UNIVERSIDAD DE EXTREMADURA

TESIS DOCTORAL

INFERENCIA CAUSAL: APLICACIONES AL ÁMBITO DE LA EDUCACIÓN SECUNDARIA EN ESPAÑA

ROSA FERMINA SIMANCAS RODRÍGUEZ

DEPARTAMENTO DE ECONOMÍA

2016

UNIVERSIDAD DE EXTREMADURA

TESIS DOCTORAL

INFERENCIA CAUSAL: APLICACIONES AL ÁMBITO DE LA EDUCACIÓN SECUNDARIA EN ESPAÑA

ROSA FERMINA SIMANCAS RODRÍGUEZ

DEPARTAMENTO DE ECONOMÍA

Conformidad de los Directores:

Fdo: Francisco Pedraja Chaparro

Fdo: Daniel Santín González

2016

A mi familia y a Javi

AGRADECIMIENTOS

Siempre vi la meta tan lejana que hoy, que casi la puedo tocar, me parece mentira. Ha sido un viaje largo, con curvas, baches y algún pinchazo que otro, pero apasionante e inmensamente enriquecedor, profesional y personalmente. Ahora toca ser honesta conmigo misma y con la gente que me rodea, es el momento de agradecer a todo aquel sin el cual, hoy no podría estar escribiendo estas líneas.

Gracias a mis directores Francisco Pedraja y Daniel Santín por ser mis guías en este viaje, por vuestros sabios consejos y vuestra inmensa paciencia. Gracias Paco y Dani por regalarme vuestro tiempo, tiempo que espero haber sabido aprovechar e invertir de la forma correcta. Gracias Paco por aceptarme, hace ya bastantes años, como becaria de colaboración; sin saberlo, sería la semilla de lo que soy ahora. Gracias por aconsejarme profesionalmente incluso antes de iniciar este viaje y estar ahí para escuchar mis avances y retrocesos en el mundo laboral, los cuales me han llevado hasta la senda de hoy y en la que no podría estar más feliz. Gracias Dani por aceptar, a ciegas, co-dirigir mi tesis. Sé a ciencia cierta que tuve buenos ‘vendedores’ y que a tus oídos sólo llegaron cosas buenas sobre mí, pero aun así gracias por arriesgarte. Espero haber cumplido las expectativas. Gracias de nuevo por ser los mejores en el plano profesional sin olvidar en ningún momento el personal, por hacer tan gran equipo; soy realmente afortunada de haber podido trabajar bajo vuestra dirección.

Aprovecho estas líneas para agradecer a ciertas personas que han sido puntos claves en el camino. En primer lugar, gracias a Inés P. Murillo quien, en su inmensa humildad y considerando que era lo mejor para mí, rechazó tutorizar mi beca de colaboración y me llevó de la mano a conocer al que se convertiría en años siguientes en mi director, Paco. Gracias por creer en mí y descubrirme este mundo lleno de personas maravillosas. En segundo lugar quisiera agradecer a José Manuel Cordero, mi ‘co-director en la sombra’. Gracias Jose por hacerme pensar sobre la posibilidad de ser docente universitario, por mostrarme los pasos a seguir para conseguirlo, por ayudarme a preparar la documentación para optar a la beca pre-doctoral, y por un sinfín de cosas más, gracias. Nunca podré agradecerte lo suficiente como te has involucrado y te involucras en mi desarrollo profesional diario, respondiendo siempre con una sonrisa hasta las dudas más absurdas. No puedo olvidar a Mª del Mar Salinas y Agustín García, gracias por vuestro apoyo en todo lo relacionado con la docencia. Gracias a vuestros consejos, experiencias, libros y por qué no, material docente, he podido defender decentemente mis clases hasta ahora. ¡Gracias!

A todos aquellos que en su día tuve la suerte de tener como profesores y que ahora puedo llamar compañeros, aunque reconozco que aún no me siento a la altura, gracias por vuestra acogida. A mis compañeros del despacho de becarios, con vosotros empecé esta aventura y no podía cerrarla sin recordaros, gracias por todos los momentos vividos y por las experiencias compartidas. A las nuevas incorporaciones al grupo de investigación, gracias por los ratitos de 'alegrías' y 'penas', por el apoyo mutuo, por hacer crecer esta gran familia. Además, tengo que agradecer a esta profesión el haber conocido a personas estupendas en el camino, de toda la geografía nacional e internacional, que se han interesado y han seguido mi trayectoria, alentándome durante todo el proceso. Como podéis ver, así da gusto trabajar todos los días.

A la familia que he escogido, mis amigos, simplemente por ser como son y por estar ahí siempre, gracias. La vida me ha regalado muy buenos amigos que han confiado en mí incluso más que yo misma, que defienden mi trabajo como suyo propio y aceptan, sin rechistar, el poco tiempo que les dedico. ¡Gracias!, sin vosotros tampoco hubiera sido posible.

Por último agradecer a los más importantes, mi familia. En primer lugar a mis padres, sin lugar a dudas, sin vosotros yo no estaría hoy aquí. Habéis fomentado la formación de vuestros hijos desde la infancia, dejándonos decidir qué estudiar y dónde, siempre que eso nos hiciera feliz. Gracias por aceptar mis fracasos cuando yo no era capaz de hacerlo y mostrarme vuestro apoyo absoluto e incondicional. Gracias papá, eres mi ejemplo a seguir en la lucha por los sueños. Gracias mamá, por ser mi modelo de fortaleza y por recordarme día a día que la vida es para vivirla, no para sufrirla. Gracias a mi hermano por estar cuando realmente hace falta, por aceptar de buen grado que, aunque ya sea todo un hombre, le siga tratando como a un niño, porque me saca de mis casillas pero no lo cambio por nada, gracias hermanito. Y por supuesto a Javi, gracias por ser el mejor 'sufridor' y compañero de viaje que podía encontrar. Gracias por compartir esta aventura conmigo desde sus inicios, por ser mi soporte diario, el que aguanta mis días malos en casa y me alienta a seguir. Gracias por ser mi taxista, mi secretario, mi informático, mi ayudante particular, esta tesis también tiene parte tuya. A vosotros, mi familia, millones de gracias por ceder todo el tiempo que os correspondía y que esta tarea en la que decidí aventurarme os ha robado, sólo espero que los frutos merezcan la pena. ¡Os quiero!

