

Resumen

El presente artículo realiza un análisis de la evidencia reciente sobre las consecuencias no intencionadas de las políticas sanitarias y educativas. Mediante varios ejemplos reales, examinamos algunos de los mecanismos comunes que provocan efectos inesperados y potencialmente indeseables. Entre los mecanismos analizados se encuentran las respuestas conductuales heterogéneas ante las mismas políticas públicas, las externalidades y los efectos de arrastre, la interacción entre distintas políticas, y la incorrecta generalización de los resultados de determinadas investigaciones. Nuestro trabajo pone de manifiesto la necesidad de considerar múltiples mecanismos y resultados al evaluar el efecto causal de las políticas públicas.

Palabras clave: consecuencias no intencionadas, análisis de políticas públicas, educación, sanidad.

Abstract

This paper reviews recent evidence on the unintended consequences of health and education policies. Through a series of topical examples, we explore common mechanisms driving unexpected and potentially undesirable effects. The mechanisms we discuss include: heterogenous behavioural responses; externalities and spillovers; policy interactions; and the incorrect generalization of research findings. Our exploration suggests the need to account for a multiplicity of mechanisms and outcomes when evaluating the causal impact of policies.

Keywords: unintended consequences, public policy analysis, education, health.

JEL classification: H50, I18, I28.

CONSECUENCIAS NO INTENCIONADAS DE LAS MEDIDAS DE POLÍTICA PÚBLICA: EJEMPLOS DE SANIDAD Y EDUCACIÓN

Jennifer GRAVES (*)

Universidad Autónoma de Madrid (UAM)

Eugenio ZUCHELLI

*Madrid Institute for Advanced Study (MIAS),
Universidad Autónoma de Madrid (UAM), University of Lancaster, IZA*

I. LA IMPORTANCIA DE TOMAR DECISIONES BASADAS EN LA EVIDENCIA

DETERMINAR correctamente los efectos causales de las políticas públicas resulta fundamental para comprender su capacidad para conseguir los objetivos que constituyen la esencia motivadora de dichas políticas. Por ejemplo, cuando un país aplica una política sanitaria resulta fundamental conocer si tiene un impacto real en la salud, o cuando aplica una política educativa resulta importante determinar si ha mejorado el aprendizaje del alumnado según lo previsto. En este artículo revisamos algunos ejemplos provenientes de la literatura académica sobre políticas públicas en materia de sanidad y educación para mostrar la importancia de determinar e interpretar correctamente los efectos causales de las políticas públicas.

Aunque muchas de estas políticas funcionan según lo esperado, en cuyo caso la cuestión pertinente a la hora de formular políticas es si los beneficios justifican los costes de tales programas, en ocasiones las medidas de política pública pueden aca-

rear consecuencias no intencionadas. Cuando se dan estos casos, la política en cuestión puede tener un impacto contrario al que esperaban los responsables políticos, o quizá alcance el objetivo marcado, pero también cause «efectos secundarios» imprevistos que deben tenerse en cuenta en el cálculo de la relación coste-beneficio de la medida. En este trabajo proporcionamos ejemplos de cómo la mejor de las voluntades puede acabar dando lugar a un mal resultado en la práctica y de cómo tanto investigadores como responsables políticos pueden aprender de los errores pasados para evitar problemas comunes que den lugar a consecuencias no intencionadas en la aplicación de políticas.

En la primera parte del trabajo, revisamos una serie de políticas sanitarias que han tenido importantes consecuencias no intencionadas debido a que la forma en la que se pusieron en práctica no interpretaba correctamente la evidencia científica. En concreto, mostramos la importancia que tiene en la implementación de políticas públicas tener en cuenta los siguientes aspectos: i) la heterogeneidad de tratamiento; ii) los efectos arrastre y las potenciales exter-

nalidades; y iii) la interacción entre regulaciones en distintos niveles de Gobierno (por ejemplo, autonómico/local vs. nacional). En la segunda parte del trabajo, revisamos distintos ejemplos de políticas educativas para mostrar la importancia de tener en cuenta los siguientes aspectos: i) considerar el abanico de respuestas conductuales que se derivarán de la política pública; ii) comprender adecuadamente cómo se generalizarán los resultados en otros contextos distintos a los específicamente investigados en los trabajos de investigación; y iii) tener en cuenta la multiplicidad de objetivos que puede cumplir una política educativa, y en particular cómo la educación de los niños pequeños también hace las veces de servicio de cuidado infantil durante la jornada laboral y cómo esto puede afectar a las decisiones familiares.

Los ejemplos proporcionados en las siguientes secciones (sección segunda sobre políticas sanitarias y sección tercera sobre políticas educativas) se escogieron con el objeto de mostrar que no tener en cuenta determinados factores a la hora de determinar los efectos causales de una política pública implicaría una incorrecta interpretación de la repercusión total real de esa política y podría conducir a tomar decisiones inadecuadas. Los estudios académicos seleccionados no pretenden realizar una revisión exhaustiva de la literatura académica sobre políticas sanitarias y educativas, sino ilustrar con determinados ejemplos algunos aspectos comunes que deberían tenerse en cuenta en todas las medidas de política pública. De hecho, cualquiera de los mecanismos analizados mediante el uso de políticas sa-

nitarias podría haberse expuesto de igual manera usando medidas de política educativa, y viceversa. Nuestro objetivo con el presente estudio es centrar la atención en lo habituales y relevantes que son estas cuestiones en la aplicación de cualquier política pública en general.

II. EJEMPLOS DE POLÍTICAS SANITARIAS

Pese al rápido aumento de la bibliografía acerca de la evaluación del impacto de las políticas sanitarias, sigue habiendo desafíos metodológicos y prácticos que deberían tenerse en cuenta para poder interpretar correctamente la evidencia disponible sobre el impacto causal de las políticas sanitarias. Como hemos mencionado anteriormente, en esta sección nos centraremos en la importancia de considerar: i) la heterogeneidad de respuestas conductuales ante una misma política pública; ii) las potenciales externalidades y efectos de arrastre; y iii) la interacción entre leyes a distintos niveles de Gobierno (por ejemplo autonómico/local vs. nacional). Para ello, analizamos las consecuencias no intencionadas de varias medidas de salud pública de gran alcance, incluidas las destinadas a reducir el daño provocado por el consumo de alcohol y tabaco, así como determinadas políticas de seguridad pública.

1. Efectos heterogéneos del tratamiento

Adda y Cornaglia (2010), posiblemente, uno de los estudios más citados acerca de las respuestas conductuales no intencionadas y potencialmente perniciosas de las políticas anti-tabaco. Los autores analizan los

efectos de las prohibiciones de fumar y de los impuestos especiales en la exposición al humo del tabaco de los *no fumadores* en Estados Unidos. Su principal conclusión es que la prohibición de fumar en espacios públicos (lugares de trabajo, bares y restaurantes) desplazó a los fumadores de los espacios públicos a los privados (sus hogares) y en realidad *incrementó la exposición al humo de los fumadores pasivos*, incluidos menores. Cabe destacar que los autores examinan específicamente la *heterogeneidad* de los efectos de las políticas antitabaco y que los resultados se presentan en base al consumo de tabaco en el hogar (familias fumadoras vs. no fumadoras), grupo de edad (menores vs. adultos) y día de la semana (días laborables vs. fines de semana).

