



**Publicación de los resultados de las pruebas estandarizadas externas:  
¿Tiene ello un efecto sobre los resultados escolares?**

by

**Brindusa Anghel\***  
**Antonio Cabrales\*\***  
**Jorge Sainz\*\*\***  
**Ismael Sanz\*\*\*\***

**Documento de Trabajo 2013-07**

September 2013

\* FEDEA  
\*\* Universidad Carlos III de Madrid.  
\*\*\* Universidad Rey Juan Carlos.  
\*\*\*\* Universidad Rey Juan Carlos.

---

Los Documentos de Trabajo se distribuyen gratuitamente a las Universidades e Instituciones de Investigación que lo solicitan. No obstante están disponibles en texto completo a través de Internet: <http://www.fedea.es>.

These Working Paper are distributed free of charge to University Department and other Research Centres. They are also available through Internet: <http://www.fedea.es>.

ISSN:1696-750

# **PUBLICACIÓN DE LOS RESULTADOS DE LAS PRUEBAS ESTANDARIZADAS EXTERNAS: ¿TIENE ELLO UN EFECTO SOBRE LOS RESULTADOS ESCOLARES?**

**Brindusa Anghel**

FEDEA

**Antonio Cabrales**

Universidad Carlos III de Madrid

**Jorge Sainz**

Universidad Rey Juan Carlos

**Ismael Sanz**

Universidad Rey Juan Carlos

Diciembre 2012

## **Resumen**

Estudiamos el efecto de pruebas estandarizadas externas sobre los resultados académicos de los estudiantes. Explotamos el hecho de que sólo una de las 17 regiones españolas empezó realizando y publicando los resultados de las pruebas estandarizadas en 2005 para aplicar una metodología de diferencias en diferencias, usando los resultados del estudio de PISA de 2000 a 2009. A continuación, confirmamos lo obtenido usando métodos de control sintéticos. La utilización de datos de un único país nos permite minimizar los sesgos derivados de diferencias en marcos legales, entornos sociales o culturales. Nuestro análisis econométrico aporta credibilidad a la hipótesis de que este tipo de prueba mejora significativamente los resultados académicos. Una novedad clave es que nuestros exámenes no tienen consecuencias académicas para los alumnos, por lo tanto, los efectos tienen que venir directamente del impacto en los profesores y administradores.

Palabras clave: pruebas externas y estandarizadas, PISA, diferencias en diferencias, métodos de control sintéticos.

# 1. INTRODUCCIÓN

Las pruebas estandarizadas externas permiten a la administración supervisar mejor el proceso de educación y el resultado de los colegios. En la mayoría de los países que tienen estas pruebas, los resultados del examen son públicos y pueden ser usados por los padres para tomar decisiones. Esta supervisión más cercana por parte de padres y administradores proporciona una motivación adicional para maestros y directores para mejorar los resultados escolares de sus estudiantes. Este potencial de mejora ha animado a un número de países cada vez mayor a usar los exámenes externos como una herramienta para aumentar la rendición de cuentas en los colegios. El informe de PISA (Programa para la Evaluación de los Estudiantes Internacionales) (OCDE, 2010) recoge el hecho de que 22 de los 34 países de la OCDE han introducido exámenes externos estandarizados en la mayoría de sus colegios. Dos países más, Alemania y EE.UU., tienen este tipo de exámenes sólo en algunos de sus “Ländern” y estados. En conjunto, dos terceras partes de los estudiantes de 15 años de la OCDE asisten a escuelas donde hay una prueba externa y estandarizada.

La evidencia empírica existente apoya la hipótesis de que los países con sistemas de exámenes finales externos tienen un mejor comportamiento en las pruebas internacionales de rendimiento de los estudiantes. La primera evidencia de esto fue dada por Bishop (1997) para los estudiantes que realizaron las pruebas IAEP en 1991 de matemáticas, ciencias y geografía y por Bishop (2006) con los resultados de PISA del año 2000. En general, la evidencia existente entre países sugiere que el efecto de los exámenes finales externos en el rendimiento estudiantil podría ser la mitad o más de un nivel equivalente, o entre 20 y 40 por ciento de una desviación estándar de

<sup>1</sup> Incluso en los países donde las zonas escolares comprenden una sola escuela, los padres afectados pueden decidir donde vivir según la calidad de la escuela como un factor para su elección.

las respectivas pruebas internacionales (OECD, 2010 y 2012 y Hanushek y Woessmann, 2011).

Esta evidencia ha sido criticada por dos motivos. Primero, estos estudios utilizan datos de corte transversal y por lo tanto la introducción de un examen estandarizado por país es endógena, y la heterogeneidad no observable podría sesgar los resultados. Segundo, la introducción de pruebas externas puede dar lugar a “enseñar para el examen”. Sin embargo, algunos estudios han encontrado la misma asociación positiva entre exámenes centrales y los logros de estudiantes dentro de países donde algunas regiones cuentan con sistemas de exámenes externos y otras donde no los tienen<sup>2</sup>. Esta evidencia excluye la posibilidad de que factores a nivel nacional no observables correlacionados con la existencia de los exámenes conducen a una correlación positiva observable entre esos exámenes y los rendimientos de los estudiantes. Además, se ha encontrado que los estudiantes en países con exámenes nacionales externos han alcanzado mejores resultados en otras pruebas internacionales tales como PISA, PIRLS o TIMMS. En la medida en que esas pruebas son de naturaleza distinta a las nacionales, esto puede descartar el hecho de que “enseñar para la prueba” es un factor principal que lleva a unos mejores resultados de los estudiantes en países o regiones con pruebas nacionales externas.

Una dificultad diferente presente en estudios anteriores es que no está muy claro cuales son los canales por medio de los cuales las pruebas finales son efectivas. Esto es porque en la mayoría de los casos estos exámenes tienen consecuencias académicas para los estudiantes, así pues, proporcionan razones para la mejora tanto de los profesionales como de los estudiantes. El presente estudio utiliza una característica especial del sistema educativo español para separar los incentivos del colegio y de los estudiantes, mientras al mismo tiempo se controlan los sesgos que provienen de la heterogeneidad no observable a nivel nacional y posiblemente también del hecho de “enseñar para la prueba”.

El rasgo especial al cual nos referimos antes es que la principal ley de educación española (Ley Orgánica de la Educación, LOE 2006) permite a las

<sup>2</sup> Ver, e.g. Bishop (1997) para provincias canadienses, Jürges et al. (2005) y Woessman (2010) para Alemania, y Bishop et al. (2001), para EE.UU.

regiones llevar a cabo evaluaciones del sistema educativo siempre y cuando los resultados no se utilicen para evaluar a los estudiantes o clasificar colegios (artículo 140). Esto significa que los exámenes españoles no son “Curriculum-Based External Exit Examination (CBEEE)” como definido por Bishop (1997) que significa examen externo final basados en el plan de estudios, así pues dichos exámenes deberían “ofrecer señales de los logros del alumno que tengan consecuencias reales para el alumno y definen los avances de acuerdo a un estándar externo, no en relación con otros estudiantes en el aula o en el colegio”. Esto significa que los efectos de tales exámenes en España, en su caso, deban notarse solo directamente en cambios en incentivos de los colegios, aunque al final probablemente puedan tener un impacto en el esfuerzo de los estudiantes.

La Comunidad de Madrid introdujo una prueba estándar externa, llamada “prueba de Conocimientos y Destrezas Indispensables” (también conocida por sus siglas como prueba CDI) en el año académico 2004/05. La nota alcanzada por los estudiantes en este examen no tiene “consecuencias académicas reales” para la mayoría de los estudiantes, por lo que no puede ser considerada un CBEEE<sup>3</sup>. De modo que los efectos de esta iniciativa irán necesariamente de forma directa a través sólo de cambios en la motivación de los profesores. La Comunidad de Madrid es también la única que publica y pone a disposición de la opinión pública los resultados medios sobre la prueba externa en cada una de las escuelas<sup>3</sup>.

Todas las regiones en España operan bajo el mismo marco legal que regula los principios, objetivos y la organización de los diferentes niveles escolares (pre-primaria, primaria, secundaria obligatoria, post secundaria obligatoria), además de aumentar al 65% (55% en regiones históricas) los contenidos y asignaturas estudiadas. Por lo tanto, junto con el importe de la financiación pública de los colegios, los cuales podemos controlar, la otra

<sup>3</sup> Un estudiante con buenas notas en la educación secundaria obligatoria y una buena nota en la prueba CDI obtiene un certificado con mérito o con distinción, en lugar de solo un certificado, pero esto no tiene ninguna implicación para admisiones a escuelas después de la educación obligatoria, o para becas, tampoco hay evidencia de que los directores tengan en cuenta esas distinciones. Los estudiantes con notas realmente extraordinarias (solo 25 por año en una región con más de 50.000 estudiantes en el año pasado en educación secundaria obligatoria) pueden obtener un Premio Extraordinario que consiste en un premio en metálico de 1.000 euros y un viaje a un “destino cultural”.