Invertir en conocimientos produce siempre los mejores beneficios

Benjamin Franklin

Uno de los más grandes errores es juzgar a las políticas y los programas por sus intenciones, en vez de por sus resultados

Milton Friedman

CONTENIDO

INTRODUCCIÓN GENERAL

Desde una perspectiva microeconómica, la existencia de fallos de mercado legitima la intervención del sector público en la economía sobre la base de los principios de eficiencia y equidad. Tampoco el sector público, en su actuación, está libre de fallos por lo que resulta necesario evaluar esas intervenciones controlando la calidad de las mismas en relación a aquellos mismos principios (Albi y Onrubia, 2015).

Si la necesidad de evaluación de las políticas públicas está justificada en cualquier situación, en la actual de crisis económica y de la obligada consolidación presupuestaria, dicha necesidad de evaluación resulta ineludible. Analizar el grado de cumplimiento de los objetivos de las diferentes políticas de gasto es un paso imprescindible para la correcta asignación de los recursos escasos (Gertler *et al.*, 2011) y previo para decidir la continuidad de ciertas políticas y descartar otras que no resulten efectivas. Pero no sólo es importante evaluar para tomar esas decisiones, sino también lo es por su efecto de retroalimentación en el proceso de elaboración de las políticas públicas. De esta forma se genera información relevante que permite introducir las oportunas modificaciones en los programas con el fin de conseguir los objetivos propuestos o cambiarlos si las necesidades de la población han evolucionado de forma diferente a la inicialmente prevista (García-Pérez, 2009).

El interés por conocer el impacto de las políticas públicas, tanto cualitativa como cuantitativamente, ha dado lugar al desarrollo de métodos econométricos complejos que permiten ir más allá de las simples correlaciones tratando de comprobar la existencia o no de relaciones causales (Schlotter *et al.*, 2011) en lo que comúnmente se conoce como evaluación de impacto. Estas técnicas hacen posible evaluar el efecto de una determinada intervención comparando los indicadores clave entre los individuos beneficiarios del programa, conocidos como *grupo de tratados*, y los que no se benefician del mismo, identificados como *grupo de control o contrafactual.* La simple comparación entre unos y otros individuos genera resultados sesgados por problemas de autoselección o selección endógena, de ahí que el principal problema en las evaluaciones sea encontrar el *contrafactual* adecuado. Por lo tanto, el éxito en este tipo de evaluaciones dependerá de la correcta delimitación de ambos grupos, tratados y de control, lo que en la práctica se consigue utilizando diferentes metodologías que suelen clasificarse en *diseños experimentales* (aleatorios) y *diseños cuasi-experimentales* (no aleatorios).

En los *diseños experimentales*, se selecciona aleatoriamente una muestra de la población y, dentro de ella, los individuos son asignados al azar entre el grupo de tratados y el de control, consiguiendo dos conjuntos con características estadísticamente equivalentes salvo en lo que respecta a su participación en el programa objeto de evaluación. La principal ventaja de esta técnica es que el impacto de la política se mide comparando directamente los resultados medios obtenidos en ambos grupos. En el ámbito de las ciencias sociales, a diferencia de lo que ocurre en otros campos como el de ciencias médicas, existe una enorme dificultad a la hora de realizar experimentos aleatorios, bien porque las políticas públicas no son diseñadas para ser evaluadas, bien por otros motivos de tipo económicos, éticos, etc. Estas razones obligan a acudir a *diseños cuasi-experimentales*.

Los *diseños cuasi-experimentales* se emplean cuando la política objeto de evaluación ya ha sido implementada y no se diseñó para ser evaluada. En estos casos, los posibles grupos de tratamiento y control no se seleccionaron de manera aleatoria y se hace necesario conseguir un adecuado contrafactual. Aunque estos métodos son más rápidos y baratos, resultan menos fiables si no se controlan adecuadamente todas las diferencias existentes entre los grupos de tratamiento y de control. Para ello son necesarios modelos econométricos más complejos tales como variables instrumentales, diferencias en diferencias, regresiones en discontinuidad o técnicas de emparejamiento (*matching*)[1]. La decisión entre las anteriores alternativas, dependerá de la naturaleza del programa a evaluar, del tipo y calidad de los datos disponibles y de las condiciones exigidas a los individuos para participar en los correspondientes programas (Santín y Sicilia, 2015a).

La evaluación de las políticas públicas constituye un campo en pleno desarrollo en ámbitos tan distintos como la salud, el mercado laboral, la reducción de la pobreza o la educación. La cultura de evaluación está muy arraigada en el mundo anglosajón que lidera su desarrollo y difunde esas experiencias evaluadoras al resto de países. En España, sin embargo, no existe una tradición de evaluación por parte de los responsables políticos y sólo encontramos algunas experiencias de evaluación fruto del trabajo de determinados grupos de investigación. Concretamente, en el campo de la educación, a pesar de la importancia del capital humano en el crecimiento económico, las experiencias evaluadoras han sido escasas, a pesar de que la literatura viene poniendo de manifiesto que un incremento sistemático de los recursos no garantiza una mejora del rendimiento académico de los alumnos si tal aumento no lleva asociado una correcta utilización de los mismos (Woessmann, 2016). En la mayoría de los casos, esas pocas experiencias consisten en experimentos naturales que

[1] Para más información sobre las diferentes técnicas de evaluación de impacto véase Khander *et al.* (2010) y Angrist y Pischke (2008, 2014).

aprovechan la modificación de una política pública o que aplican técnicas cuasi-experimentales para analizar el impacto de políticas implantadas sin un diseño de evaluación previo.

Entre los primeros se encuentra el trabajo de Azmat e Iriberri (2010), que aprovecharon que durante el curso 1990/91 en un instituto del País Vasco se informó a los alumnos, de forma casual, de sus resultados académicos en términos relativos, es decir, en relación a los de la media de su clase. Utilizando datos de panel entre 1986 y 1994, comprobaron que el rendimiento de los alumnos que disponían de la información complementaria (resultados relativos), se incrementaba en un 5% de media, mientras que en aquellos que contaban solo con información individual, la mejora no se producía.