Entre los diversos hallazgos interesantes presentados en el estudio, los autores constatan que *ninguna* de las dos políticas antitabaco –prohibiciones de fumar e impuestos especiales– parecen tener un impacto en las personas que viven en hogares no fumadores si se consideran conjuntamente todos los grupos de edad. Sin embargo, esta fusión de los grupos de edad enmascara una importante heterogeneidad en los efectos. En concreto, también observan aumentos estadísticamente significativos en los niveles de cotinina: i) durante los días laborables, en los menores que conviven con fumadores tras la implantación de prohibiciones de fumar más estrictas en lugares de trabajo; y ii) durante los fines de semana, en los menores que viven en familias no fumadoras tras la implantación de prohibiciones en bares y restaurantes.

Estos resultados evidencian el desplazamiento de los fumadores desde los espacios públicos a los domicilios particulares. Los autores apuntan que, tras la implantación de estas prohibiciones, los fumadores tienden a pasar menos tiempo fumando en lugares de trabajo, bares y restaurantes y más tiempo en domicilios particulares, lo cual incrementa la exposición pasiva al tabaco de sus propios hijos y de los hijos de las familias no fumadoras. A la vista del nivel de cotinina detectado en los menores, este efecto desplazamiento supone un aumento de alrededor de 2-4 cigarrillos diarios fumados en su presencia. A su vez, a diferencia de los resultados que se obtienen al analizar las prohibiciones de fumar, este mismo estudio concluye que la subida de los impuestos del tabaco reduce la exposición media al humo de los fumadores pasivos. En consecuencia, los resultados constatan efectos contrarios a los previstos en menores y *efectos potencialmente contrapuestos* de dos políticas públicas que tienen el mismo objetivo inicial.

Un estudio posterior de Carpenter, Postolek y Warman (2011) emplea un enfoque de diferencias en diferencias similar al utilizado en Adda y Cornaglia (2010), para analizar la variación temporal y geográfica provocada por la implantación de las prohibiciones de fumar en Canadá. Este estudio obtiene un resultado distinto, ya que no encuentra evidencias de efecto desplazamiento del humo ambiental del tabaco hacia los hogares. El estudio, sin embargo, se apoyaba en datos de encuesta reportados por los individuos y no incluía marcadores biológicos como la cotinina. Por este motivo, el uso de la concentración de cotinina

en fluidos corporales (un biomarcador usado para medir la exposición al humo ambiental) utilizado por Adda y Cornaglia (2010) resulta más fiable desde el punto de vista científico ya que los efectos nocivos en los menores del hogar se identificaron mediante los aumentos estadísticamente significativos de los niveles de cotinina. Dichos aumentos aparecían de forma clara en el análisis estadístico, pero es posible que fueran imperceptibles para los padres y, en consecuencia, no serían recogidos en los datos de encuesta.

Cabe señalar que Adda y Cornaglia pretendían evaluar los efectos de las políticas antitabaco en no fumadores, ya que en el momento de la publicación de su estudio no había pruebas sólidas sobre la cuestión, especialmente en la literatura económica. Desde entonces, la literatura económica sobre los efectos causales, tanto previstos como no previstos, de las políticas antitabaco ha crecido de manera notable y ahora ya existen numerosos estudios sobre esta cuestión. Hoy en día se reconoce ampliamente que las prohibiciones de fumar en espacios públicos parece que únicamente reducen el consumo de tabaco en determinados subgrupos, como es el caso de las personas que acuden con frecuencia a bares y restaurantes o que fuman mucho (p. ej., Anger, Kvasnicka y Siedler, 2011; De Chaisemartin, Geoffard y Le Faou, 2011; Jones *et al.*, 2015).

Los resultados de la literatura sobre las políticas antitabaco son un ejemplo ilustrativo, pero existen otros ejemplos de política pública en materia de salud en los que también se han encontrado respuestas conductuales no previstas. Por ejemplo,

Carpenter y Stehr (2011) utilizan un diseño cuasiexperimental basado en los distintos momentos en los que se aprobaron las leyes que introducían la obligatoriedad del uso del casco para montar en bicicleta en los distintos estados de Estados Unidos. Los autores se centran en los efectos de estas políticas en el uso del casco y de la bicicleta, y también en el número de víctimas mortales. Aunque sus resultados confirman que la legislación de uso obligatorio del casco reduce las víctimas mortales e incrementa el uso de esta medida de protección, estas medidas también parecen causar una reducción del uso de la bicicleta entre los jóvenes. Los autores mencionan como posibles mecanismos causantes de este resultado no intencionado el coste de los cascos, la existencia de alternativas más baratas que la bicicleta, y el estigma asociado a llevar un casco. Finalmente, y como consecuencia de todos esos efectos, sugieren que convendría realizar un análisis exhaustivo de la «relación coste-beneficio» antes de implantar este tipo de políticas.

2. Interacción entre distintas políticas y entre distintos niveles de Gobierno

La interacción entre políticas en distintos niveles de Gobierno también puede derivar en respuestas conductuales no intencionadas. Un ejemplo es el trabajo de Adams y Cotti (2008) sobre la legislación antitabaco a nivel estatal y local en Estados Unidos. Su trabajo usa como fuente de identificación causal la variación geográfica entre las prohibiciones de fumar locales y estatales. Contrariamente a lo esperado por los legisladores, observan un aumento en los accidentes mortales relacionados con la conduc-

ción bajo los efectos del alcohol. Este resultado, sorprendente a primera vista, pero robusto a una variedad de especificaciones, parece venir explicado por el hecho de que una parte de la población fumadora reaccionó ante la legislación antitabaco en su condado conduciendo distancias más largas hasta condados limítrofes en los que sí estaba permitido fumar en bares. Sus conclusiones muestran la utilidad, al aprobar una política pública, de tener en cuenta otras políticas relacionadas como lo es, en este caso, el posible impacto en la conducción bajo los efectos del alcohol (por ejemplo, impuestos del alcohol y legislación sobre la tasa de alcohol en sangre).

En general, la conclusión que puede obtenerse de estos trabajos es que el éxito de una política pública depende de los contextos normativos en el que se aplica y en concreto de que exista coordinación entre distintos niveles de Gobierno y distintas políticas públicas. Por ejemplo, al implantar una política local, los responsables políticos han de considerar las políticas estatales existentes y la posible interacción entre ambas normativas. Asimismo, estos trabajos también ponen de manifiesto la importancia de tener en cuenta las políticas existentes en territorios geográficamente limítrofes, las cuales podrían proporcionar una vía para que los ciudadanos eludan el cumplimiento de la nueva norma, haciéndola ineficaz y dando lugar, en ocasiones, a resultados contraproducentes.

3. Efectos de sustitución y arrastre

Al igual que ocurre con las conductas descritas en el apartado anterior, al implementar

políticas sanitarias eficaces han de tenerse en cuenta una multitud de efectos de sustitución y arrastre. Los Gobiernos deben ser conscientes de estos efectos, especialmente al implantar políticas destinadas a restringir el consumo de sustancias perjudiciales para la salud. Un ejemplo muy conocido es el estudio de DiNardo y Lemieux (2001), que analiza el impacto del incremento de la edad mínima para consumir alcohol sobre el consumo de bebidas alcohólicas y marihuana en alumnos estadounidenses del último curso de instituto. Los autores comparan distintos momentos del tiempo en los que gradualmente se aumentó la edad mínima de consumo de alcohol desde los 18 a los 21 años en diversos estados del país entre 1980 y 1989. Los datos utilizados proceden de una amplia muestra de alumnos extraídos de la encuesta *Monitoring the Future (MTF*, «Observando el futuro»). Los resultados de este estudio muestran de manera robusta que, si bien el aumento de la edad mínima redujo el consumo de alcohol, también incrementó el consumo de marihuana entre los estudiantes. En concreto, el destacado efecto positivo en términos estadísticos sobre el consumo de marihuana es sistemáticamente mayor en valores absolutos que el efecto negativo sobre el consumo de alcohol en todos los modelos (3,7 por 100 frente al 2,6 aproximadamente, en función de la especificación y las variables observables incluidas). No obstante, los autores señalan que dado que el porcentaje de estudiantes que consumen alcohol (67 por 100) es mucho mayor que en el caso de la marihuana (25 por 100), en realidad el efecto agregado es mayor para las bebidas alcohólicas que en la marihuana.