<sup>4</sup> Otras regiones tienen pruebas externas estandarizadas donde todos los colegios son evaluados, pero Madrid es la única que publica los resultados.

principal diferencia observable en educación entre regiones españolas es la aparición en el período de estudio de este examen estandarizado externo en Madrid cuyos resultados son publicados.

Esta característica nos permite llevar a cabo un análisis de diferencias en diferencias (“diff-in-diff”) comparando los resultados de PISA de la región analizada (Madrid) antes y después de que la prueba CDI fuera introducida en el resto de Comunidades españolas antes y después del análisis. Este enfoque de diferencias en diferencias nos permite controlar los factores no observables invariantes en el tiempo que afectan a Madrid. Al trabajar con regiones del mismo país también excluimos algunos efectos no observables que aparecen en estudios entre países con diferentes legislaciones y culturas.

El hecho de que se trata de un solo país también nos permite aplicar los nuevos métodos de inferencia de control sintético para estudios de casos comparativos propuestos por Abadie y Gardeazabal (2003) y Abadie et al. (2010). Usamos una combinación de otras regiones españolas para construir una región de control sintético, la cual imita las características importantes de la educación de Madrid antes de la introducción de la prueba CDI. La evolución resultante de la educación subsiguiente a esta situación hipotética en Madrid sin CDI es comparada con la experiencia real de Madrid. La idea detrás del enfoque de control sintético es que una combinación de unidades a menudo proporciona una mejor comparación para la unidad expuesta a la intervención que cualquier unidad sola. La transparencia y la protección contra la extrapolación son dos características atractivas del método de control sintético relacionadas con los métodos de regresión tradicionales.

Nuestros resultados también están más protegidos que otros de la crítica de que son el resultado de “enseñar para el examen”. Esto se debe a que nuestra medida del rendimiento, es decir, los resultados en el examen de PISA, tienen en cierto modo objetivos distintos y miden cosas diferentes que el examen CDI en cuyos efectos estamos interesados. Las preguntas del examen CDI de Madrid evalúan el conocimiento y están directamente relacionadas con el material visto en clases de Lengua y Matemáticas durante el curso académico. Por el contrario, las preguntas del examen de PISA (llamado estímulo) están más relacionadas con procesos cognitivos (el dominio de los procesos, el entendimiento de los conceptos y la habilidad de actuar o fun-

cionar en varias situaciones dentro de cada dominio), además de cómo usar el conocimiento en contextos particulares. Esto es, la evaluación de PISA está más relacionada con competencias mientras que el CDI de Madrid está más relacionado con el conocimiento.

El artículo se encuentra organizado de la siguiente manera. La Sección 2 describe con cierto detalle la estructura institucional y la prueba CDI externa y estándar. La Sección 3 describe los datos. La Sección 4 trata sobre la metodología econométrica y contiene los principales resultados del artículo. La Sección 5 muestra los resultados de los métodos de control sintéticos. En la Sección 6 se concluye.

## 2. ESTRUCTURA INSTITUCIONAL

El gobierno de la Comunidad de Madrid ha llevado a cabo desde el año académico 2004/2005 un examen externo estandarizado para todos los estudiantes de sexto grado en la región, quienes están por lo tanto en el último año de escuela primaria (en torno a los 11-12 años de edad). Tres años más tarde, la región introdujo otro examen externo y estandarizado en el noveno grado (el tercer año de la escuela secundaria, el cual es el último año académico común para los estudiantes). Estos exámenes son obligatorios para todos los colegios primarios y secundarios (públicos y privados). El examen mide lo que las autoridades consideran conocimiento básico en matemáticas (ejercicios y problemas) y en lengua (dictado, lectura, conocimiento general y preguntas relacionadas con un texto).

Nuestro objetivo es evaluar si la introducción de estos exámenes ha mejorado los resultados académicos de los estudiantes en Madrid. Utilizamos como medida de rendimiento de los estudiantes las notas de los exámenes realizados por el Programa para la Evaluación Internacional de Estudiantes (PISA) de la OCDE. PISA analiza las capacidades básicas en lectura, matemáticas y ciencias de estudiantes de 15 años en países miembros de la OCDE y países/economías asociadas a través de encuestas trienales. La medida de la escala global de lectura está basada en un promedio de los países de la OCDE establecida en 500 en PISA 2000, con una desviación estándar de 100. PISA llevó a cabo sus primeras pruebas en 2000 considerando la lectura como un área de evaluación importante, y proporcionando un resumen de las habilidades en matemáticas y ciencias. En 2003, las matemáticas fueron el área principal y en 2006 fue la ciencia. En 2009, PISA empezó otro ciclo, centrándose de nuevo en la lectura. Cuando un área es el principal objetivo del examen, dos terceras partes del tiempo del examen está dedicado a dicho área, lo que permite su análisis más profundo. Dado que tanto PISA

2000 como PISA 2009 se centraron en la lectura, es posible obtener comparaciones muy detalladas de cómo el desempeño del estudiante en esa área cambió durante ese tiempo. Comparaciones en el tiempo de las materias de matemáticas y ciencias son algo más limitadas.

En la prueba de PISA, cada estudiante participante tarda dos horas en realizar las tareas con lápiz y papel de lectura, matemáticas y ciencias. La evaluación incluye tareas requiriendo a los estudiantes elaborar sus propias respuestas además de incluir preguntas de elección múltiple. Asimismo, los estudiantes también responden a una encuesta que tardan aproximadamente 30 minutos en completar y que incluye preguntas sobre su entorno personal.

### 3. DESCRIPCIÓN DE LOS DATOS

El primer examen CDI tuvo lugar en el año académico 2004/05, por lo que consideramos éste como el año en el que la evaluación (la introducción de un examen estandarizado) fue por primera vez implementada. Por esta razón, comparamos los resultados de los estudiantes de la región de Madrid en (i) Lectura, utilizando PISA 2000 y PISA 2009 y (ii) en Matemáticas, utilizando PISA 2003 y PISA 2009<sup>5</sup>.

Nuestra primera metodología para el análisis será un método de regresión de diferencias en diferencias. Construimos el tratamiento y los grupos de control de la siguiente forma: el grupo de tratamiento antes del tratamiento (la introducción de la prueba CDI) es el grupo de estudiantes de la Comunidad de Madrid que realizó el examen PISA en 2000 de lectura o en 2003 de matemáticas, el grupo de tratamiento después del tratamiento es el grupo de estudiantes que realizó el examen de PISA en 2009, y el grupo de control está formado por estudiantes de otras regiones de España antes del tratamiento (PISA 2000 o 2003) y después (PISA 2009).

El cuestionario de PISA nos permite controlar por varias características relacionadas con el estudiante, la familia y el colegio. Las características del estudiante y de la familia son: sexo, edad, nacionalidad (inmigrante o español), nacionalidad de los padres, idiomas diferentes al castellano hablados en casa, estructura de la familia (familia monoparental, familia nuclear, familia mixta), tiempo de aprendizaje en horas por semana dedicado a la lectura y las matemáticas (horas a la semana de clases de lengua o matemáticas), índice del estatus económico, social y cultural (índice ESCS) calculado por la

<sup>5</sup>No vamos a utilizar las notas de PISA en ciencias, ya que el primer año en el que la materia de ciencias fue el principal objetivo fue 2006 y esto es después de aplicar nuestro análisis.

OCDE<sup>6</sup>. Las características del colegio son: tipo de escuela (pública, concertada o privada), ubicación de la escuela (pueblo, ciudad pequeña, ciudad, urbe o urbe grande), ratio estudiante/profesor, tamaño del colegio, si el colegio utiliza evaluaciones para comparar con el rendimiento del distrito/nacional, si el colegio utiliza evaluaciones para hacer juicios sobre la efectividad del profesor, proporción de niñas en el colegio, promedio del colegio respecto al índice ESCS, porcentaje de alumnos inmigrantes en la escuela, y el tiempo medio de aprendizaje en la escuela en lectura y matemáticas.

Las siguientes tablas contienen estadísticas descriptivas de estos cuatro grupos, para las características más relevantes de los estudiantes y de los colegios. La Tabla 1 describe el tratamiento y el control de grupos en PISA 2000 y en PISA 2009 en lectura y la Tabla 2 describe los dos grupos en PISA 2003 y en PISA 2009 en matemáticas.

<sup>6</sup> El índice PISA de estatus económico, social y cultural (ESCS) se derivó de los siguientes tres índices: mayor nivel de ocupación de los padres, mayor nivel de educación de los padres en años de enseñanza según ISCED y las posesiones del hogar (OECD, 2010).