También dentro de este grupo están los trabajos de Nollenberger y Rodríguez-Planas (2015) y Felfe *et al.* (2015), quienes explotaron el experimento natural producido por la aprobación de la LOGSE de 1990 (*Ley de Ordenación General del Sistema Educativo*) que, entre otros cambios, ampliaba la gratuidad de la educación infantil a los niños de 3 años de edad (anteriormente dicha gratuidad afectaba a los niños de entre 4 y 6 años). Aunque la reforma fue nacional, al posibilitar su aplicación en las CCAA a lo largo de un plazo de diez años, permitió disponer de situaciones distintas en la implementación de esa política. Nollenberger y Rodríguez-Planas (2015) estudiaron el efecto de esa modificación legal en el empleo femenino, concluyendo que se incrementó en más de un 90% en mujeres de 30 años o más y en aquellas con dos o más hijos. Por su parte, Felfe *et al.* (2015) analizaron el impacto de la misma reforma sobre el resultado de los alumnos al final de la etapa de educación obligatoria. Sus resultados muestran que un incremento en la tasa de escolarización a los 3 años afectó positivamente a la promoción de curso durante la educación primaria y supuso una mejora en el desarrollo cognitivo que se mantiene en el largo plazo.

Por último, en el estudio de Santín y Sicilia (2015b), se comprueba el efecto de asistir a educación infantil sobre los resultados obtenidos en Educación Primaria utilizando la información proporcionada en la Evaluación General de Diagnóstico de 2009. En esa base de datos, un amplio grupo de escuelas declararon distribuir a sus alumnos de cuarto de primaria de forma aleatoria entre las diferentes aulas, lo que permitió a los autores explotar el hecho de que, por azar, en cada escuela hubiera un aula en la que, en media, sus alumnos contaran con un mayor número de años de asistencia a educación infantil. Los resultados mostraron un impacto positivo y significativo de asistir a educación infantil sobre las calificaciones en matemáticas y lectura.

Por otro lado, entre los trabajos que han aplicado técnicas cuasi-experimentales para determinar el impacto de diferentes políticas educativas, están los de Villaplana (2014) y Cabras y Tena (2013), que analizaron el impacto del incremento del número de ordenadores por alumno (política denominada *Programa 2.0*) y del uso de esos ordenadores en la docencia sobre el rendimiento en matemáticas. Partiendo de la información proporcionada por diferentes oleadas de PISA, en ambos trabajos, se concluye que el aumento del número de ordenadores en los centros es positivo siempre que se combine con una correcta metodología docente. Además, ese efecto resultó ser mayor en aquellos estudiantes que provenían de entornos más desfavorecidos.

En García-Pérez e Hidalgo-Hidalgo (2014) se analizó, mediante técnicas de *matching*, el impacto del Programa de Refuerzo, Orientación y Apoyo (PROA) sobre el rendimiento de los estudiantes españoles con datos de PISA 2012, encontrando un efecto positivo tanto en el corto como en el largo plazo.

De igual forma, el trabajo de Mediavilla (2014) estudió cómo las políticas de becas, llevadas a cabo en España durante el curso académico 2004-2005, incidieron sobre las tasas de graduación en educación postobligatoria. Aplicando *propensity score matching* a los datos proporcionados por la Encuesta de Condiciones de Vida de 2006, comprobó que este tipo de incentivos incrementaba, en un 40% (de media), la probabilidad de que el individuo beneficiario de la beca finalizase sus estudios postobligatorios.

Por último, Anghel *et al.* (2016) evaluaron un programa educativo bilingüe puesto en marcha en educación primaria en los colegios públicos de la comunidad autónoma de Madrid durante el curso académico 2004/05. El programa consistía en impartir las asignaturas de conocimiento del medio, plástica, educación física y/o música en inglés, además de la asignatura de inglés como lengua extranjera. Utilizando la información proporcionada por los exámenes externos anuales realizados por los alumnos de 6º de primaria en dicha región, los autores no encontraron efecto alguno del programa sobre los resultados en matemáticas y lengua (asignaturas impartidas en español). Sin embargo, sí encontraron un impacto negativo sobre los resultados obtenidos en aquellas asignaturas impartidas en inglés (conocimiento del medio), sobre todo entre los niños cuyos padres tenían un menor nivel educativo.

Este es el contexto en el que se desarrollan los ensayos que componen la tesis doctoral cuyo objetivo es aportar varias experiencias de evaluación en el ámbito educativo español, con el fin de ir enriqueciendo la escasa cultura evaluadora existente y la necesidad de la rendición de cuentas por parte de nuestro sector público.

La tesis se compone de tres ensayos. Los dos primeros se centran en cuestiones que han caracterizado al sistema educativo español en los últimos años, como sucede con la repetición de curso o la llegada de alumnos inmigrantes, y que son analizadas aplicando técnicas de evaluación de impacto. El tercero evalúa una política educativa específica puesta en marcha por el Gobierno de la Comunidad Autónoma de Extremadura.

En el primer capítulo se analiza la influencia de distintos factores en la repetición de curso, poniendo especial atención en el efecto del mes de nacimiento. Como se sabe, el sistema escolar español asigna a los alumnos que nacen en el mismo año natural a un mismo curso académico sin tener en cuenta las diferencias de madurez relativa entre los nacidos al principio y al final del año. Tras comprobar que la población española no sigue ningún tipo de comportamiento estratégico, basado en los resultados académicos esperados, a la hora de tener hijos, podemos considerar el mes de nacimiento como una variable exógena. En este sentido y utilizando los datos de PISA 2009 de la OCDE, se cuenta con un experimento natural para evaluar el impacto del mes de nacimiento sobre la probabilidad que tiene un alumno español de ser repetidor a los 15 años. Los resultados obtenidos refuerzan nuestra hipótesis inicial sobre la exogeneidad del mes de nacimiento y muestran que su impacto sobre la probabilidad de repetir se produce de forma bimestral. De hecho, los estudiantes nacidos en el último bimestre del año tienen una probabilidad de repetir un 80% superior a la de sus compañeros de clase nacidos en el primer bimestre.