Este estudio es un ejemplo de cómo la consecuencia no intencionada de esta medida (aumento del consumo de marihuana) podría deberse a simples efectos de sustitución.

Las conclusiones del estudio anterior han sido matizadas posteriormente por otros trabajos que muestran que además del posible efecto sustitución también puede darse un efecto arrastre entre el alcohol y la marihuana. A este respecto, Crost y Guerrero (2012) constatan que si bien el consumo de alcohol en EE. UU. aumenta, como es de esperar, a los 21 años, el de marihuana también lo hace en torno a la misma edad. Es decir, cuando la normativa sobre bebidas alcohólicas se vuelve más permisiva (a los 21 años) y ambos productos están potencialmente disponibles, los resultados indican que el alcohol y la marihuana podrían ser productos complementarios. Por otro lado, Anderson, Hansen y Rees (2013) analizan los efectos que tiene la legislación sobre marihuana medicinal (LMM), que elimina las sanciones estatales por el consumo, posesión y cultivo de marihuana, en las muertes en accidentes de tráfico en Estados Unidos. Observan que, aunque la LMM reduce el número de víctimas mortales en carretera (aproximadamente entre un 8 y un 11 por 100 tras un año de su entrada en vigor), los efectos son mayores en las muertes relacionadas con el consumo de alcohol (disminución del 13,2 por 100), especialmente de los casos en los que al menos un conductor registra una tasa de alcoholemia superior al 0,10 (disminución del 15,5 por 100). Esto también respaldaría, en cierto modo, los potenciales efectos de interacción entre el alcohol y la marihuana.

A su vez, Carpenter y Dobkin (2011) examinan una variedad de consecuencias de la legislación que impone una edad mínima para consumir alcohol. En concreto analizan los efectos sobre el consumo de alcohol, la mortalidad, las lesiones no mortales y la delincuencia utilizando una amplia variedad de bases de datos estadounidenses. Concluyen que limitarse únicamente a reducir la edad mínima legal para consumir alcohol probablemente daría lugar a un aumento de los problemas relacionados con el alcohol. También indican que no existen evidencias suficientes para evaluar, ni por separado ni en conjunto, los efectos de políticas complementarias a las de establecimiento de una edad mínima, como pueden ser aquellas que educan sobre las consecuencias del consumo de alcohol, o las que afectan a la concesión de licencias para su venta. Por tanto, harían falta más datos y pruebas para determinar con certeza el efecto neto de las iniciativas relativas al alcohol.

Por último, las políticas relacionadas con el alcohol también pueden tener efectos *positivos* no intencionados. Un estudio reciente de Wada *et al.* (2017) emplea un modelo estructural que simula los efectos de un aumento del 5 por 100 en el IVA de las bebidas alcohólicas utilizando datos de Estados Unidos. Los autores constatan que, aunque subir los impuestos a las bebidas alcohólicas provocaría un desplazamiento de puestos de trabajo entre sectores de la economía (desde los sectores relacionados con el consumo de alcohol a los sectores no relacionados), el impacto en el número total de puestos laborales sería positivo, aunque escaso. Esto implicaría

que, aparte de los efectos intencionados sobre el consumo de alcohol según demuestran varios estudios (por ejemplo, Xu y Chaloupka, 2011), una subida de impuestos de este tipo también podría tener un efecto positivo general en el empleo.

En resumen, los resultados que hemos descrito acerca del impacto de las políticas sanitarias generan dudas sobre la eficacia global de algunas medidas de salud pública. Además, subrayan las ventajas que supone combinar *diversas fuentes de datos* que permitan analizar un *amplio rango* de resultados y que permitan el análisis de subgrupos poblacionales heterogéneos. Los Gobiernos que deseen reducir el consumo de productos perjudiciales para la salud, como el alcohol, mediante políticas que establezcan edades mínimas, prohibiciones e impuestos, no han de considerar los efectos en el consumo del producto en cuestión de manera aislada, sino que también deben tener en cuenta los posibles *efectos de sustitución* entre productos similares y los *efectos de arrastre* sobre otros sectores de la economía. Esta advertencia no solo se aplica a productos claramente perjudiciales para la sociedad como son el alcohol, la marihuana y el tabaco, sino también a cualquier producto para el que la legislación intente disuadir de su consumo. Por ejemplo, varios estudios acerca de los efectos indirectos de gravar con mayores impuestos los alimentos poco saludables y las bebidas azucaradas también recomiendan analizar detalladamente si las sustituciones podrían socavar los beneficios para la salud esperados de dichas medidas (por ejemplo, Cornelsen *et al.*, 2015). También se han observado con-

secuencias no intencionadas en materia de desigualdad, dado que la subida de impuestos de los alimentos poco saludables podría tener un carácter regresivo al favorecer a los consumidores con rentas más altas, que suelen consumir alimentos más saludables y tener mayor capacidad de respuesta ante las variaciones de precios (por ejemplo, Muller *et al.*, 2017).

III. EJEMPLOS DE POLÍTICAS EDUCATIVAS

En esta sección continuamos con la revisión de los efectos no intencionados de determinadas políticas públicas utilizando en este caso como ejemplo diversas medidas de política educativa que han dado lugar a resultados no previstos inicialmente.

1. Comportamientos no anticipados inicialmente

Nuestro primer ejemplo procede de la extensa literatura existente sobre lo que definiremos, en sentido amplio, como «programas de rendición de cuentas». Dos ejemplos de este tipo de políticas, provenientes de Estados Unidos, son el programa *No Child Left Behind (NCLB)*, «Que ningún niño se quede atrás», implantado durante la administración Bush, y el programa *Race to the Top* («Carrera a la cima»), implementado durante la administración Obama y que también incluye el uso de parámetros para la evaluación y remuneración de centros educativos y personal docente, como son medidas del valor añadido del profesor y la medición del progreso anual de cada centro. A nivel de centros educativos, el objetivo de estas políticas es mejorar el rendimiento educati-

vo, para lo que se utilizan datos sobre el rendimiento académico del alumnado en exámenes externos para premiar o castigar a los centros en función de su cumplimiento de las expectativas. El argumento es que, si existe una capacidad limitada de realizar un seguimiento eficaz de la calidad del centro y de su producción educativa, los docentes podrían comportarse de manera contraria a los intereses de la comunidad (esto es, padres o responsables políticos locales).

En principio, proporcionar información independiente y fiable acerca del buen funcionamiento de los centros, y asociar consecuencias al incumplimiento de las expectativas por parte del centro, puede incentivar a los docentes a ajustarse a las metas educativas establecidas (véase Figlio y Loeb, 2011). Sin embargo, poner en práctica la rendición de cuentas en los centros trae consigo una serie de decisiones prácticas; y es en estas decisiones donde surgen las consecuencias no intencionadas. Por ejemplo, para implementar una política de este tipo es necesario tomar una decisión normativa en cuanto a qué contenido incluir en los exámenes estandarizados, pues se plantea la disyuntiva de centrar la evaluación del alumno en un conjunto reducido de asignaturas o en uno más amplio. Por un lado, evaluar un conjunto amplio de asignaturas periódicamente tiene la ventaja de representar de manera más precisa una variedad de materias socialmente importantes, aunque al mismo tiempo puede ser difícil de cuantificar o puede tener un coste prohibitivo. Por otro lado, administrar estos exámenes resta tiempo adicional a la docencia lo que puede tener efectos en el rendimiento acadé-

mico, por lo que, aunque existe cierta heterogeneidad en los programas de rendición de cuentas estadounidenses, la mayoría se centran más específicamente en la lectura y las matemáticas (Figlio y Loeb, 2011).