Tabla 1: Estadísticas Descriptivas de estudiantes y colegios en PISA 2000 y PISA 2009 (Lectura)

Variable	Grupo de tratamiento antes del cambio - Colegio Madrid en el examen de PISA 2000			Grupo de control antes del cambio - Colegio NO Madrid en el examen de PISA 2000			Grupo de tratamiento después del cambio - Colegio Madrid en el examen de PISA 2009			Grupo de control después del cambio - Colegio NO Madrid en el examen de PISA 2009						
	Media	Desv. Est.	Mín	Máx	Media	Desv. Est.	Mín	Máx	Media	Desv. Est.	Mín	Máx				
<b>Asignaturas - Valores Plausibles</b>																
Lectura - VP1	510.16	79.75	245.12	697.46	491.56	84.70	150.64	777.08	504.38	85.14	74.61	721.54	483.75	89.49	6.65	814.71
Lectura - VP2	509.94	82.20	204.02	711.64	491.47	84.37	151.55	756.66	504.39	85.52	91.12	822.19	483.82	89.90	60.42	824.33
Lectura - VP3	510.29	81.44	237.66	716.00	491.89	84.88	157.91	768.72	503.43	83.68	60.18	733.45	483.84	89.65	19.36	814.71
Lectura - VP4	508.55	81.72	229.34	710.26	491.36	84.87	117.01	751.45	503.58	85.14	80.22	702.07	483.67	90.16	29.15	877.24
Lectura - VP5	508.86	81.96	176.62	709.59	491.23	85.21	112.65	782.35	504.67	84.54	124.31	729.32	483.96	89.75	64.42	904.50
<b>Características individuales</b>																
Hembra	0.51	0.50	0	1	0.51	0.50	0	1	0.50	0.50	0	1	0.49	0.50	0	1
Edad	15.80	0.28	15.33	16.25	15.79	0.28	15.33	16.25	15.87	0.28	15.33	16.33	15.86	0.29	15.33	16.33
Inmigrante	0.03	0.18	0	1	0.02	0.15	0	1	0.16	0.37	0	1	0.09	0.28	0	1
Madre inmigrante	0.07	0.25	0	1	0.04	0.19	0	1	0.21	0.41	0	1	0.12	0.33	0	1
Padre inmigrante	0.05	0.22	0	1	0.03	0.18	0	1	0.20	0.40	0	1	0.11	0.31	0	1
Idioma diferente al castellano hablado en casa	0.01	0.09	0	1	0.17	0.37	0	1	0.05	0.22	0	1	0.16	0.37	0	1
Familia monoparental	0.17	0.37	0	1	0.17	0.37	0	1	0.16	0.37	0	1	0.13	0.34	0	1
Familia nuclear	0.79	0.41	0	1	0.79	0.41	0	1	0.83	0.38	0	1	0.85	0.35	0	1
Familia mixta	0.04	0.21	0	1	0.05	0.21	0	1	0.01	0.09	0	1	0.01	0.11	0	1
Tiempo de estudio (horas/semana) en lengua	3.30	0.74	0.00	4.58	3.04	0.74	0.00	4.33	3.66	0.68	1.67	9.00	3.37	0.70	0.00	9.00
ESCS (índice del nivel económico, social y cultural)	-0.16	1.01	-2.79	2.13	-0.40	1.05	-4.05	2.21	-0.09	1.05	-3.40	2.85	-0.26	1.05	-5.34	3.41
ESCS al cuadrado	1.05	1.14	0.00	7.80	1.26	1.42	0.00	16.42	1.11	1.30	0.00	11.57	1.18	1.44	0.00	28.54
<b>Características del colegio</b>																
Colegio público	0.56	0.50	0	1	0.63	0.48	0	1	0.60	0.49	0	1	0.63	0.48	0	1
Colegio privado	0.16	0.36	0	1	0.08	0.27	0	1	0.08	0.27	0	1	0.03	0.18	0	1
Colegio concertado	0.28	0.45	0	1	0.29	0.46	0	1	0.32	0.47	0	1	0.34	0.47	0	1
Colegio en pueblo	0.00	0.00	0	0	0.02	0.15	0	1	0.02	0.14	0	1	0.06	0.24	0	1
Colegio en pequeña ciudad	0.00	0.00	0	0	0.24	0.42	0	1	0.11	0.32	0	1	0.26	0.44	0	1
Colegio en ciudad	0.18	0.38	0	1	0.34	0.47	0	1	0.27	0.44	0	1	0.35	0.48	0	1
Colegio en urbe	0.36	0.48	0	1	0.36	0.48	0	1	0.16	0.36	0	1	0.32	0.47	0	1

**Tabla 1: Estadísticas Descriptivas de estudiantes y colegios en PISA 2000 y PISA 2009 (Lectura)**

Variable	Grupo de tratamiento antes del cambio - Colegio Madrid en el examen de PISA 2000			Grupo de control antes del cambio - Colegio NO Madrid en el examen de PISA 2000			Grupo de tratamiento después del cambio - Colegio Madrid en el examen de PISA 2009			Grupo de control después del cambio - Colegio NO Madrid en el examen de PISA 2009		
	Media	Desv. Est.	Mín	Máx	Media	Desv. Est.	Mín	Máx	Media	Desv. Est.	Mín	Máx
<b>Asignaturas - Valores Plausibles</b>												
Colegio en una gran urbe	0.47	0.50	0	1	0.04	0.21	0	1	0.44	0.50	0	1
Ratio estudiante/profesor	16.93	4.39	10.90	26.80	14.15	4.66	5.66	27.60	12.76	4.23	1.18	20.27
Tamaño del colegio	1,048.43	493.49	335.00	2,139.00	725.64	345.15	93.00	1,742.00	851.82	416.07	100.00	2,288.00
Evaluación para comparar el colegio con el resultado en el distrito/país	0.32	0.47	0	1	0.18	0.39	0	1	0.50	0.50	0	1
Evaluación para juzgar la efectividad del profesor	0.47	0.50	0	1	0.40	0.49	0	1	0.62	0.48	0	1

Nº de observaciones - colegios	20	165	51	838
Nº de observaciones - estudiantes	679	5535	1453	24434

**Tabla 2: Estadísticas Descriptivas de estudiantes y colegios en PISA 2003 y PISA 2009 (Matemáticas)**

Variable	Grupo de tratamiento antes del cambio - Colegio Madrid en el examen de PISA 2000			Grupo de control antes del cambio - Colegio NO Madrid en el examen de PISA 2000			Grupo de tratamiento después del cambio - Colegio Madrid en el examen de PISA 2009			Grupo de control después del cambio - Colegio NO Madrid en el examen de PISA 2009		
	Media	Desv. Est.	Mín	Máx	Media	Desv. Est.	Mín	Máx	Media	Desv. Est.	Mín	Máx
<b>Asignaturas - Valores Plausibles</b>												
Matemáticas - VP1	493.98	92.37	238.46	706.68	495.25	85.81	122.01	807.17	497.47	88.34	179.79	773.49
Matemáticas - VP2	492.34	91.10	197.49	701.70	495.11	86.17	56.27	788.63	497.52	88.68	205.49	820.23
Matemáticas - VP3	494.73	87.87	216.34	694.69	494.23	85.59	137.67	793.61	496.24	88.22	222.63	749.35
Matemáticas - VP4	495.29	90.48	229.35	731.61	495.14	85.91	117.34	770.01	496.53	87.95	196.22	740.23
Matemáticas - VP5	491.91	92.66	202.94	723.35	494.47	85.62	143.82	777.49	497.89	87.76	193.03	777.39
Individual characteristics	0.50	0.50	0	1	0.51	0.50	0	1	0.50	0.50	0	1