El objetivo del segundo capítulo es evaluar el impacto del incremento exógeno de alumnos inmigrantes producido entre 2003 y 2009 partiendo de los datos españoles proporcionados por los sucesivos informes PISA de la OCDE. Para ello utilizamos el método de diferencias en diferencias, que nos permitirá comprobar si la concentración de inmigrantes tuvo algún efecto significativo sobre el desempeño de los alumnos. En nuestro caso, el *grupo de control* lo forman aquellos colegios que no cuentan con ningún estudiante de origen inmigrante en la muestra y que han mantenido esta situación a lo largo del tiempo, y el *grupo de tratamiento,* aquellos colegios con presencia de alumnado inmigrante y que además han visto incrementada la composición de ese tipo de alumnado con el paso del tiempo. Debido a que el porcentaje de alumnos inmigrantes difiere de unas escuelas a otras, ampliamos la metodología de diferencias en diferencias para introducir un tratamiento por dosis ya que no solo queremos comprobar el efecto medio de tener estudiantes inmigrantes en las escuelas, sino el efecto que produce su acumulación sobre los resultados. La conclusión fundamental derivada del análisis empírico es que la llegada de alumnos inmigrantes no da lugar, en media, a una reducción de las tasas de promoción de curso en las escuelas, sino que incluso es beneficioso para los nativos. Sin embargo, la acumulación de inmigrantes sí tiene un impacto negativo sobre los porcentajes de repetidores tanto de inmigrantes como de nativos, aunque, en este último caso, es necesaria una acumulación mayor para poder observar un efecto similar.

Por último, en el tercer capítulo evaluamos una política implantada por el Gobierno de la Comunidad Autónoma de Extremadura. Debido al elevado número de personas desempleadas que carecían de formación básica en la región donde más del 40% en 2012 no tenían el título de Educación Secundaria Obligatoria (ESO), el gobierno autonómico promovió el plan conocido como *Programa 18-25* cuyo fin era la obtención del título de la ESO mediante un incentivo económico de 1.000 euros dirigido a desempleados de entre 18 y 25 años. En este tercer capítulo se trata de evaluar el efecto de dicho programa sobre la población masculina durante el curso académico 2013/2014. Para ello aplicamos la técnica de regresiones en discontinuidad. Los resultados obtenidos ponen de manifiesto que el *Programa 18-25* no tuvo efecto significativo sobre los objetivos previstos.

Referencias Bibliográficas

Albi, E. y Onrubia, J. (Eds) (2015). *Economía de la gestión pública: cuestiones fundamentales*. Editorial Centro de Estudios Ramón Areces.

Anghel, B., Cabrales, A. y Carro, J.M. (2015). Evaluating a bilingual education program in Spain: the impact beyond foreign language learning. *Economic Inquiry*, 54(2), 1202-1223.

Angrist, J.D. y Pischke, J.S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Angrist, J. D. y Pischke, J.S. (2014). *Mastering' metrics: the path from cause to effect*. Princeton University Press.

Azmat, G. y Iriberri, N. (2010). The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics*, 94(7), 435-452.

Cabras, S. y Tena, J.D. (2013). Estimación del efecto causal del uso de ordenadores en los resultados de los estudiantes en la prueba PISA 2012. En INEE (ed.): *PISA 2012. Informe Español. Volumen II: Análisis Secundario*. Madrid: Instituto Nacional de Evaluación Educativa.

Felfe, C., Nollenberger, N. y Rodríguez-Planas, N. (2015). Can't buy mommy's love? universal childcare and children's long-term cognitive development. *Journal of Population Economics*, 28(2), 393-422.

García-Pérez, J. y Hidalgo-Hidalgo, M. (2014). Evaluación de PROA: su efecto sobre el rendimiento de los estudiantes. Evaluación de programas de cooperación territorial. Madrid: Instituto Nacional de Evaluación Educativa.

Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B. y Vermeersch, C.M. (2011). *Impact evaluation in practice*. Washington DC: World Bank.

García-Pérez, J. I. (2009). *Metodología y diseño de estudios para la evaluación de políticas públicas*. Antoni Bosch editor.

Khandker, S.R., Koolwal, G.B., y Samad, H.A. (2010). *Handbook on impact evaluation: quantitative methods and practices*. Washington DC: World Bank.

Mediavilla, M. (2014). ¿Son efectivas las becas en España? Una evaluación de impacto para el nivel secundario post-obligatorio. *RIEE. Revista Iberoamericana de Evaluación Educativa*, 7(1), 121-139.

Nollenberger, N. y Rodríguez-Planas, N. (2015). Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from Spain. *Labour Economics*, 36, 124-136.

Santín, D. y Sicilia, G. (2015a). Evaluar para mejorar: hacia el seguimiento y la evaluación sistemática de las políticas educativas. En Esteban, M. (ed.): *Reflexiones sobre el sistema educativo español*. Madrid: Fundación Europea Sociedad y Educación.

Santín, D. y Sicilia, G. (2015b). El impacto de la educación infantil en los resultados de primaria: Evidencia para Españar a partir de un expermiento natural. En Esteban, M. (ed.): *Reflexiones sobre el sistema educativo español*. Madrid: Fundación Europea Sociedad y Educación.

Schlotter, M., Schwerdt, G. y Woessmann, L. (2011). Econometric methods for causal evaluation of education policies and practices: a non-technical guide. *Education Economics*, 19(2), 109-137.

Villaplana, C. (2014). Evaluación del programa Escuela 2.0 para la asignatura de Matemáticas a partir de PISA 2012. *Investigaciones de Economía de la Educación*, 9, 631-652.

Woessmann, L. (2016). The Importance of School Systems: Evidence from International Differences in Student Achievement. Documento de trabajo 10001, IZA.

INTRODUCTION

From a microeconomic perspective, the existence of market failures justifies the intervention of the public sector in the economy, based on the principles of efficiency and equity. However, the public sector is not free to fail in its action, so it is necessary to evaluate these interventions by controlling the quality of them with regards to those same principles (Albi and Onrubia, 2015).

If the need for evaluation of public policy is justified in any situation, in the current economic crisis and the necessary budgetary consolidation, such need becomes essential. Analysing the degree of compliance with the objectives of the different spending policies, it is a compulsory step for the proper allocation of resources which are more than scarce (Gertler *et al.*, 2011) and also a previous step to decide the continuity of certain policies and rule out others that are not effective. But it is not only important to evaluate with the aim of guiding decision-making process in public policies, but also in its feedback effect during the monitoring process creating relevant information that allows the decision-makers to make proper modifications in the programs, in order to achieve the goals or change them if the needs of the population have evolved differently than initially expected (García-Pérez, 2009).