Desafortunadamente, está ampliamente documentado que cuando los exámenes son decisivos para la carrera del estudiante (*high-stakes testing*), algo que suele ser inherente a los programas de rendición de cuentas, estos provocan un fenómeno llamado «enseñar para el examen» (*teaching to the test*), en el que los centros educativos adaptan la formación para centrarse más activamente en las asignaturas evaluadas y menos en las que no se evalúan (por ejemplo, Ladd y Zelli, 2002; Stecher *et al.*, 2000; Linn, 2000). Poner el foco en las asignaturas evaluadas podría tener un efecto positivo si dichas asignaturas no recibieran la atención suficiente y representaran adecuadamente los objetivos de aprendizaje más importantes. Sin embargo, esta no parece ser la tónica dominante, puesto que la frase «enseñar para el examen» ha adquirido en sí misma una connotación peyorativa en las políticas educativas. Un enfoque limitado a unas pocas materias podría descuidar el tiempo dedicado al aprendizaje de conocimientos y habilidades valiosos, pero más difíciles de medir, como el razonamiento crítico o la creatividad. De la misma manera, incluso en las asignaturas evaluadas se ha observado que los docentes se limitan a impartir el material que se espera que aparezca en el examen, perdiendo así algunas de las competencias generales en dichas materias (por ejemplo, Koretz y Barron, 1998; Figlio y Loeb, 2011). También se ha documentado otro compor-

tamiento estratégico por parte de los docentes consistente en dedicar tiempo de enseñanza a mejorar la habilidad para realizar exámenes en lugar de dedicarlo al contenido curricular, así como centrar la formación únicamente en las áreas más fácilmente mejorables (Chakrabarti, 2013). También se han documentado comportamientos todavía más llamativos como los analizados en el trabajo de Jacob y Levitt (2003) en el que se constata que algunos docentes de Chicago cometieron fraude en los exámenes oficiales, llegando a modificar las respuestas de los alumnos para alterar los resultados académicos, lo que, lógicamente, repercutía también, en los resultados obtenidos por sus centros en los programas de rendición de cuentas.

Otro comportamiento estratégico que surge como respuesta directa a las presiones propias de los programas de rendición de cuentas de los centros educativos es la manipulación de a quién incluir en los exámenes. Resulta pertinente que en las políticas públicas se cuestione a quién incluir o excluir de los exámenes decisivos. Por ejemplo, ¿un centro debe rendir cuentas por los alumnos recién transferidos, por aquellos con discapacidades o por los que no sean hablantes nativos de la lengua de docencia? Han de establecerse normas para saber cómo contabilizar a dichos estudiantes con el fin de que los resultados de los exámenes sean fiables, aunque también se ha abusado de tales normas con fines estratégicos. Por ejemplo, se ha encontrado evidencia de que algunos centros educativos han reclasificado a los alumnos con bajo rendimiento académico como estudiantes con discapacidad para manipular la lista de examinados (por ejem-

plo, Figlio y Getzler, 2007; Cullen y Rebak, 2006), mientras que otros centros han modificado sus medidas disciplinarias particularmente durante los períodos de evaluación para que los alumnos con bajo rendimiento se quedaran en casa y, por tanto, no computaran en los promedios de los exámenes (Figlio, 2006).

Un objetivo adicional de estos programas de rendición de cuentas suele ser la mejora de los resultados académicos de los grupos poblacionales más desfavorecidos. No obstante, en muchas ocasiones, los modelos de evaluación empleados en los programas de rendición de cuentas no capturan de manera precisa las causas potencialmente importantes del bajo rendimiento académico de dichos grupos. Un sistema de evaluación que no permita tener en cuenta los problemas propios del perfil del estudiante ni la falta de recursos en centros más desfavorecidos atribuirá injustamente el incumplimiento de las expectativas al centro y a los docentes, que resultarán además injustamente castigados por el sistema, lo que provoca efectos adicionales contraproducentes (Ladd y Walsh, 2002). Si la falta de recursos escolares es el motivo del bajo rendimiento académico del alumno, un castigo que reduzca los recursos tan solo agravaría el problema. Figlio (2003) ha documentado algunas de estas consecuencias de financiación en distritos con muchas minorías y alumnos de pocos recursos en lo relativo a la normativa para el uso de los fondos incluidos en el programa *NCLB*.

Aunque sigue habiendo muy pocos datos acerca de los programas de mejora escolar en el marco de *Race to the Top* en

comparación con la amplia literatura existente sobre anteriores programas de rendición de cuentas, estudios recientes del estado de Carolina del Norte han constatado que la medida fue infructuosa y que, de hecho, es posible que haya mermado el rendimiento académico del alumnado, al haber restado tiempo de docencia por añadir carga administrativa, y haber derivado en rotaciones no anticipadas entre colegios de directores y profesorado (Heissel y Ladd, 2018). Una política que incentive cambios en la manera de enseñar en los centros educativos también modifica, evidentemente, las condiciones laborales del profesorado y, en consecuencia, puede alterar sus preferencias respecto a dónde trabajar. Nótese que esto no implica necesariamente que la rotación produzca efectos negativos, sino que podría tenerlos. De hecho, los efectos de esa rotación en la calidad docente y en el rendimiento del alumnado parecen ser heterogéneos (por ejemplo, Rockoff *et al.*, 2012; Clotfelter *et al.*, 2004).

En general, las conclusiones extraídas de los programas de rendición de cuentas que hemos resumido destacan la importancia de anticipar comportamientos estratégicos provocados por los incentivos incorporados en las políticas educativas. También muestran que un parámetro destinado a la medición, si se convierte en un parámetro que acarrea consecuencias académicas o laborales, puede alterar el comportamiento en formas deseadas y no deseadas. En general, el efecto neto de los programas de rendición de cuentas parece positivo, ya que dan lugar a un aumento en gran medida del rendimiento medio del alumnado en las asigna-

turas evaluadas tanto a corto plazo como a lo largo del tiempo (Figlio y Loeb, 2011). Sin embargo, pese a que los incentivos logren los objetivos marcados, se ha de ser consciente de que dichos incentivos también producen efectos negativos que limitan los beneficios de la medida.

2. Problemas asociados a la incorrecta generalización de resultados

Otro tipo de trabajos que arroja luz acerca de consecuencias no intencionadas de las políticas públicas son las investigaciones sobre el impacto académico del número de alumnos por aula. Al realizar una evaluación de esta medida, lo idóneo es que el investigador emplee algún tipo de variación exógena en el número de alumnos por aula, puesto que el tamaño de las clases es un factor endógeno en el que participan los propios centros y que tiene efectos en el alumnado que opta por un colegio u otro. Por ejemplo, podría darse el caso de que los centros con una menor concentración de alumnos por aula logren mejores resultados en los exámenes debido a que son muy selectivos o porque disponen de más recursos económicos, lo cual podría en ambos casos repercutir directamente en el rendimiento académico medio del alumnado. De la misma manera, si se asignan a propósito clases de menor tamaño a los programas dirigidos a determinados subgrupos de estudiantes, la relación entre el número de alumnos por aula y las calificaciones del alumnado en el entorno académico podría ser un simple reflejo de la selección del alumnado admitido en dichos programas (por ejemplo, cursos de nivel avanzado o programas de apoyo académico). Si los investi-

gadores o responsables políticos observan una correlación entre el número de alumnos por aula y el rendimiento académico del alumnado y lo consideran erróneamente un indicativo de que existe una relación causa-efecto entre ambas variables, esta incorrecta interpretación podría llevar a decisiones normativas equivocadas. Por ello, los trabajos más rigurosos sobre la cuestión suelen buscar variaciones exógenas en el número de alumnos por aula; esto es, variaciones en el tamaño de las clases que no guarden relación con factores que puedan repercutir en el rendimiento académico.