<b>Individual characteristics</b>															
Edad	15.85	0.29	15.25	16.33	15.86	0.29	15.87	0.28	15.33	16.33	15.86	0.29	15.33	16.33	16.33
Inmigrante	0.07	0.25	0	1	0.03	0.17	0	1	0.16	0.37	0	1	0.09	0.28	0
Madre inmigrante	0.12	0.33	0	1	0.04	0.20	0	1	0.21	0.41	0	1	0.12	0.33	0
Padre inmigrante	0.09	0.29	0	1	0.04	0.19	0	1	0.20	0.40	0	1	0.11	0.31	0
Idioma diferente al castellano hablado en casa	0.04	0.19	0	1	0.14	0.35	0	1	0.05	0.22	0	1	0.16	0.37	0
Familia monoparental	0.18	0.38	0	1	0.13	0.33	0	1	0.16	0.37	0	1	0.13	0.34	0
Familia nuclear	0.79	0.41	0	1	0.83	0.38	0	1	0.83	0.38	0	1	0.85	0.35	0
Familia mixta	0.02	0.15	0	1	0.02	0.15	0	1	0.01	0.09	0	1	0.01	0.11	0
Tiempo de estudio (horas/semana) en matemáticas	3.05	0.60	0.92	7.33	2.97	0.73	0.00	15.00	3.20	0.81	1.67	9.00	3.44	0.67	0.00
ESCS (Índice del nivel económico, social y cultural)	-0.12	0.97	-3.74	2.10	-0.20	0.98	-3.40	2.39	-0.09	1.05	-3.40	2.85	-0.26	1.05	-5.34
ESCS al cuadrado	0.94	1.38	0.00	13.96	1.00	1.28	0.00	11.55	1.11	1.30	0.00	11.57	1.18	1.44	0.00
<b>School characteristics</b>															
Colegio público	0.54	0.50	0.00	1.00	0.55	0.50	0.00	1.00	0.60	0.49	0.00	1.00	0.63	0.48	0.00
Colegio privado	0.07	0.26	0	1	0.06	0.24	0	1	0.08	0.27	0	1	0.03	0.18	0
Colegio concertado	0.39	0.49	0	1	0.39	0.49	0	1	0.32	0.47	0	1	0.34	0.47	0
Colegio en pueblo	0.00	0.00	0	0	0.05	0.22	0	1	0.02	0.14	0	1	0.06	0.24	0
Colegio en pequeña ciudad	0.05	0.23	0	1	0.24	0.43	0	1	0.11	0.32	0	1	0.26	0.44	0
Colegio en ciudad	0.23	0.42	0	1	0.33	0.47	0	1	0.27	0.44	0	1	0.35	0.48	0
Colegio en urbe	0.13	0.33	0	1	0.36	0.48	0	1	0.16	0.36	0	1	0.32	0.47	0
Colegio en una gran urbe	0.59	0.49	0	1	0.02	0.15	0	1	0.44	0.50	0	1	0.01	0.12	0
Ratio estudiante/profesor	15.61	5.48	9.44	24.36	13.23	5.31	1.38	44.30	12.76	4.23	1.18	20.27	11.29	4.69	0.82
Tamaño del colegio	880.37	401.45	389	1951	710.70	432.90	95	2819	851.82	416.07	100	2268	685.72	392.01	44
Evaluación para comparar el colegio con el resultado en el distrito/pais	0.12	0.33	0	1	0.26	0.44	0	1	0.50	0.50	0	1	0.28	0.45	0.00
Evaluación para juzgar la efectividad del profesor	0.24	0.42	0	1	0.40	0.49	0	1	0.62	0.48	0	1	0.45	0.50	0.00
<b>Nº de observaciones - colegios</b>	<b>18</b>				<b>365</b>				<b>51</b>				<b>838</b>		
<b>Nº de observaciones - estudiantes</b>	<b>511</b>				<b>10280</b>				<b>1453</b>				<b>24434</b>		

Las dos tablas muestran una evolución muy similar de las características de los alumnos y de los colegios cuando comparamos PISA 2000 y PISA 2009 en lectura y cuando comparamos PISA 2003 y PISA 2009 en matemáticas.

Si comparamos el grupo de tratamiento y el grupo de control antes y después del cambio en lectura, podemos ver tendencias que son muy similares entre ambos grupos: la proporción de niñas y estudiantes provenientes de familias monoparentales disminuye ligeramente mientras que la edad, los inmigrantes, el tiempo de aprendizaje, y el índice de estatus económico, social y cultural (ESCS) aumenta. Esto es consistente con el hecho de que España ha experimentado una gran afluencia de inmigrantes en la última década y ha convergido hacia el PIB per cápita de la UE y la OCDE, un proceso que desde 2009 se ha invertido<sup>7</sup>. Sin embargo, el aumento en la proporción de inmigrantes entre 2000 y 2009 fue mayor en la Comunidad de Madrid (desde el 3% al 16%) que en el resto de las Comunidades (desde el 2% al 9%). Además, la proporción de estudiantes hablando idiomas extranjeros diferentes al español aumentó en Madrid durante el período (desde el 1% al 5%) mientras que permaneció constante en el grupo de control (17% frente al 16%).

Si nos fijamos en las características de las escuelas, observamos una disminución en el número de escuelas privadas y un aumento en el número de escuelas concertadas en el período 2000-2009. Esto podría ser debido al hecho de que algunos colegios privados han solicitado y logrado de la Administración Pública su transformación en colegios concertados, lo que disminuye las tasas que pagan las familias de los alumnos y evita que desciendan las inscripciones en esos colegios. Sin embargo, los datos oficiales de la Oficina de Estadística del Ministerio Español de Educación, Cultura y Deporte muestra que el aumento de los alumnos de las escuelas concertadas procede de una reducción en el número de estudiantes en los colegios públicos. Esto contrasta con la muestra de PISA, la cual muestra que el aumento de los estudiantes en colegios concertados procede de un descenso en los colegios privados. Esto es, la cobertura de PISA en las escuelas privadas disminuyó desde 2000 a 2009 mientras que la cobertura de escuelas públicas aumentó. Esto podría ser debido al hecho de que la muestra de las escuelas

<sup>7</sup> España pasó de menos del 1% de inmigrantes en la población a casi el 10% durante este período. El PIB per cápita español en términos PPA aumentó desde el 97% del correspondiente al de la UE-27 en 2000 al 103% en 2009 (Source: Eurostat).

en las ciudades o grandes urbes en 2009 disminuyó mientras que en ciudades pequeñas y pueblos aumentó.

Estudiante/profesor y el ratio del tamaño de la escuela disminuyó a lo largo de los dos períodos. También, observamos que el porcentaje de colegios que declararon que llevaron a cabo evaluaciones utilizadas para comparar el colegio con el rendimiento del distrito/nacional o evaluaciones utilizadas para hacer juicios sobre la efectividad del profesor aumentó durante los dos períodos tanto en el grupo de control como en el de tratamiento. En resumen, las estadísticas descriptivas muestran que las tendencias que observamos cuando comparamos el grupo de tratamiento y de control son similares.

El grupo de control y de tratamiento tienen también patrones similares en PISA 2003 y en PISA 2009, que son los años que estamos utilizando para el análisis de matemáticas. Las únicas excepciones son la proporción de niñas y, sobre todo, el índice ESCS. En ambos casos, el indicador de la Comunidad de Madrid aumentó mientras que el del grupo de control disminuyó ligeramente.

## 4. METODOLOGÍA ECONOMETRICA Y RESULTADOS

Con el fin de estimar el impacto de la introducción de un examen estandarizado en la Comunidad de Madrid sobre el rendimiento de los estudiantes, proponemos un enfoque de diferencias en diferencias. Utilizamos como resultado del desempeño de los estudiantes, sus notas presentadas por PISA. Éstas se calculan usando métodos de imputación, designados los valores plausibles (OCDE, 2009). Así, para un año dado, la nota del estudiante  $i$  en el colegio  $j$  viene dada por:

$$y_{ij} = \alpha_0 + a_1 \text{Madrid}_j + a_2 \text{PISA2009} + \delta \text{Madrid}_j \cdot \text{PISA2009} + \beta x_i + \beta x_j + u_i + \varepsilon_{ij}$$

donde  $x_i$  son características observables de los estudiantes y sus familias descritas anteriormente,  $x_j$  son características observables de los colegios,  $\text{Madrid}_j$  es una variable dummy para las escuelas ubicadas en la Comunidad de Madrid (es decir, toma el valor 1 para el grupo tratado),  $\text{PISA2009}$  es una variable dummy para los estudiantes que realizaron el examen de PISA en 2009 (después de la introducción del examen estandarizado en la Comunidad de Madrid),  $\text{Madrid}_j \cdot \text{PISA2009}$  indica si la escuela  $j$  está en la Comunidad de Madrid y participa en el examen de PISA de 2009 (es decir, toma el valor 1 para el grupo tratado después del tratamiento),  $u_i$  son características no observables de los estudiantes, tales como esfuerzo o capacidad,  $\varepsilon_{ij}$  son características no observables del colegio, como calidad del director y de los maestros, y  $\varepsilon_{ij}$  es un shock aleatorio.

Nuestro parámetro de interés es  $\delta$ , correspondiente a la variable  $\text{Madrid}_j \cdot \text{PISA2009}$ , que coincide con la introducción de un examen estandarizado (el examen CDI) en la Comunidad de Madrid.

Las tablas 3 y 4 de abajo muestran los resultados de la estimación de diferencias en diferencias para la lectura, usando las muestras de los estudian-

tes de España en PISA 2000 y PISA 2009, y para matemáticas, usando las muestras de estudiantes de España en PISA 2003 y PISA 2009. Dado que la base de datos de PISA proporciona cinco valores plausibles, que se asignan a cada estudiante, usamos la metodología propuesta por la OCDE para el cálculo de los coeficientes de regresión y sus respectivos errores estándar. Según la OCDE (2009), los análisis estadísticos deberían comportarse independientemente en cada uno de estos cinco valores plausibles y los resultados deberían ser agregados para obtener las estimaciones finales de las estadísticas y sus errores estándar respectivos.