The need to know the impact of a public policy, both qualitatively and quantitatively, has resulted in the development of complex econometric methods to go beyond simple correlations and thus being able to verify the existence of causal relationships (Schlotter *et al.*, 2011), in what is commonly known as impact evaluation. These techniques allow us to assess the effect of an intervention by comparing key outcomes among individual beneficiaries of the program, known as *treated group*, and those who do not benefit from it, identified as a *control group* or *counterfactual*. A simple comparison between these individuals creates biased result due to self-selection problems or endogenous selection, so we can infer that the main problem in evaluations is finding the right *counterfactual.* Therefore, the success in this type of assessment will depend on the correct definition of the two groups, both the treated and the control, which in practice is achieved by using different methodologies that are often classified in *experimental designs* (randomized) and *quasi-experimental designs* (non-randomized).

In *experimental designs*, a sample of the population is selected randomly and, within the selection, individuals are randomized again between the treated group and the control group, getting two groups of individuals with features statistically equivalent except for their participation in the program under evaluation. The main advantage of this technique is that the impact of the policy is directly measured by comparing the average results obtained in both groups. In the field of social sciences, unlike what happens in other fields such as medical science, it is difficult to carry out randomized experiments, due to several reasons (economic, ethical, etc.). Other times, public policies are not designed to be evaluated. Both reasons forces in most occasions to resort to *quasi-experimental designs*.

Quasi-experimental designs are used when policy under evaluation has already been implemented and was not intended to be evaluated. In these cases, possible treatment and control groups were not selected randomly and getting a proper counterfactual becomes necessary. Although these methods are faster and cheaper, they are less reliable if all differences between the treatment and control groups are not properly controlled. For this, more complex econometric models such as instrumental variables, difference in differences, regression discontinuity designs or matching [2] are require d. The choice between the aforementioned methods will depend on the nature of the programme under evaluation, the type and quality of data available and the requirements for individuals to participate in the programmes (Santín and Sicilia, 2015a).

The evaluation of public policies is a field in full development in areas as diverse as health, labour market, poverty reduction or education. The evaluation culture is deeply rooted in the Anglo-Saxon world, which leads its development and disseminates these experiences into other countries. In Spain, however, there is no tradition of impact evaluation by the policymakers, and only some experiences can be found as a result of the work of certain research groups.

Specifically, in the field of education, despite the importance of human capital in economic growth, evaluation experiences are scarce and literature has proven that a systematic increase in resources does not guarantee the improvement of the students' academic performance if it does not go together with a proper use (Woessmann, 2016). In most cases, they consist of natural experiments that take advantage of changing public policy or quasi-experimental techniques applied to analyse the impact of policies implemented without prior evaluation design.

[2] For further information about impact evaluation methods, see Khander *et al*. (2010) and Angrist and Pischke (2008, 2014).

Among the first ones, we can find the study of Azmat and Iriberri (2012), who took advantage of what happened during the 1990/91 academic year at an institute of the Basque Country in which students were told their academic performance in regards to the average of their class at the end of each semester. Using panel data between 1986 and 1994, they found that the performance of students who possessed the additional information (relative performance), was increased by 5% on average, while those who had only individual information did not experience such improvement.

Also, Nollenberger and Rodríguez-Planas (2015) and Felfe *et al.* (2015) exploited the natural experiment produced by the adoption of LOGSE of 1990 (*Ley de Ordenación General del Sistema Educativo Español*, a Spanish Education Act) which extended the gratuity of pre-primary education to 3-year-old children (prior to such measure it was free for 4- to 6-year-old children) belong to this group. Although the reform was at national level, due to the fact that its application in the different regions could happen over a period of ten years, it was possible to have different situations in the implementation of that policy. Nollenberger and Rodríguez-Planas (2015) studied the effect of this legal amendment in female employment, concluding that it increased by more than 90% for 30 year-old women or older and for those with two or more children. Meanwhile, Felfe *et al.* (2015) analysed the impact of the reform itself on student results at the end of compulsory secondary education. Their results show that an increase in enrolment rate for 3-year-old children affected positively to promotion in primary education and it also meant an improvement in cognitive development, which was maintained in the long term.

Finally, in Santín and Sicilia (2015b), the effect of attending pre-primary education on the results obtained in primary education using the information provided by General Diagnostic Assessment 2009 database (*Evaluación General de Diagnóstico* in Spanish) is proved. In that database, a large group of schools stated that students of fourth grade were randomly assigned into classrooms, which allowed the authors to exploit the fact that, by chance, in every school there was a classroom in which, on average, their students had a greater number of years attending pre-primary education. The results showed a positive and significant impact of attending pre-primary education on reading and math scores.

On the other hand, Villaplana (2014) and Cabras and Tena (2014) can be found among those studies that have applied a quasi-experimental design to determine the impact of different educational policies. Both papers analysed the effect of increasing the number of computers per student (policy known as *Programa 2.0*) and the use of those computers in teaching on mathematics performance. Using information provided by different waves of PISA, both studies concluded that the increase in the number of computers in schools is positive provided that combined with proper teaching methodology. Furthermore, this effect was greater in those students who came from more disadvantaged backgrounds.

In García-Pérez and Hidalgo-Hidalgo (2014), the impact of Programme for Reinforcement, Support and Guidance (*Programa de Refuerzo, Orientación y Apoyo*) on Spanish students' performance using data from PISA 2012 applying matching was analysed, finding a positive effect in the short and long term.

Likewise, Mediavilla (2014) studied how scholarship policies, carried out in Spain during the 2004/05 academic year, influenced on graduation rates in post-compulsory education. Applying propensity score matching to data provided by Living Conditions Survey 2006, it was found that such incentives increased by 40% (on average), the likelihood that the beneficiaries would complete their post-compulsory studies.

Finally, a bilingual education programme introduced in primary education in the 2004/2005 academic year in the region of Madrid was evaluated by Anghel *et al.* (2016). This programme consisted on teaching some other subjects in English additionally to English as a foreign language. Using the information provided by standardised exams administered each year in all primary school in that region to 6th-grade students, the authors found out no effect on mathematics or reading scores (those subjects taught in Spanish) and a substantial negative impact on the results obtained in subjects taught in English, above all among children from less educated parents.