Un influyente estudio en esta línea de investigación evalúa el conocido proyecto STAR del estado de Tennessee, cuyas siglas en inglés significan «relación de rendimiento alumno/profesor» (student/teacher achievement ratio), que puso en marcha un experimento a gran escala en el que alumnos y profesores de educación infantil fueron asignados aleatoriamente a aulas de distintos tamaños y evaluados a lo largo del tiempo. Gracias a su diseño experimental y al gran tamaño de la muestra, este estudio se ha considerado un importante referente sobre los efectos de la reducción del número de alumnos por aula. Los datos sobre el impacto de esta política muestran que los resultados en los exámenes estandarizados fueron mejores en las clases de menor tamaño y que el beneficio fue mayor entre grupos desfavorecidos y con menos recursos económicos (Krueger, 1999). Asimismo, se ha observado que los efectos de este programa pueden ser duraderos. En este sentido, existe evidencia de que asistir a clases menos numerosas a edades tempranas incrementó

las probabilidades de que los alumnos pertenecientes a minorías se presentaran a pruebas de acceso a la universidad (Krueger y Whitmore, 2001).

También existen trabajos que estiman el efecto causal de las clases menos numerosas utilizando escenarios cuasiexperimentales. Estos trabajos llegan a conclusiones similares, ya que observan efectos positivos en las clases menos numerosas, particularmente para alumnos desfavorecidos (por ejemplo, Angrist y Lavy, 1999; Rivkin, Hanushek y Kain, 2005). No obstante, también existe evidencia sobre casos en los que la reducción del número de alumnos por aula en centros educativos (como es el caso del estado de Connecticut, estudiado en Hoxby [2000]) no parece tener un efecto causal.

Todos estos estudios tienen en común que aíslan el efecto causal del número de alumnos por aula en el contexto de estudio en cuestión. Es decir, analizan el impacto del tamaño de las clases *manteniendo todo lo demás constante*. Son investigaciones relevantes, puesto que pueden atribuir claramente sus efectos estimados exclusivamente al número de alumnos por aula. Sin embargo, en la aplicación real de tales políticas, esta estimación causal no es suficiente para saber qué repercusiones tendrá una potencial política en el alumnado, puesto que es muy improbable que el resto de factores educativos permanezcan invariables.

Un ejemplo ilustrativo de cómo la implementación a gran escala puede producir resultados contrarios a los documentados en estudios experimentales es la evaluación de Jepsen y Rivkin

(2009) del programa Class Size Reduction (CSR, «Reducción del número de alumnos por aula») en California. La implementación de un programa a gran escala de reducción del número de alumnos por clase implicará necesariamente la contratación de más profesorado y dicho efecto en el mercado laboral docente puede causar consecuencias no intencionadas. La iniciativa CSR de California, implementada en 1996, redujo el número de alumnos de 30 a 20 en las clases de dicho estado desde educación infantil hasta tercero de primaria. Jepsen y Rivkin (2009) observan que esta política aumentó ligeramente las calificaciones obtenidas en matemáticas y lectura, aunque los efectos de la reducción del número de alumnos por clase se vieron contrarrestados en gran medida por la disminución de la calidad del profesorado. El motivo de ello es que tal reducción a gran escala de alumnos por clase requiere una repentina contratación, también a gran escala, de personal docente, lo cual llevó a contratar profesores sin la debida titulación ni experiencia docente previa.

Detrás de estos resultados inesperados del programa CSR, se encuentra el hecho de que *no todos los demás factores se mantienen constantes en un contexto normativo a gran escala*. En este caso, disminuyó el número de alumnos por aula, pero también lo hicieron las cualificaciones del personal docente. Este es un importante efecto secundario de la política a tener en cuenta. En este sentido, Hanushek (2011) demuestra que el *efecto profesor* es enorme en el rendimiento académico y que la calidad docente es posiblemente uno de los aspectos más importantes para el aprendizaje del alumnado. El mencionado trabajo de

Hanushek puso de manifiesto que «reemplazar entre un 5 y un 8 por 100 del personal docente de peor calidad por profesores de calidad media podría posicionar a Estados Unidos cerca de los mejores puestos en los ránquines internacionales de matemáticas y ciencias, lo que supondría un valor actual de 100 billones de dólares» (Hanushek, 2011: p. 466). Comprender que en la práctica existe un equilibrio al menos a corto plazo entre la calidad docente y las clases menos numerosas pone de manifiesto que: i) apoyar la implementación de una política en los efectos causales que se han estimado asumiendo la condición *ceteris paribus* tan solo fundamenta parcialmente las decisiones normativas; y que ii) en los casos donde los efectos en el profesorado son incluso mayores que los efectos de la reducción del número de alumnos, se podría observar que una política bien intencionada acabe resultando contraproducente.

Cabe añadir que en el estudio de Jepsen y Rivkin (2009) acerca del programa CSR de California, los autores observaron una mayor proporción de docentes sin experiencia concentrados en los centros educativos con alumnado de estratos socioeconómicos más bajos. Esto se debe a que la contratación se vuelve aún más complicada en los centros más desfavorecidos (por ejemplo, Boyd *et al.*, 2013), por lo que cuando se convoca una gran oferta de plazas, existe tanto la necesidad de contratar a profesores nuevos y menos preparados para cubrir esa demanda, como también la necesidad de sustituir profesorado debido al traslado de algunos docentes de mayor cualificación desde entornos escolares más desfavorecidos

a centros más aventajados en los que haya plazas vacantes.

Otro estudio más reciente acerca de una política a gran escala de reducción del número de alumnos por clase en Florida también muestra resultados desalentadores para una inversión educativa tan elevada (Chingos, 2012). Chingos (2011) revisa la bibliografía acerca de la reducción del número de alumnos por aula y señala que, aunque puedan tener algunos beneficios muy concretos, por lo general, «estas políticas a gran escala también son extremadamente costosas y suponen una pérdida de oportunidades para realizar inversiones educativas más inteligentes» (Chingos, 2011: p.2).

En definitiva, los ejemplos de las políticas de California y Florida ponen de relieve la importancia de comprender las limitaciones de los resultados de un estudio científico a la hora de emplearlos para promover proyectos normativos.

3. Problemas asociados a la doble función de los centros escolares como centros educativos y como servicios de cuidado infantil

Un último ejemplo de la necesidad de prever correctamente los efectos secundarios de determinadas políticas proviene de los trabajos sobre la relación entre el mercado laboral y la doble función que desempeña a nivel familiar la educación de los niños pequeños. Además de proporcionar educación a los niños, el colegio también representa la principal fuente de cuidado infantil durante los días laborables. En consecuencia, las políticas que alteran el horario escolar

también alteran el horario del cuidado de los hijos y la jornada laboral de los padres y madres que trabajan.