**Tabla 3: Estimaciones de diferencias en diferencias para PISA 2000 y PISA 2009 en lectura**

Variable	(1)	(2)	(3)	(4)
Madrid	13.620* (8.107)	9.327 (5.841)	2.071 (8.001)	2.761 (6.869)
<b>Madrid*PISA2009</b>	<b>12.267</b> <b>(9.703)</b>	<b>13.718**</b> <b>(6.867)</b>	<b>16.985**</b> <b>(8.083)</b>	<b>15.415**</b> <b>(7.657)</b>
PISA2009	-13.185*** (3.635)	-9.499*** (2.995)	-5.448 (3.843)	-8.436** (3.744)
<b>Características de los estudiantes</b>				
Hembra		27.230*** (1.789)	26.115*** (1.781)	26.231*** (1.696)
Edad		18.521*** (3.118)	19.680*** (3.372)	19.180*** (3.366)
Inmigrante		-17.608*** (6.301)	-16.988** (6.778)	-19.294*** (6.504)
Madre inmigrante		-13.890** (6.766)	-14.851** (7.345)	-11.932* (7.027)
Padre inmigrante		-8.080 (5.609)	-7.079 (5.983)	-5.043 (6.043)
Idioma extranjero hablado en el hogar		-4.108 (2.689)	-2.888 (2.615)	-0.192 (2.721)
Familia monoparental		-4.899** (2.288)	-5.730** (2.312)	-6.146*** (2.178)
Familia mixta		-22.204*** (5.404)	-24.441*** (5.750)	-25.051*** (5.911)
Tiempo de estudio (horas/semana) en lectura		-8.492*** (1.420)	-8.690*** (1.545)	-12.495*** (1.587)
ESCS (Índice del nivel económico, social y cultural)		27.955*** (1.100)	24.845*** (1.072)	19.628*** (0.852)
ESCS al cuadrado		-1.564** (0.757)	-1.868** (0.763)	-1.656** (0.710)
<b>Características del colegio</b>				
Colegio privado			23.473*** (8.548)	-3.719 (8.095)
Colegio concertado			12.103** (5.776)	-1.032 (5.099)
Colegio en pequeña ciudad			1.078 (6.443)	1.581 (4.890)
Colegio en ciudad			1.749 (6.653)	0.086 (5.042)
Colegio en urbe			7.013 (7.233)	4.163 (5.556)
Colegio en una gran urbe			12.746 (8.965)	8.784 (6.671)

**Tabla 3: Estimaciones de diferencias en diferencias para PISA 2000 y PISA 2009 en lectura**

Variable	(1)	(2)	(3)	(4)
<b>Características del colegio</b>				
Ratio estudiante/profesor			0.038 (0.657)	0.373 (0.544)
Tamaño del colegio			0.005 (0.004)	-0.006* (0.004)
Evaluación para comparar el colegio con el resultado en el distrito/país			-1.527 (2.993)	-2.520 (2.959)
Evaluación para juzgar la efectividad del profesor			0.592 (2.669)	-0.442 (2.389)
Proporción de niñas en el colegio				-14.228 (13.844)
Media del colegio del índice ESCS				23.880*** (2.808)
Media del colegio del índice ESCS al cuadrado				-8.226*** (2.706)
% de estudiantes inmigrantes en el colegio				0.471 (0.451)
% de madres inmigrantes en el colegio				-0.109 (0.336)
% de padres inmigrantes en el colegio				-0.538 (0.415)
Promedio de la escuela del tiempo de dedicación (horas/semana) a la lectura				11.089*** (3.485)
Constante	490.770*** (2.793)	229.792*** (49.645)	197.296*** (54.414)	221.023*** (55.028)

Fuente: 1. Errores estándar en paréntesis. \* significativo al 10%; \*\* significativo al 5%; \*\*\* significativo al 1% 2. Categorías base para dummies: varón, estudiante de España, madre de España, padre de España, familia nuclear, colegio público, colegio en pueblo.

**Tabla 4: Estimaciones de diferencias en diferencias para PISA 2003 y PISA 2009 en matemáticas**

Variable	(1)	(2)	(3)	(4)
Madrid	9.580 (9.973)	7.819 (6.039)	-3.420 (7.350)	-6.879 (6.791)
<b>Madrid*PISA2009</b>	<b>5.415</b> <b>(10.559)</b>	<b>-0.084</b> <b>(6.606)</b>	<b>4.272</b> <b>(7.268)</b>	<b>7.323</b> <b>(7.509)</b>
PISA2009	-2.296 (3.360)	6.752** (2.688)	6.619** (3.227)	8.484** (3.514)
<b>Características de los estudiantes</b>				
Hembra		-13.744*** (1.626)	-14.519*** (1.605)	-15.746*** (1.584)
Edad		15.847*** (3.269)	14.596*** (3.433)	15.139*** (3.325)
Inmigrante		-21.247*** (6.481)	-21.009*** (6.973)	-20.172*** (6.948)
Madre inmigrante		-9.528** (4.416)	-8.680* (4.507)	-7.674* (4.433)
Padre inmigrante		-13.285*** (5.073)	-15.604*** (5.339)	-13.873*** (5.268)
Idioma extranjero hablado en el hogar		2.083 (2.619)	2.103 (2.852)	-0.132 (2.997)
Familia monoparental		-8.722*** (2.771)	-9.873*** (2.723)	-10.985*** (2.699)
Familia mixta		-5.591 (5.366)	-10.635 (6.748)	-11.649* (6.902)

**Tabla 4: Estimaciones de diferencias en diferencias para PISA 2003 y PISA 2009 en matemáticas**

Variable	(1)	(2)	(3)	(4)
<b>Características de los estudiantes</b>				
Tiempo de estudio (horas/semana) en lectura		-6.616*** (1.560)	-6.759*** (1.620)	-4.840*** (1.844)
ESCS (Índice del nivel económico, social y cultural)		29.471*** (1.024)	26.666*** (1.043)	21.539*** (0.853)
ESCS al cuadrado		-0.177 (0.735)	-0.644 (0.799)	-0.046 (0.657)
<b>Características del colegio</b>				
Colegio privado			14.301** (6.453)	-0.962 (6.895)
Colegio concertado			10.068** (4.193)	2.393 (4.053)
Colegio en pequeña ciudad			3.890 (4.283)	3.007 (4.901)
Colegio en ciudad			-1.624 (4.961)	-2.684 (5.075)
Colegio en urbe			4.235 (5.395)	1.968 (5.389)
Colegio en una gran urbe			17.099** (8.373)	15.854* (8.311)
Ratio estudiante/profesor			-0.024 (0.441)	-0.161 (0.446)
Tamaño del colegio			0.009* (0.005)	-0.000 (0.004)
Evaluación para comparar el colegio con el resultado en el distrito/país			-0.661 (3.322)	-3.048 (3.235)
Evaluación para juzgar la efectividad del profesor			-1.618 (3.460)	-1.442 (3.471)
Proporción de niñas en el colegio				49.186*** (15.062)
Media del colegio del índice ESCS				22.307*** (3.094)
Media del colegio del índice ESCS al cuadrado				-10.684*** (3.716)
% de estudiantes inmigrantes en el colegio				0.032 (0.333)
% de madres inmigrantes en el colegio				0.106 (0.342)
% de padres inmigrantes en el colegio				-0.420 (0.386)
Promedio de la escuela del tiempo de dedicación (horas/semana) a la lectura				-1.111 (4.785)
Constante	483.786*** (2.370)	273.629*** (52.352)	285.233*** (55.280)	281.827*** (54.504)

Fuente: 1. Errores estándar en paréntesis. \* significativo al 10%; \*\* significativo al 5%; \*\*\* significativo al 1% 2. Categorías base para dummies: varón, estudiante de España, madre de España, padre de España, familia nuclear, colegio público, colegio en pueblo.

La primera columna de las tablas muestra los resultados de la estimación sin variables de control. Este sería el efecto bruto promedio de nuestro tratamiento. La segunda columna incluye características individuales de los estudiantes y la tercera y la cuarta columna agregan gradualmente características de los colegios.

Cuando estimamos diferencias en diferencias sin variables de control, el coeficiente para el tratamiento no es estadísticamente significativo para la lectura ni para matemáticas.

Sin embargo, los resultados de la estimación de diferencias en diferencias para la lectura en la tabla 3, columnas (2)-(4), muestran un efecto positivo y estadístico significativo de nuestro tratamiento en las notas de PISA. En la segunda columna, cuando controlamos por las características individuales de los estudiantes, el coeficiente de la variable de tratamiento es positivo y significativo<sup>8</sup>. La inclusión de las características de la escuela en las columnas (3) y (4) no cambia este resultado. Encontramos una mejora relativa en las notas de PISA en lectura en la Comunidad de Madrid entre 2000 y 2009 de una magnitud de 14 a 17 puntos de PISA que no puede explicarse por variables observables. En 2009, España estaba significativamente por debajo de la media de la OCDE en lectura en 12 puntos. Si los resultados son totalmente explicados por la introducción y la publicación de exámenes externos, esto podría implicar que la generalización de esos exámenes aumentaría el nivel de España en lectura sobre la media de la OCDE.