In this context the present PhD dissertation aims to provide several evaluation experiences about the Spanish education system, in order to enrich the scarce evaluation culture and the need for accountability public sector in Spain. This PhD dissertation is composed of three separate essays. The first two focus on issues that have defined the Spanish education system in recent years and the third one assesses a specific educational policy implemented by the government of Extremadura.

The first chapter analyses the influence of several factors on the probability of repeating a grade paying special attention to the effect of the month of birth. In Spain, students born in the same calendar year start school together, although it can be argued that there are differences in relative maturity between children who were born at the beginning and at the end of the year. After verifying that parents do not plan the birth of their children based on expected educational outcomes, the month of birth can be considered as an exogenous variable. Hence, using OECD's PISA 2009 data, we have a natural experiment in order to evaluate the impact of the month of birth on the probability of having repeated a grade by the age of 15 years. The results provide evidence that reinforces our initial hypothesis about the exogenous effect of the birth month. Furthermore, a bimonthly impact on grade retention can also been identified since the students who were born in the last two months of the year present a probability to repeat a grade 80% higher than their classmates who were born in the first two months of the year.

The goal of the second chapter is to assess the impact of the exogenous increase in the number of immigrant students from 2003 to 2009 using Spanish data from consecutive OECD's PISA reports. For this purpose, we use the Difference-in-Differences method (DiD) capable of detecting whether the immigrant concentration had a significant effect on student performance. Within this framework, the *control group* will be the schools without sampled immigrants that maintained this situation over time, and the *treatment group* will be those schools with immigrant students, which experienced a significant increase of immigrants along this period. As the percentage of immigrants is different across schools, DiD methodology is adapted to deal with a *dose treatment*, so what we seek is not simply the average effect of having or not foreign students at the school, but the effect of its accumulation on results. The main conclusion derived from empirical analysis is that the arrival of immigrant students does not on average decrease school promotion rates and is even beneficial to native students. However, the concentration of immigrant students at the same school over a threshold does have a negative impact on the percentage of repeaters, for both immigrants and natives, although in the latter case greater concentrations of immigrant students are required to observe similar effects.

Finally, in the third chapter a policy implemented by the Government of the region of Extremadura is evaluated. Due to the high number of unemployed and uneducated people in such region (over 40% in 2012 according to the regional Public Employment Service), the regional government promoted the programme known as *Programa 18-25* whose purpose was to reduce those figures. This policy was targeted to unemployed people aged between 18 and 25 years who had not completed the compulsory secondary education aiming to motivate these people to get back into education system, through an economic incentive of 1,000 euros. The impact of this programme on the male population during its second call (the 2013/14 academic year) is measured, by applying a regression discontinuity design. The results show that being a beneficiary from the programme did not involve an increase in the probability of obtaining the diploma of upper secondary education and neither increased the likelihood that students passed all modules in which they were enrolled.

References

Albi, E. and Onrubia, J. (Eds) (2015). *Economía de la gestión pública: cuestiones fundamentales*. Centro de Estudios Ramón Areces.

Anghel, B., Cabrales, A. and Carro, J.M. (2015). Evaluating a bilingual education program in Spain: the impact beyond foreign language learning. *Economic Inquiry*, 54(2), 1202-1223.

Angrist, J.D. and Pischke, J.S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Angrist, J. D. and Pischke, J.S. (2014). *Mastering' metrics: the path from cause to effect*. Princeton University Press.

Azmat, G. and Iriberri, N. (2010). The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics*, 94(7), 435-452.

Cabras, S. and Tena, J.D. (2013). Estimación del efecto causal del uso de ordenadores en los resultados de los estudiantes en la prueba PISA 2012. En INEE (ed.): *PISA 2012. Informe Español. Volumen II: Análisis Secundario*. Madrid: Instituto Nacional de Evaluación Educativa.

Felfe, C., Nollenberger, N. and Rodríguez-Planas, N. (2015). Can't buy mommy's love? universal childcare and children's long-term cognitive development. *Journal of Population Economics*, 28(2), 393-422.

García-Pérez, J. and Hidalgo-Hidalgo, M. (2014). Evaluación de PROA: su efecto sobre el rendimiento de los estudiantes. Evaluación de programas de cooperación territorial. Madrid: Instituto Nacional de Evaluación Educativa.

Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B. and Vermeersch, C.M. (2011). *Impact evaluation in practice*. Washington DC: World Bank.

García-Pérez, J. I. (2009). *Metodología y diseño de estudios para la evaluación de políticas públicas*. Antoni Bosch editor.

Khandker, S.R., Koolwal, G.B., and Samad, H.A. (2010). *Handbook on impact evaluation: quantitative methods and practices*. Washington DC: World Bank.

Mediavilla, M. (2014). ¿Son efectivas las becas en España? Una evaluación de impacto para el nivel secundario post-obligatorio. *RIEE. Revista Iberoamericana de Evaluación Educativa*, 7(1), 121-139.

Nollenberger, N. and Rodríguez-Planas, N. (2015). Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from Spain. *Labour Economics*, 36, 124-136.

Santín, D. and Sicilia, G. (2015a). Evaluar para mejorar: hacia el seguimiento y la evaluación sistemática de las políticas educativas. En Esteban, M. (ed.): *Reflexiones sobre el sistema educativo español*. Madrid: Fundación Europea Sociedad y Educación.

Santín, D. and Sicilia, G. (2015b). El impacto de la educación infantil en los resultados de primaria: Evidencia para Españar a partir de un expermiento natural. En Esteban, M. (ed.): *Reflexiones sobre el sistema educativo español*. Madrid: Fundación Europea Sociedad y Educación.

Schlotter, M., Schwerdt, G. and Woessmann, L. (2011). Econometric methods for causal evaluation of education policies and practices: a non-technical guide. *Education Economics*, 19(2), 109-137.

Villaplana, C. (2014). Evaluación del programa Escuela 2.0 para la asignatura de Matemáticas a partir de PISA 2012. *Investigaciones de Economía de la Educación*, 9, 631-652.

Woessmann, L. (2016). The Importance of School Systems: Evidence from International Differences in Student Achievement. Documento de trabajo 10001, IZA.

CAPÍTULO 1

DETERMINANTS OF GRADE RETENTION IN SPAIN: DOES BIRTH MONTH MATTER?