Por ejemplo, en un intento de estimar cómo afecta el coste del cuidado de los niños en el mercado laboral de las madres, varios investigadores han considerado los programas de educación temprana como un sustituto de los servicios de guardería privados. En sus estimaciones, estas investigaciones utilizan la variación en la introducción de programas gratuitos de educación infantil y preescolar. Esto permite a los investigadores comparar a mujeres similares, algunas de ellas con la posibilidad de elegir una opción educativa pública gratuita y otras abocadas a pagar guarderías privadas para niños de la misma edad. Estos trabajos muestran que existen efectos positivos en el mercado laboral cuando el cuidado infantil se proporciona a través del sistema educativo, ya que ofrece una opción asequible para aquellas mujeres que se reincorporan al trabajo (por ejemplo, Gelbach 2002; Cascio 2009).

Otro ejemplo relevante son los cambios en el calendario escolar. Los calendarios escolares continuos modifican las fechas de asistencia a clase de los alumnos. Los colegios que adoptan estos calendarios suelen tener el mismo número de días lectivos que los colegios con calendarios normales, pero los redistribuyen a lo largo del año natural (Graves, McMullen y Rouse, 2013). Existen diferentes modelos sobre cómo implementar un calendario escolar continuo, pudiendo variar en la duración de los períodos lectivos y vacacionales alternos y en si todo el alumnado tiene el mismo horario de asistencia (*sin-*

gle-track; turno simple) o si se forman grupos de alumnos que vayan rotando según empiezan y terminan los períodos vacacionales (*multi-track*; turnos múltiples). Este tipo de calendarios suelen estar motivados por los supuestos beneficios académicos que conllevan y porque permiten combatir los problemas de saturación de la capacidad en un centro (véase Graves, McMullen y Rouse, 2013).

Desafortunadamente, los datos sobre el impacto académico no han resultado tan prometedores como se esperaba. En Carolina del Norte, McMullen y Rouse (2012a; 2012b) no observan que el cambio de calendario haya tenido un impacto académico relevante; tan solo constatan algunos efectos positivos en los centros con mayores problemas de aglomeración. Por otro lado, los datos de California indican que estos calendarios pueden afectar negativamente al aprendizaje de los alumnos, lo cual se ve aún más acrecentado en los grupos de alumnos desfavorecidos (Graves, 2010; 2011). Se ha observado que estos estudiantes desfavorecidos son quienes experimentan la mayor pérdida de conocimientos en verano, un fenómeno en el que los alumnos olvidan parte de lo aprendido en el anterior curso académico durante las extensas vacaciones de verano y en consecuencia a veces necesitan «ponerse al día» durante el primer trimestre del siguiente curso (por ejemplo, Downey, Von Hippel y Broh, 2004). Los responsables políticos que promueven calendarios académicos continuos basándose en posibles beneficios académicos derivados de acortar las vacaciones no suelen tener en cuenta que el hecho de que las vacaciones estivales no des-

aparecen en un calendario continuo, sino que se redistribuyen en descansos más frecuentes a lo largo del curso. La cuestión pertinente de cara a la formación de los alumnos es, por tanto, qué resulta más beneficioso para el aprendizaje: un único descanso o parones y comienzos más frecuentes. Y, en este sentido, parece que la evidencia no es favorable a los calendarios continuos.

La política de alterar los calendarios escolares, del mismo modo que se observó en las investigaciones sobre la reducción del número de alumnos por aula, pone de manifiesto un ejemplo de interpretación errónea de lo que puede aportarnos la información ofrecida por estudios científicos para fundamentar futuras decisiones normativas (por ejemplo, calendarios escolares). Sin embargo, también representa un claro ejemplo de que ignorar la función que desempeñan los centros educativos en el cuidado infantil puede producir efectos adversos no intencionados. En este sentido, Graves (2013a) pone de manifiesto que los calendarios escolares continuos impactan de manera negativa en el trabajo materno y que las madres que tienen algún hijo en un centro público con calendario continuo y también hijos más pequeños son quienes experimentan mayores dificultades para adaptarse. Este dato no sorprende si se tiene en cuenta que es más fácil encontrar servicios de cuidado infantil diferentes a los centros escolares durante los meses de verano, gracias a los campamentos y a que los alumnos de institutos y universidades, que aportan mano de obra a los campamentos, no tienen clase en verano. En cambio, si alguien intenta encontrar opciones para cuidar a sus hijos durante unas

vacaciones cortas en otros períodos no dispondrá de esas opciones. Graves (2103b) observa que los efectos en el trabajo materno son mayores en las madres blancas que en las minorías étnicas, lo cual la autora atribuye a que las minorías dependen en mayor medida de la familia para el cuidado de los hijos. La autora señala que «la familia quizá se adapte mejor a una distribución diferente de las necesidades de cuidado infantil, lo que permitiría a las madres pertenecientes a grupos minoritarios contrarrestar algunos (aunque no todos) de los efectos negativos sobre el empleo» (p. 279).

La conclusión de que los horarios escolares repercuten en la capacidad de las madres para trabajar no es exclusiva de los estudios sobre calendarios académicos. Como hemos mencionado antes, la implantación de programas públicos de educación infantil influye positivamente en el empleo de las madres (Gelbach, 2002; Cascio, 2009). Otras políticas como la ampliación de los programas de educación infantil a tiempo completo en lugar de a media jornada (por ejemplo, Gibbs, 2014) y los cambios en los días de la semana en que los niños asisten a clase (Duchini y Van Effenterre, 2020) también han influido en el mercado laboral de las madres. No cabe duda de que modificar la distribución del calendario escolar, ya sea a lo largo del año, de la semana o del día, supone un cambio de disponibilidad en cuanto al cuidado de los niños. Por consiguiente, si la sociedad desea facilitar la inserción laboral de la mujer, la distribución (días y horas) del calendario escolar debe ajustarse lo mejor posible al calendario y horarios propios del mundo laboral.

IV. CONCLUSIONES APLICABLES A POLÍTICAS ACTUALES Y FUTURAS

En este trabajo hemos ilustrado, con ejemplos, varios mecanismos comunes que provocan que las políticas públicas tengan resultados no intencionados. Obviamente, hemos proporcionado una muestra de ejemplos que no abarca todos los posibles mecanismos que podrían producir consecuencias no intencionadas de las políticas públicas.

Indudablemente, una recomendación que se deduce de nuestro trabajo es que resulta necesario prestar atención a las consecuencias no intencionadas de las nuevas políticas que se planteen teniendo en cuenta las políticas actuales. Un ejemplo al que cabe aplicar esta recomendación son las medidas para hacer frente a la pandemia mundial producida por la COVID-19. Las medidas de confinamiento y distanciamiento físico son, sin duda, políticas de salud pública que tienen el objetivo expreso de frenar los contagios para salvar vidas. No obstante, numerosas publicaciones recientes ya han constatado que también provocan muchos efectos secundarios que también convendría tener en cuenta. Por ejemplo, se ha demostrado que las medidas para frenar la COVID-19 mediante el confinamiento y el distanciamiento físico han provocado diversos efectos negativos en la salud mental, entre ellos depresión, ansiedad y un nivel de satisfacción con la vida cambiante a lo largo de las diversas olas de la pandemia, lo cual se observa en mayor medida entre la población joven y las mujeres (por ejemplo, Banks y Xu, 2020; Banks, Fancourt y Xu, 2021). En gene-

ral, es posible que el impacto en la salud mental de las políticas relacionadas con la COVID-19 haya acentuado las desigualdades de salud mental que ya existían por motivos de género, edad o étnicos (por ejemplo, Proto y Quintana-Domeque, 2021).