En la tabla 4 se llevan a cabo las mismas estimaciones pero para matemáticas y utilizando los resultados de PISA 2003 y PISA 2009. En este caso, no encontramos ningún impacto de nuestro tratamiento en el rendimiento de los estudiantes. El coeficiente cambia de positivo a negativo de una especificación a otra y no es estadísticamente significativo en ninguna de ellas. La falta de significatividad podría explicarse en parte por el hecho de que estamos utilizando como rendimiento del estudiante las puntuaciones de PISA en una materia, la cual fue objeto del examen de PISA sólo en 2003 y no en 2009, por lo que las comparaciones no son del todo válidas.

Los datos parecen indicar que algo diferente ha ocurrido en Madrid entre 2000 y 2009 con respecto a otras regiones españolas. Una hipótesis lógica en este contexto es que la introducción y la publicación de los resultados del examen estandarizado toma un papel importante en este cambio. Es muy difícil ofrecer una prueba definitiva con estos datos, pero podemos descartar algunas explicaciones alternativas.

<sup>8</sup> En una especificación adicional, quitamos el índice ESCS, que es un índice agregado del perfil socioeconómico de los estudiantes y controlamos por separado la situación del mercado laboral y el nivel de educación de la madre y del padre. Nuestros resultados no cambiaron significativamente. Estos resultados están disponibles bajo petición.

El gasto público en educación por alumno afecta hasta cierto punto a los rendimientos de los estudiantes (OCDE, 2010). El Ministerio español de Educación proporciona datos sobre gasto público en educación por alumno y por regiones a partir de 2004. Durante el período 2004-2009, Madrid aumentó el gasto público en educación por alumno en un 21%, menos que la media española (33%). Más importante aún, Madrid ha sido la región que menos ha aumentado el gasto en educación por alumno de entre las 17 regiones españolas. Por lo tanto, el gasto en educación no puede explicar el mejor comportamiento de las puntuaciones de PISA en Madrid.

España recibió una gran cantidad de inmigrantes entre 2000 y 2009 (los inmigrantes pasaron del 2% al 12% de la población durante el período, según el Instituto Nacional de Estadística, INE), y Madrid fue un importante lugar de destino (tiene aproximadamente el 18% de los inmigrantes y sobre el 13% de la población). Nuestros datos pueden identificar si el estudiante es inmigrante y si el número de inmigrantes varía bastante entre colegios cuyo efecto es probablemente captado a nivel escolar. Este es también un período de rápido crecimiento económico, el cual no es igual entre regiones, pero el índice ESCS tiene suficiente información sobre esta variable a nivel individual para controlar adecuadamente el efecto de los datos económicos. Otros factores afectan a los colegios más directamente. Madrid tiene un mayor número de colegios concertados que otras regiones. Madrid también ha aumentado la proporción de colegios concertados pero su tendencia ha sido similar a la del resto de Comunidades, si acaso un poco más pequeña. En cualquier caso, ya que la identidad de las escuelas es observable su efecto puede ser controlado.

La otra reforma institucional importante en los colegios de Madrid en este período, más allá de la introducción y publicación de exámenes externos, es la introducción de colegios bilingües en la región, donde inglés es un medio de enseñanza durante al menos una tercera parte del tiempo<sup>9</sup>. Aunque esto es claramente una importante reforma, se ha llevado a cabo sólo gradualmente a partir del primer grado, y los alumnos mayores expuestos al programa ahora tienen 13 años. Además, Anghel, Cabrales and Carro (2012) han encontrado efectos significativos del programa tanto en lengua como en matemáticas, y posiblemente un efecto negativo en las ciencias sociales y naturales (las materias enseñadas en inglés).

<sup>8</sup> Estudiantes no sólo estudian inglés como lengua extranjera sino que también algunas asignaturas (al menos ciencias, historia y geografía) son impartidas en inglés. Lengua española y matemáticas se imparten sólo en español.

## 5. MÉTODO DE CONTROL SINTÉTICO

En esta sección, utilizamos la metodología propuesta por Abadie y Gardeazabal (2003) y Abadie et al. (2010), que aplica métodos de control sintéticos para los estudios de casos comparativos. Su metodología está motivada por el hecho de que, en estudios de casos comparativos, el investigador se ve obligado a encontrar similitudes entre unidades tratadas y no tratadas utilizando características observables, algo que a menudo es difícil en la práctica. Para resolver este problema proponen construir una combinación de unidades con fines comparativos, ya que la combinación normalmente se parecerá a la unidad tratada mucho mejor que cualquier unidad sola.

En nuestro caso construimos una combinación de regiones españolas que se asemeja a la región de Madrid en términos de varias características antes del tratamiento y observamos la evolución de esta combinación en ausencia del tratamiento. Esta combinación se denomina grupo de control sintético. Se construye buscando una combinación ponderada de las regiones españolas no tratadas, en términos de diferentes variables de predicción, las cuales son la media de todo el período antes de la intervención. Según Abadie et al. (2010) “debido a que la elección del control sintético no requiere acceso a los resultados posteriores a la intervención, el método de control sintético permite a los investigadores decidir sobre el diseño del estudio sin conocer cómo esas decisiones afectarán a las conclusiones de sus estudios”.

Para construir el grupo de control sintético (el Madrid sintético), tenemos que agregar los datos a nivel del colegio y después a nivel regional. El año en el que el examen estandarizado CDI fue llevado a cabo en la Comunidad de Madrid fue 2004/05, por lo tanto tenemos dos años de datos de PISA anteriores al tratamiento (PISA 2000 y 2003). PISA 2006 y PISA 2009 será

nuestro período posterior al tratamiento. El Madrid sintético se construye como una media ponderada del grupo de regiones no tratadas. Este grupo incluye 15 regiones<sup>10</sup>. Las ponderaciones se eligen de manera que el sintético de Madrid resultante se parezca a la verdadera región de Madrid lo máximo posible en términos de los valores de un conjunto de predicciones del rendimiento de los estudiantes antes de la introducción del examen CDI, esto es, antes del tratamiento.

Incluimos en la lista de las variables de predicción para calcular los pesos las siguientes variables: el ratio profesor/estudiante, el tamaño del colegio, el índice del colegio ESCS, la proporción de inmigrantes en el colegio y la proporción de repetidores en el colegio. Todas las variables son medias a nivel regional y durante el período antes de la intervención (2000 y 2003).

Utilizando estas variables de predicción construimos el Madrid sintético como la combinación convexa de las regiones, que se parezca más a la región de Madrid en el período antes del tratamiento. La Tabla 5 muestra las características de la región de Madrid real, de la región sintética de Madrid y del grupo de las regiones (la media de las 15 regiones las cuales forman el grupo), en términos de las variables de control. Las cifras demuestran que el Madrid sintético construido es mucho más parecido al real, tanto en lectura como en matemáticas, que el promedio simple de las regiones que forman el grupo. En lectura, el ratio estudiante/profesor en la región real de Madrid es 16,18 y en el Madrid sintético es 15 (la media del grupo de control es 14,03). La media del tamaño del colegio es 933,74 en el Madrid real y 852,74 en el Madrid sintético (el promedio del grupo es 720,08). La media del colegio del índice ESCS es -0,16 en el Madrid real y -0,19 en el Madrid sintético (el promedio del grupo de control es -0,31). En términos del porcentaje de repetidores encontramos que la región sintética y el grupo de control son ambos bastante similares a la región de Madrid real. Finalmente, hay una diferencia sustancial en el porcentaje de estudiantes inmigrantes entre la región real de Madrid y la sintética. Por otra parte, encontramos también afinidades entre el Madrid sintético y el real en los resultados de PISA de los estudiantes.

<sup>10</sup> Hay 17 regiones (incluyendo la región de Madrid) y dos ciudades autónomas (Ceuta y Melilla) en España. Tuvimos que quitar Baleares y Ceuta y Melilla por falta de datos, por lo que esto nos deja con 15 Comunidades o regiones.