1.1 INTRODUCTION

The phenomenon of grade retention in Spain affects around one third of all students and has become a considerable obstacle to future economic growth. According to the OECD's PISA 2009 report, 36% of 15-year-old Spanish students had repeated at least one grade. Several studies provide empirical evidence that students who repeat an academic year (from now on repeaters) are at greater risk of school failure, and this variable is a good predictor for early school leaving (Benito, 2007; Calero *et al*., 2010; Jimerson *et al*., 2002). According to Eurostat, the school failure figures for Spain have been consistently high over the last decade at 30%, which is twice the current EU-27 average. This is a central issue for European governments, and the Europe 2020 strategy encourages educational policy measures to reduce school dropout rates to under 10% by 2020.

School failure has negative consequences for both individuals and the efficiency of the whole economy. On the one hand, early school leaving generates major labour insertion problems and a higher risk of social and economic exclusion. This is a major concern given the severity of the current economic crisis in Spain, where, according to Eurostat, the unemployment level among young people (population aged under 25 years) was 53.2% in 2012 (8.2% in Germany). On the other hand, school failure is associated with a lower stock of human capital and lower labour force productivity, higher social public expenditure, and lower economic growth prospects (Asteriou and Agiomirgianakis, 2001; Duval and de la Maisonneuve, 2010; Hanushek and Kimko, 2000; Psacharopoulos 2007).

In the Spanish education system, students born in the same calendar year start school in the same academic year. By law, pre-primary education is optional and free for three- to five-year-olds. Compulsory education starts with primary education; pupils enter primary school in September of the year that pupils turn six, continuing for just ten years until pupils reach the age of 16 at the end of secondary education. As the cut-off date is January 1st, students born in January are almost one year older than their classmates born in December. Previous research provides evidence of a maturity gap between children born in January and December; therefore, this policy could potentially have an impact on students' future academic performance if teachers confuse maturity with learning ability (Allen and Barnsley, 1993).

In this context, the aim of this chapter is to evaluate whether a pupil's relative age with his/her age cohort may have a significant long-term effect on the specific problem of grade retention in the Spanish education system. For the purpose of causal identification, one of the key issues of this research is to show that birth month is an exogenous variable in the analysis. To do this, we demonstrate that parents do not target birth dates on the basis of their children's expected future academic performance; besides, there are legal constraints preventing parents from choosing their children's enrolment cohort. Therefore, we have a natural experiment framework in which we can distinguish the cause-effect relationship between birth month and the probability of repeating any year from any accidental correlation.

As mentioned above, grade retention has a number of negative effects on students, ranging from problems of self-esteem to higher school dropout rates (Agasisti and Cordero, 2013; Jimerson *et al.*, 2002; Manacorda, 2012). Were birth month found to matter, this would place a constraint not only on the efficiency of the Spanish economy but also on the equal opportunities policy established by LOE[1] (*Ley Orgánica de Educación*, the Spanish Education Act), and it would justify the search for public educational policies designed to avoid or reduce this problem.

The chapter is structured as follows. The following section summarizes the existing literature about the analysed topic. Section 1.3 presents and justifies the experimental design together with the database. Section 1.4 presents the empirical results, and the article winds up with the main findings of this research, as well as with some educational policy proposals to reduce the birth month effect.

1.2 LITERATURE REVIEW

There is a lot of literature on the determinants of academic performance and school failure. However, literature evaluating whether or not birth month-induced age differences in the same age cohort have a direct influence on grade retention like us is harder to find. Most

[1] LOE article 1.b. states that education must ensure equal opportunities, educational inclusion and non-discrimination. Education should act as a means to offset personal, cultural, economic and social inequalities, especially any caused by disability.

previous papers investigate the impact of birth month on the student outcomes[2]. This literature can be classified according to how the authors defined the birth month variable. Some researchers compare achievement for children born in different quarters of a year. Allen and Barnsley (1993), for example, compare percentages and apply chi-squared tests to data from a specific survey of Canadian and English schools, showing that there are educational differences by birth quarter and claiming that these differences persist and even increase in the long term. Bedard and Dhuey (2006) apply instrumental variables and also observe that initial maturity differences have long-lasting effects on student performance across some OECD countries using data from TIMSS 1995-1999. Strom (2004) compares mean scores of Norwegian students using PISA 2000 data and proves that children born in the fourth quarter of the year have lower educational results and a higher likelihood of being held back or requiring special education. Along the same lines, Sprietsma (2010) identifies a long-term (non-linear) age effect on both the probability of repeating a grade and academic outcomes using PISA 2003 data. Gutierrez-Domenech and Adserà (2012) ran a multivariate model on primary student data from a 2005 *Família i Educació a Catalunya* survey and likewise found that performance by younger students is poorer than for their peers and that this disadvantage does not disappear over time.

An alternative line of research followed by different authors is to restrict the sample to only older and younger students, classified according to a specific cut-off date (established by the government in order to determine children's entry to the education system). Some examples are papers by Kawaguchi (2011), based on a regression discontinuity design with data from TIMSS 2003 and the Employment Status Survey, Crawford *et al*. (2011), using the English National Pupil Database, and Ponzo and Scoppa (2014), exploiting the information provided by PIRLS 2006, TIMSS 2007 and PISA 2009 about Italian students. They find evidence of significant differences between children born before and after the cut-off date in terms of their educational attainment, i.e. older children in the same school cohort do better than younger ones and although these differences decrease over time, they are still significant among students aged from 16 to 18 years.

The grade retention variable is usually considered as a major determinant rather than the dependent variable of educational performance. This is the line taken by Manacorda (2012), who exploits specific data from Uruguay, and Eide and Showalter (2001), also using

[2] The Appendix to this chapter summaryzes previous contributions relating birth month, educational performance and grade retention.

a particular database (High School and Beyond). Both papers report the negative impact of grade retention on educational performance, causing higher dropout rates and lower future earnings. The same issue is analysed by Calero *et al*. (2010) for Spain concluding that grade retention significantly explains low educational performance calling into question grade retention as an efficient strategy for reducing the risk of school failure.