También se han observado efectos secundarios de las medidas de confinamiento y distanciamiento físico en el empleo y la educación. Se han documentado desigualdades raciales y de género en el desempleo (por ejemplo, Stevenson, 2020), así como también efectos sobre el reparto entre hombres y mujeres del cuidado de los hijos y las tareas domésticas, poniéndose de manifiesto que las mujeres que viven en entornos rurales han experimentado un aumento de la brecha de género tanto en el empleo como en términos de trabajo no remunerado, agudizando el fenómeno de la «doble jornada» (por ejemplo, Sevilla y Smith, 2020; Farré *et al.*, 2021). Asimismo, como los centros educativos han permanecido cerrados en algunos países en varias fases de la pandemia, la educación ha pasado a realizarse en el hogar o a distancia; y los datos preliminares muestran que esto también repercute en el aprendizaje de los estudiantes (Chatterji y Li, 2021; Musaddiq *et al.*, 2021).

No cabe duda de que estos ejemplos no aportan una visión completa de los efectos secundarios provocados por las medidas de política pública para frenar el avance de la COVID-19 (por ejemplo, cambios en los hábitos de consumo, impacto medioambiental, etc.), aunque sí que ilustran una esfera de acción actualmente en desarrollo donde investigadores y responsables políticos han de reflexionar dete-

nidamente acerca de todas las repercusiones que pueden tener estas medidas. También es importante destacar que aún es pronto para extraer conclusiones definitivas de muchas de las investigaciones en curso. Gran parte de la bibliografía documenta cambios, aunque no es capaz de identificar un efecto causal. En los casos en los que puede establecerse un efecto causal, este ha de interpretarse con cautela, puesto que obviamente no solo determina el impacto de una política, sino más bien el efecto conjunto de la medida política y también de las consecuencias directas de simplemente vivir en una pandemia. Por ejemplo, aunque las medidas de confinamiento por la COVID-19 tuvieron efectos en el empleo, no se sabe con certeza cuánto se habría resentido el mercado laboral si los trabajadores hubieran decidido abandonar sus puestos por miedo a la pandemia en caso de no haberse decretado un confinamiento. En las investigaciones a menudo se comparan los resultados del período previo a la pandemia y al confinamiento con los resultados durante el período tanto de pandemia como de confinamiento.

En definitiva, una de las principales contribuciones del presente trabajo consiste en poner de relieve que los diversos mecanismos descritos en este artículo pueden ser tenidos en cuenta para mejorar el diseño de futuros proyectos normativos, tanto en relación con la pandemia como en otras áreas. No obstante, es igualmente importante comprender y reconocer que muchos de los efectos secundarios de políticas analizados en este artículo pasaron desapercibidos para los investigadores durante varios años tras la entrada en vigor de aquellas. Este hecho pone de manifiesto

la importancia de realizar evaluaciones periódicas de las políticas públicas, así como la necesidad de establecer alianzas entre investigadores y responsables políticos en las primeras etapas de su implementación con el fin de poder identificar y rectificar cualquier consecuencia desfavorable en una fase temprana del proceso.

NOTA

(*) Los autores agradecen el apoyo financiero de la Línea de Estímulo de Jóvenes Doctores a través del Convenio Plurianual entre la Comunidad de Madrid y la Universidad Autónoma de Madrid (referencia: S11-PJI-2019-00326).

BIBLIOGRAFÍA

- ADAMS, S. y COTTI, C. (2008). Drunk driving after the passage of smoking bans in bars. *Journal of Public Economics*, 92(5-6), pp. 1288-1305.
- ADDA, J. y CORNAGLIA, F. (2010). The Effect of Bans and Taxes on Passive Smoking. *American Economic Journal: Applied Economics*, 2(1), pp. 1-32.
- ANDERSON, M. D., HANSEN, B. y REES, D. I. (2013). Medical marijuana laws, traffic fatalities, and alcohol consumption. *The Journal of Law and Economics*, 56(2), pp. 333-369.
- ANGER, S., KVASNICKA, M. y SIEDLER, T. (2011). One last puff? Public smoking bans and smoking behavior. *Journal of Health Economics*, 30(3), pp. 591-601.
- ANGRIST, J. D. y LAVY, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*, 114(2), pp. 533-575.
- BANKS, J., FANOURT, D. y XU, X. (2021). Mental health and the COVID-19 pandemic. *World Happiness Report 2021*, pp. 107-130.
- BANKS, J. y XU, X. (2020). The Mental Health Effects of the First Two Months of Lockdown during the COVID-19 Pandemic in the UK. *Fiscal Studies*, 41(3), pp. 685-708.
- BOYD, D., LANKFORD, H., LOEB, S. y WYCKOFF, J. (2013). Analyzing the determinants of the matching of public school teachers to jobs: Disentangling the preferences of teachers and employers. *Journal of Labor Economics*, 31(1), pp. 83-117.
- CARPENTER, C. y DOBKIN, C. (2011). The minimum legal drinking age and public health. *Journal of Economic Perspectives*, 25(2), pp. 133-156.
- CARPENTER, C., POSTOLEK, S. y WARMAN, C. (2011). Public-place smoking laws and exposure to environmental tobacco smoke (ETS). *American Economic Journal: Economic Policy*, 3(3), pp. 35-61.
- CARPENTER, C. S. y STEHR, M. (2011). Intended and unintended consequences of youth bicycle helmet laws. *The Journal of Law and Economics*, 54(2), pp. 305-324.
- CORNELSEN, L., GREEN, R., DANGOUR, A. y SMITH, R. (2015). Why fat taxes won't make us thin. *Journal of Public Health*, 37(1), pp. 18-23.
- CASCIO, E. U. (2009). Maternal labor supply and the introduction of kindergartens into American public schools. *Journal of Human Resources*, 44(1), pp. 140-170.
- CHAKRABARTI, R. (2013). Vouchers, public school response, and the role of incentives: Evidence from Florida. *Economic Inquiry*, 51(1), pp. 500-526.
- CHATTERJI, P. y LI, Y. (2021). Effects of COVID-19 on school enrollment. *Economics of Education Review*, 83, p. 102128.
- CHINGOS, M. M. (2011). *The False Promise of Class-Size Reduction*. Center for American Progress.
- CHINGOS, M. M. (2012). The impact of a universal class-size reduction policy: Evidence from Florida's statewide mandate. *Economics of Education Review*, 31(5), pp. 543-562.
- CLOTFELTER, C., LADD, H., VIGDOR, J. y DÍAZ, R. (2004). Do school accountability systems make it more difficult for low performing schools to attract and retain high quality teachers? *Journal of Policy Analysis and Management*, 23(2), pp. 251-271.
- CROST, B. y GUERRERO, S. (2012). The effect of alcohol availability on marijuana use: Evidence from the minimum legal drinking age. *Journal of Health Economics*, 31(1), pp. 112-121.
- CULLEN, J. y REBACK, R. (2006). Tinkering towards accolades: School gaming under a performance accountability system. En T. GRONBERG y D. JANSEN (eds.), *Advances in Applied Microeconomics*, 14 (pp. 1-34). Emerald Group Publishing Limited, Improving School Accountability: Check-ups or Choice.
- DE CHAISEMARTIN, C., GEOFFARD, P. Y. y LE FAOU, A. L. (2011). Workplace smoking ban effects on unhappy smokers. *Health Economics*, 20(9), pp. 1043-1055.
- DINARDO, J. y LEMIEUX, T. (2001). Alcohol, marijuana, and American youth: the unintended consequences of government regulation. *Journal of Health Economics*, 20(6), pp. 991-1010.
- DOWNEY, D. B., VON HIPPEL, P. T. y BROH, B. A. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review*, 69(5), pp. 613-635.
- DUCHINI, E. y VAN EFFENTERRE, C. (2020). School Schedule and the Gender Pay Gap. *IZA Discussion Papers*, 13791.
- FARRÉ, L., FAWAZ, Y., GONZÁLEZ, L. y GRAVES, J. (2021). Gender inequality in paid and unpaid work during COVID-19 times. *Review of Income and Wealth*. Forthcoming.
- FIGLIO, D. (2003). Fiscal implications of school accountability initiatives. *Tax Policy and the Economy*, 17, pp. 1-36.
- FIGLIO, D. (2006). Testing, crime and punishment. *Journal of Public Economics*, 90(4-5), pp. 837-851.
- FIGLIO, D. y GETZLER, L. (2007). Accountability, ability and disability: Gaming the system? En T. GRONBERG, D. JANSEN (eds.), *Advances in Applied Microeconomics*, 14, pp. 35-49. Emerald Group Publishing Limited, Improving School Accountability: Check-ups or Choice.
- FIGLIO, D. y LOEB, S. (2011). School Accountability. En A. ERIC, HANUSHEK, STEPHEN MACHIN y LUDGER WOESSMANN (ed.), *Handbooks in Economics*, Vol. 3. The Netherlands: North-Holland, pp. 383-421.