**Tabla 5: Medias de predicción para el rendimiento de los estudiantes en PISA**

Variable	LECTURA		METEMÁTICAS		LECTURA	MATEMÁTICAS
	Madrid		Madrid		Media de 15 regiones	
	Real	Sintético	Real	Sintético		
Ratio estudiante/profesor	16,18	15,00	16,18	14,85	14,03	14,03
Tamaño del colegio	933,74	852,74	933,74	907,70	720,08	720,08
Media del colegio del índice ESCS	-0,16	-0,19	-0,16	-0,16	-0,31	-0,31
% de estudiantes inmigrantes en el colegio	5,43	3,35	5,43	2,52	2,77	2,77
% de repetidores en el colegio	0,36	0,31	0,36	0,29	0,27	0,27
PISA calificación 2000	506,80	500,46	490,83	493,04	496,92	482,30
PISA calificación 2003	490,66	480,00	492,11	485,63	488,63	493,23
PISA calificación 2006	477,27	474,42	500,11	502,07	471,05	492,25
PISA calificación 2009	500,86	483,61	494,63	495,33	484,70	490,64

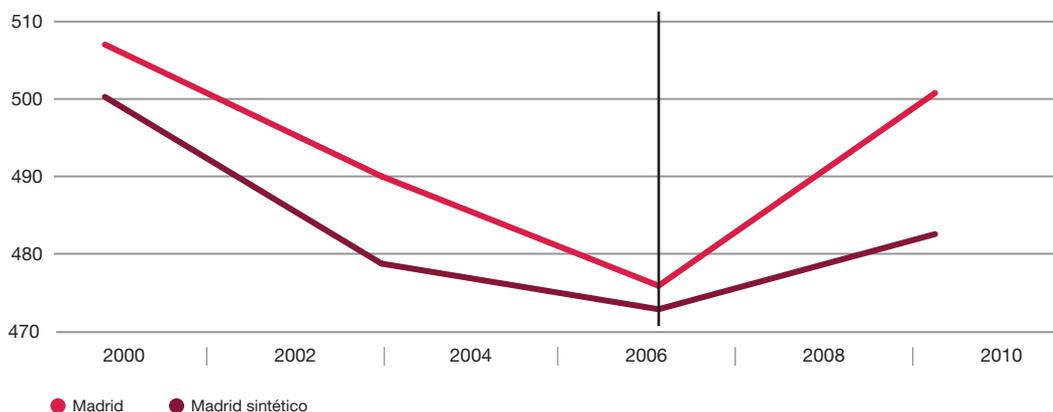
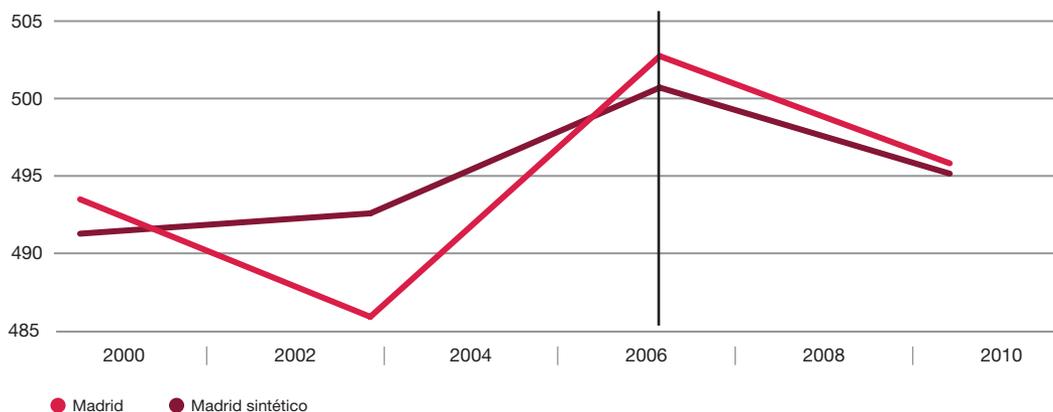
Fuente: Nota: Todas las variables, excepto las calificaciones de PISA, son la media a nivel regional para el período anterior al tratamiento (2000 y 2003).

La Tabla 6 proporciona los pesos de las 15 regiones que forman el grupo en el Madrid sintético. Muestra que el rendimiento de los estudiantes en la región de Madrid está mejor aproximada por una combinación de Aragón, Asturias y las Islas Canarias en lectura y de Aragón y las Islas Canarias en matemáticas. El resto de regiones en el grupo se les asigna ponderaciones cero.

**Tabla 6: Pesos por región en el Madrid Sintético (PISA2000-PISA2009)**

Región	Pesos lectura	Pesos matemáticas
Andalucía	0	0
Aragón	0,454	0,801
Asturias	0,286	0
Canarias	0,259	0,199
Cantabria	0	0
Castilla - La Mancha	0	0
Castilla-León	0	0
Cataluña	0	0
Extremadura	0	0
Galicia	0	0
La Rioja	0	0
Murcia	0	0
Navarra	0	0
País Vasco	0	0
Valencia	0	0

Los próximos dos gráficos muestran la evolución de la región de Madrid real y de la sintética en 2000 y 2003 (los años antes de la intervención) y en 2006 y 2009 (los años posteriores a la intervención), por separado la lectura y las matemáticas.

**Gráfico 1: La evolución del Madrid real y del Madrid sintético en lectura en 2000, 2003, 2006 y 2009****Gráfico 2: La evolución del Madrid real y del Madrid sintético en matemáticas en 2000, 2003, 2006 y 2009**

Para la lectura, el gráfico muestra que el Madrid sintético aproxima muy bien la evolución del Madrid real en el período antes del tratamiento. Después del tratamiento, el cual nos lleva a la introducción del examen estandarizado CDI en la región de Madrid, en 2006 y 2009, la evolución del grupo de control sintético es diferente que la del grupo de tratamiento. En particular vemos que, aunque tanto la región de Madrid real como la sintética han tenido una tendencia creciente después de 2006, el Madrid sintético está comportándose peor que la región de Madrid real: en 2009, la diferencia del rendimiento es 17,24 puntos PISA a favor de la región de Madrid real. Esta diferencia podría atribuirse a la introducción de un examen estandarizado en la región de Madrid, al menos con la información que podemos observar.

Esto confirma, incluso cuantitativamente, los resultados que obtuvimos previamente con la metodología de diferencias en diferencias, donde encontramos que controlando las características de la escuela, la región de Madrid mejoró su rendimiento relativo a otras regiones de España en el período de 2000 a 2009, entre 14 y 17 puntos PISA. La principal publicación sobre educación de la OCDE, *Panorama de la Educación*, llega a una conclusión similar en la última edición de 2012: “Los estudiantes en los sistemas escolares que utilizan pruebas externas basadas en estándares obtuvieron 16 puntos más de media entre los países de la OCDE, que los estudiantes en sistemas escolares donde no se usan esas pruebas (*Education at a Glance*, 2012, página 527). Nuestra estimación es un poco más pequeña que el rango encontrado en la literatura por Hanushek y Woessmann (2011) del 20% al 40% de la desviación estándar (20 a 40 puntos de PISA).

Para matemáticas, sin embargo, la metodología del grupo de control sintético no funciona bien. El Madrid sintético no se aproxima muy bien a la evolución del Madrid real en 2000 y 2003, el período anterior al tratamiento. En el período después del tratamiento, el Madrid sintético está funcionando algo mejor que el Madrid real. No obstante, la estimación de diferencias en diferencias no mostró ningún impacto estadístico en la materia de matemáticas.

Como prueba de robustez, introducimos otras variables entre las variables de predicción que se han especificado anteriormente. En particular, incluimos el porcentaje medio de niñas en el examen PISA y la proporción media de escuelas públicas, concertadas y privadas. Nuestros resultados no cambian significativamente.

El resultado diferente para lectura y para matemáticas podría explicarse por el hecho de que la lectura fue el principal objetivo de la prueba de PISA en 2000 y 2009, mientras que las matemáticas fueron el principal objetivo en 2003 pero no en 2009. Ser el objetivo del estudio significa, como se indicó en el apartado de Estructura Institucional, que dos terceras partes del estudio se dedican a esta materia, mientras que la otra tercera parte se divide por igual entre las otras dos áreas. Este análisis más detallado en una de las materias permite desagregar los rendimientos de los estudiantes por sub-áreas de la materia elegida. En 2009, PISA calculó en detalle los niveles de competencia de los estudiantes en varios aspectos de la lectura,

tales como la habilidad de los estudiantes para acceder y conseguir datos, la incorporación y la interpretación de los mismos, y la reflexión y evaluación de la información que ellos obtienen por medio de la lectura. También examinó la capacidad de los estudiantes para leer y entender textos continuos y discontinuos. Por el contrario, PISA 2009 sólo calculó un rendimiento agregado para matemáticas y ciencias. Por lo tanto, el rendimiento en lectura es una estimación más precisa que el de matemáticas en 2009 mientras que en 2012 será al revés.

Somos conscientes de las limitaciones de nuestros datos en la realización de la estimación usando métodos de control sintéticos. Una de las principales limitaciones es que debido a que el estudio de PISA empezó en 2000 y es llevado a cabo cada tres años, tenemos solo dos años de datos antes de la intervención (2000 y 2003), lo que complica el cálculo de los pesos de las regiones para el grupo de control sintético. El resultado en matemáticas, donde el Madrid sintético no es tan similar al Madrid real en el año anterior al tratamiento puede estar en parte explicado por este hecho.

## 6. CONCLUSIONES

Este trabajo trata de identificar si la implementación y la publicación de los resultados de pruebas externas y estandarizadas podrían tener algún impacto en el rendimiento de los estudiantes. Utilizamos el hecho de que en la Comunidad de Madrid se realizó un examen estandarizado por primera vez (y sus resultados se publicaron) en 2004/2005 para todos los alumnos de 6º de primaria, mientras que en otras regiones de España, no existía dicho examen. Empleando un enfoque de diferencias en diferencias encontramos un efecto positivo en la lectura del orden de 14 a 17 puntos PISA. El método de control sintético conduce a un efecto parecido incluso en términos cuantitativos. Nuestros resultados están en línea con las investigaciones previas en el área, pero nuestro estudio proporciona una innovación importante, ya que los exámenes externos en Madrid no tienen consecuencias para los estudiantes, por lo que el efecto tiene que provenir del impacto en maestros y directores de las escuelas.

