lo que favorece su éxito a largo plazo. En última instancia, fortalecer esta cooperación entre científicos del comportamiento y Administraciones públicas contribuye a una gestión más eficiente y justa, beneficiando a la sociedad y fomentando la confianza en las instituciones.

Sin embargo, aún estamos lejos de que esta colaboración alcance todo su potencial. Factores como la falta de formación y el desconocimiento sobre los resultados provenientes de las ciencias del comportamiento, fomentados además por la reciente «crisis de replicabilidad» de los resultados de ciertos experimentos de psicología social, nos alejan de este objetivo. La polarización y los propios sesgos ideológicos de los responsables públicos no favorecen un clima de confianza entre científicos y políticos, más cuando desde ciertos sectores se han cuestionado los resultados científicos. También los científicos somos responsables de fortalecer esta confianza. Especialmente en el campo de las ciencias del comportamiento, donde los resultados anecdóticos son llamativos, es importante mantener el rigor científico para poder hacer entender al administrador público las sutilezas y los detalles de una intervención conductual. No todo va a funcionar y, por ello, es necesario que esa colaboración se base en la evaluación de dichas intervenciones, favorecida por el acceso a datos y la creación de una colaboración estrecha que entienda las diferencias en objetivos, enfoques y tiempos de unos y otros agentes.

Referencias bibliográficas

Almunia, M. y Rey-Biel, P. (2021). La gestión de los datos administrativos en España: diagnóstico y retos de futuro.

- Cuadernos Económicos de ICE*, (102), 103-131. <https://doi.org/10.32796/cice.2021.102.7313>
- Banerjee, A. V., Duflo, E., Glennerster, R., & Kothari, D. (2010). Improving immunisation coverage in rural India: Clustered randomised controlled evaluation of immunisation campaigns with and without incentives. *BMJ*, 340. <https://doi.org/10.1136/bmj.c2220>
- Bonargent, A. (2024). *Can research with policymakers change the world?* (IGC Working Paper XXX-23018). International Growth Centre. <https://www.theigc.org/sites/default/files/2025-04/Bonargent-Working-Paper-February-2024.pdf>
- Cole, S., Dhaliwal, I., Sautmann, A., & Vilhuber, L. (Eds.). (2020). *Handbook on using administrative data for research and evidence-based policy*. Abdul Latif Jameel Poverty Action Lab.
- Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives*, 25(4), 191-210.
- Gneezy, U., & Rustichini, A. (2000). A fine is a price. *Journal of Legal Studies*, 29(1), 1-17.
- Mehmood, S., Naseer, S., & Chen, D. (2024). *Training policymakers in econometrics* (NBER Working Paper). National Bureau of Economic Research. https://users.nber.org/~dlchen/papers/Training_Policy_Makers_in_Econometrics.pdf
- Slemrod, J. B. (2016). *Tax compliance and enforcement: New research and its policy implications* (SSRN Ross School of Business Paper No. 1302). Social Science Research Network. <https://doi.org/10.2139/ssrn.2726077>