<p>GELBACH, J. B. (2002). Public schooling for young children and maternal labor supply. <i>American Economic Review</i>, 92(1), pp. 307-322.</p> <p>GIBBS, C. R. (2014). A matter of time? Impact of statewide full-day kindergarten expansions on later academic skills and maternal employment. <i>The Department of Labor 2013-2014 Scholars Program</i>, p. 2.</p> <p>GRAVES, J. (2010). The Academic Impact of Multi-Track Year-Round School Calendars: A Response to School Overcrowding. <i>Journal of Urban Economics</i>, 67, pp. 378-391.</p> <p>GRAVES, J. (2011). Effects of Year-Round Schooling on Disadvantaged Students and the Distribution of Standardized Test Performance. <i>Economics of Education Review</i>, 30, pp. 1281-1305.</p> <p>GRAVES, J. (2013a). School Calendars, Child Care Availability and Maternal Employment. <i>Journal of Urban Economics</i>, 78, pp. 57-70.</p> <p>GRAVES, J. (2013b). The Effects of School Calendar Type on Maternal Employment across Racial Groups: A Story of Child Care Availability. <i>American Economic Review</i>, 103, pp. 279-283.</p> <p>GRAVES, J., McMULLEN, S. y ROUSE, K. (2013). Multi-Track Year-Round Schooling as Cost Savings Reform: Not Just a Matter of Time. <i>Education Finance and Policy</i>, 8, pp. 300-315.</p> <p>HANUSHEK, E. A. (2011). The economic value of higher teacher quality. <i>Economics of Education Review</i>, 30(3), pp. 466-479.</p> <p>HEISSEL, J. y LADD, H. (2018). School turnaround in North Carolina: A regression discontinuity analysis. <i>Economics of Education Review</i>, 62, pp. 302-320.</p> <p>HOXBY, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. <i>The Quarterly Journal of Economics</i>, 115(4), pp.1239-1285.</p> <p>JACOB, B. A. y LEVITT, S. D. (2003). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. <i>Quarterly Journal of Economics</i>, 118(3), pp. 843-878.</p>	<p>JEPSEN, C. y RIVKIN, S. (2009). Class size reduction and student achievement: The potential tradeoff between teacher quality and class size. <i>Journal of Human Resources</i>, 44(1), pp. 223-250.</p> <p>JONES, A. M., LAPORTE, A., RICE, N. y ZUCHELLI, E. (2015). Do public smoking bans have an impact on active smoking? Evidence from the UK. <i>Health Economics</i>, 24(2), pp. 175-192.</p> <p>KORETZ, D. M. y BARRON, S. (1998). <i>The validity of gains on the Kentucky Instructional Results Information System (KIRIS)</i>. Santa Monica, CA: RAND Corporation, 1998.</p> <p>KRUEGER, A. B. (1999). Experimental estimates of education production functions. <i>The Quarterly Journal of Economics</i>, 114(2), pp. 497-532.</p> <p>KRUEGER, A. B. y WHITMORE, D. M. (2001). The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from Project STAR. <i>The Economic Journal</i>, 111(468), pp. 1-28.</p> <p>LADD, H. y WALSH, R. (2002). Implementing value-added measures of school effectiveness: Getting the incentives right. <i>Economics of Education Review</i>, 21(1), pp. 1-17.</p> <p>LADD, H. y ZELLI, F. (2002). School-based accountability in North Carolina: The responses of school principals. <i>Educational Administration Quarterly</i>, 38(4), pp. 494-529.</p> <p>LINN, R. L. (2000). Assessments and accountability. <i>Educ. Researcher</i>, 29(2), pp. 4-16.</p> <p>MULLER, L., LACROIX, A., LUSK, J. L. y RUFFIEUX, B. (2017). Distributional impacts of fat taxes and thin subsidies. <i>The Economic Journal</i>, 127(604), pp. 2066-2092.</p> <p>McMULLEN, S. C. y ROUSE, K. E. (2012a). School Crowding, Year-Round Schooling and Mobile Classroom Use: Evidence from North Carolina. <i>Economics of Education Review</i>, 31, pp. 812-823.</p> <p>McMULLEN, S. C. y ROUSE, K. E. (2012b). The Impact of Year-Round Schooling</p>	<p>on Academic Achievement: Evidence from Mandatory School Calendar Conversions. <i>American Economic Journal: Economic Policy</i>, 4, pp. 230-252.</p> <p>MUSADDIQ, T., STANGE, K. M., BACHER-HICKS, A. y GOODMAN, J. (2021). The Pandemic's effect on demand for public schools, homeschooling, and private schools. <i>Working Paper</i>, 29262. National Bureau of Economic Research.</p> <p>PROTO, E. y QUINTANA-DOMEQUE, C. (2021). COVID-19 and mental health deterioration by ethnicity and gender in the UK. <i>PLoS one</i>, 16(1), p.e0244419.</p> <p>RIVKIN, S. G., HANUSHEK, E. A. y KAIN, J. F. (2005). Teachers, schools, and academic achievement. <i>Econometrica</i>, 73(2), pp. 417-458.</p> <p>ROCKOFF, J. E., STAIGER, D. O., KANE, T. J. y TAYLOR, E. S. (2012). Information and employee evaluation: Evidence from a randomized intervention in public schools. <i>American Economic Review</i>, 102(7), pp. 3184-3213.</p> <p>SEVILLA, A. y SMITH, S. (2020). Baby steps: the gender division of childcare during the COVID-19 pandemic. <i>Oxford Review of Economic Policy</i>, 36(Supplement_1), pp. S169-S186.</p> <p>STECHEK, B., BARRON, S., CHUN, T. y ROSS, K. (2000). <i>The Effects of the Washington State Education Reform on Schools and Classrooms</i>. Center for Research on Evaluation, Standards and Student Testing.</p> <p>STEVENSON, B. (2020). The initial impact of COVID-19 on labor market outcomes across groups and the potential for permanent scarring. <i>The Hamilton Project</i>. Washington, DC: Brookings Institution.</p> <p>XU, X. y CHALOUKKA, F. J. (2011). The effects of prices on alcohol use and its consequences. <i>Alcohol Research & Health</i>, 34(2), p. 236.</p> <p>WADA, R., CHALOUKKA, F. J., POWELL, L. M. y JERNIGAN, D. H. (2017). Employment impacts of alcohol taxes. <i>Preventive Medicine</i>, 105, pp. S50-S55.</p>
---	---	---