Hemos identificado un posible efecto en lengua, pero no en matemáticas. Esto es sorprendente, ya que muchos programas educativos han observado efectos que son mayores en matemáticas que en lengua (véase por ejemplo Abdulkadiroglu et al. 2011). Una posible explicación puede venir del diferente énfasis que tiene el plan de estudios en los niveles educativos del colegio primario en España con respecto a otros países, pero esta idea merece una investigación más a fondo, la cual aplazamos a futuros estudios.

## 7. BIBLIOGRAFÍA

- Abadie, Alberto y Javier Gardeazabal (2003), “The Economic Costs of Conflict: A Case Study of the Basque Country”, *The American Economic Review*, 93 (1), 113-132
- Abadie, Alberto, Alexis Diamond y Jens Hainmueller (2010), “Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program”, *Journal of the American Statistical Association*, 105 (490), 493-505
- Abdulkadiroglu, Attila, Joshua Angrist, Susan M. Dynarski, Thomas J. Kane, y Parag A. Pathak (2011), “Accountability and Flexibility in Public Schools: Evidence from Boston’s Charters and Pilots,” *The Quarterly Journal of Economics*, vol. 126, pág. 699–748.
- Anghel, Brindusa y Antonio Cabrales (2010), “The Determinants of Success in Primary Education in Spain,” *Documento de trabajo FEDEA*, N°. 2010-20.
- Anghel, Brindusa, Antonio Cabrales y Jesús M. Carro (2012), “Evaluating a bilingual education program in Spain: the impact beyond foreign language learning”, *CEPR Documento de trabajo*, N°. 8995.
- Bishop, John H. (1997), “The Effect of National Standards and Curriculum-Based Examinations on Achievement”, *American Economic Review*, vol. 87 (2), pág. 260-264.
- Bishop John H. (2006), “Drinking from the fountain of knowledge: student incentive to study and learn,” In: Hanushek EA, Welch F (eds) *Handbook of the economics of education*. North-Holland, Amsterdam
- Bishop, John H., Ferran Mane, Michael Bishop y Joan Moriarty (2001), “The Role of End-of-Course Exams and Minimum Competency Exams in Standards-Based Reforms,” *Brookings Papers on Education Policy* vol. 4, pág. 267-345.

- Hanushek, Eric y Ludger Woessmann (2011): “The Economics of International Differences in Educational Achievement”, in Hanushek, E, Machin, S. and Woessmann, L. (Eds.) Handbook of the Economics of Education, vol. 3, pág. 89-200, Amsterdam: North Holland
- Jürges, Hendrik, Kerstin Schneider y Felix Büchel (2005), “The Effect Of Central Exit Examinations On Student Achievement: Quasi-Experimental Evidence From TIMSS Germany”, Journal of the European Economic Association, vol. 3(5), pág. 1134-1155, 09.
- OECD (2009), “PISA Data Analysis Manual: SPSS Second Edition”.  
OECD (2010), “PISA 2009 Results: What Makes a School Successful?” Volume IV.
- Wößmann, Ludger (2010), “Institutional Determinants of School Efficiency and Equity: German States as a Microcosm for OECD Countries”, Jahrbücher für Nationalökonomie und Statistik vol. 230 (2), pág. 234–270,



## ÚLTIMOS DOCUMENTOS DE TRABAJO

- 2013-07: Publicación de los resultados de las pruebas estandarizadas externas: ¿Tiene ello un efecto sobre los resultados escolares?, **Brindusa Anghel, Antonio Cabrales, Jorge Sainz e Ismael Sanz.**
- 2013-06: "DYPES: A Microsimulation model for the Spanish retirement pension system", **F. J. Fernández-Díaz, C. Patxot y G. Souto.**
- 2013-05: "Vertical differentiation, schedule delay and entry deterrence: Low cost vs. full service airlines", **Jorge Validoa, M. Pilar Socorroa y Francesca Medda.**
- 2013-04: "Dropout Trends and Educational Reforms: The Role of the LOGSE in Spain", **Florentino Felgueroso, María Gutiérrez-Domènech y Sergi Jiménez-Martín.**
- 2013-03: "Understanding Different Migrant Selection Patterns in Rural and Urban Mexico", **Simone Bertoli, Herbert Brücker y Jesús Fernández-Huertas Moraga.**
- 2013-02: "Understanding Different Migrant Selection Patterns in Rural and Urban Mexico", **Jesús Fernández-Huertas Moraga.**
- 2013-01: "Publicizing the results of standardized external tests: Does it have an effect on school outcomes?", **Brindusa Anghel, Antonio Cabrales, Jorge Sainz y Ismael Sanz.**
- 2012-12: "Visa Policies, Networks and the Cliff at the Border", **Simone Bertoli, Jesús Fernández-Huertas Moraga.**
- 2012-11: "Intergenerational and Socioeconomic Gradients of Child Obesity", **Joan Costa-Fonta y Joan Gil.**
- 2012-10: "Subsidies for resident passengers in air transport markets", **Jorge Valido, M. Pilar Socorro, Aday Hernández y Ofelia Betancor.**
- 2012-09: "Dual Labour Markets and the Tenure Distribution: Reducing Severance Pay or Introducing a Single Contract?", **J. Ignacio García Pérez y Victoria Osuna.**
- 2012-08: "The Influence of BMI, Obesity and Overweight on Medical Costs: A Panel Data Approach", **Toni Mora, Joan Gil y Antoni Sicras-Mainer.**
- 2012-07: "Strategic behavior in regressions: an experimental", **Javier Perote, Juan Perote-Peña y Marc Vorsatz.**
- 2012-06: "Access pricing, infrastructure investment and intermodal competition", **Ginés de Rus y M. Pilar Socorro.**
- 2012-05: "Trade-offs between environmental regulation and market competition: airlines, emission trading systems and entry deterrence", **Cristina Barbot, Ofelia Betancor, M. Pilar Socorro y M. Fernanda Viacens.**
- 2012-04: "Labor Income and the Design of Default Portfolios in Mandatory Pension Systems: An Application to Chile", **A. Sánchez Martín, S. Jiménez Martín, D. Robalino y F. Todeschini.**
- 2012-03: "Spain 2011 Pension Reform", **J. Ignacio Conde-Ruiz y Clara I. Gonzalez.**
- 2012-02: "Study Time and Scholarly Achievement in PISA", **Zöe Kuehn y Pedro Landeras.**
- 2012-01: "Reforming an Insider-Outsider Labor Market: The Spanish Experience", **Samuel Bentolila, Juan J. Dolado y Juan F. Jimeno.**
- 2011-13: "Infrastructure investment and incentives with supranational funding", **Ginés de Rus y M. Pilar Socorro.**
- 2011-12: "The BCA of HSR. Should the Government Invest in High Speed Rail Infrastructure?", **Ginés de Rus.**
- 2011-11: "La rentabilidad privada y fiscal de la educación en España y sus regiones", **Angel de la Fuente y Juan Francisco Jimeno.**
- 2011-10: "Tradable Immigration Quotas", **Jesús Fernández-Huertas Moraga y Hillel Rapoport.**
- 2011-09: "The Effects of Employment Uncertainty and Wealth Shocks on the Labor Supply and Claiming Behavior of Older American Workers", **Hugo Benítez-Silva, J. Ignacio García-Pérez y Sergi Jiménez-Martín.**
- 2011-08: "The Effect of Public Sector Employment on Women's Labour Market Outcomes", **Brindusa Anghel, Sara de la Rica y Juan J. Dolado.**
- 2011-07: "The peer group effect and the optimality properties of head and income taxes", **Francisco Martínez-Mora.**
- 2011-06: "Public Preferences for Climate Change Policies: Evidence from Spain", **Michael Hanemann, Xavier Labandeira y María L. Loureiro.**
- 2011-05: "A Matter of Weight? Hours of Work of Married Men and Women and Their Relative Physical Attractiveness", **Sonia Oreffice y Climent Quintana-Domeque.**
- 2011-04: "Multilateral Resistance to Migration", **Simone Bertoli y Jesús Fernández-Huertas Moraga.**
- 2011-03: "On the Utility Representation of Asymmetric Single-Peaked Preferences", **Francisco Martínez Mora y M. Socorro Puy.**
- 2011-02: "Strategic Behaviour of Exporting and Importing Countries of a Non-Renewable Natural Resource: Taxation and Capturing Rents", **Emilio Cerdá y Xiral López-Otero.**
- 2011-01: "Politicians' Luck of the Draw: Evidence from the Spanish Christmas Lottery", **Manuel F. Bagues y Berta Esteve-Volart.**
- 2010-31: "The Effect of Family Background on Student Effort", **Pedro Landeras.**
- 2010-29: "Random-Walk-Based Segregation Measures", **Coralio Ballester y Marc Vorsatz.**