There is substantially less research and literature targeting our objective: analyse the influence of birth month on the probability of repeating a year. Corman (2003) studies the influence of certain variables (including students' birth month) on the probability of retention through a multivariate probit analysis using the United States National Household Education Survey database; the conclusion is that children born in the fourth quarter of the year are six percentage points more likely to repeat a grade than children born in the first quarter. In Spain, Calero (2006) applies a multinomial logistic model using EU Household Panel data and sets a four-level dependent variable: 'in compulsory secondary education or primary education' (this represents the group of repeater students)[3], 'in post-compulsory education', 'in intermediate vocational training' or 'in work'. The birth month is introduced in the model as a dummy variable (children born in the last quarter of the year compared to others), and the research reports that the youngest students in the same school cohort have greater learning difficulties, which increases their probability of repeating a year.

All things considered, this chapter contributes to existing research by providing evidence of the effect of birth month on the probability of 15-year-old Spanish students repeating a grade; besides, we introduce a bimonthly aggregation of this explanatory variable, since this is the smallest time interval that showed statistically significant differences.

1.3 RESEARCH DESIGN

1.3.1 Data

The dataset used for the research comes from the PISA (*Programme for International Student Assessment*) survey, designed and implemented by the OECD in the late 1990s as a

[3] The reason is that these students should have completed compulsory education by that age.

comparative, international, regular and continuous study of certain characteristics and skills of students worldwide (Turner, 2006). The PISA target population is composed of students aged between 15 and 16 years at the time of the assessment, all of whom are born in the same year and who have completed at least six years of formal schooling. PISA measures their performance in math, reading, and science. It also gathers information about students' personal background and school environment, for which purpose two questionnaires are administered, one addressed to school principals and another to students themselves[4]. These surveys have taken place every three years since the year 2000 focusing on one of the above three competences each year.

An important aspect to be taken into account in an empirical analysis using PISA data is that the data are gathered by means of a two-stage sampling procedure. First, a sample of schools is selected in every country from the full list of schools containing the total student population. Then, a sample of 35 students is randomly selected within each school. As a result, statistical analyses have to consider sampling weights to ensure that students adequately represent the analysed total population (Rutkowski *et al*., 2010)[5]. A sample of about 25,998 students from 889 schools from Spain took part in PISA 2009 (OECD, 2010).

1.3.2 Is birth month exogenously distributed?

The key variable in this research is the students' birth month (BM). First of all, we need to find out whether this variable is exogenously distributed among students regardless of other factors or, on the contrary, parents target their children's birth date with the aim of maximizing their future academic performance. In this last case, most births should occur in the first few months of the natural year, and fewer births should be observed at the end.

We use a Kolmogorov-Smirnov test to check for a uniform distribution that would demonstrate the exogeneity of the birth month variable. We compare the observed distribution of births in 1993 (birth year of the student respondents) to the expected

[4] Parents complete a third questionnaire. However, this information is only available for a limited number of countries and, unfortunately, Spain is not one (OECD, 2010).

[5] These weights include adjustments for non-response by some schools and students within schools and weight cutting to prevent a small set of schools or students having undue influences. These processes are based on intensive calculation methods, known as 'resampling' methods, which consist of taking multiple samples from the original sample. Specifically, PISA uses the Balanced Repeated Replication (BRR) with 80 replicates. For an extensive description of this procedure, see OECD (2005; 2009).

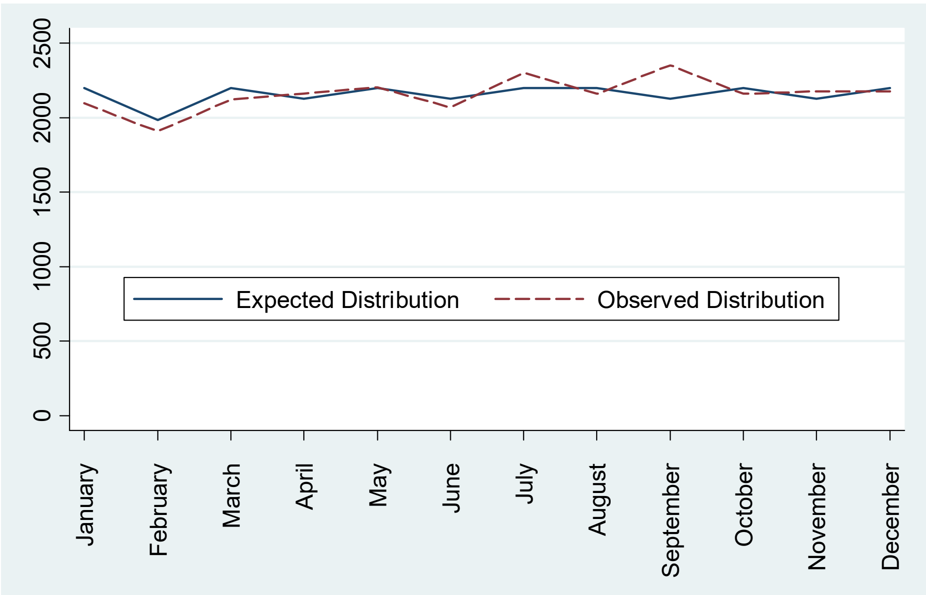
distribution of births according to the average daily births in that year based on the information provided by PISA 2009. Table 1 reports both distributions, and Figure 1 plots the monthly deviation in the number of observed births with respect to their expected distribution over the year.

Table 1. Observed and expected distribution of births by months in Spain

Month	Expected Distribution	Observed Distribution
January	2,199	2,096
February	1,986	1,908
March	2,199	2,122
April	2,128	2,162
May	2,199	2,204
June	2,128	2,069
July	2,199	2,303
August	2,199	2,160
September	2,128	2,352
October	2,199	2,160
November	2,128	2,177
December	2,199	2,176
Total	**25,887**	**25,887**

Source: Authors' own elaboration using data from PISA 2009 database (OECD, 2010).

Figure 1. Monthly deviation in observed births with respect to expected births



Source: Authors' own elaboration using data from PISA 2009 database (OECD, 2010).

As Table 1 shows, the expected number of births is not exactly the same every month because months contain different numbers of days. The Kolmogorov-Smirnov test finds that both distributions are similar at a 95% confidence level (*p*-value = 0.391). This finding confirms that births are randomly distributed throughout the year, and hence we consider birth month to be an exogenous random variable. From this result, we can conclude that parents in Spain do not plan their children's birth date with the aim of improving their educational outcomes[6].

Figure 1 shows that the observed and expected distributions of births are clearly similar. The real number of births is slightly higher than expected in July and September. These births may be planned for the purpose of adding the holiday month on to the maternity leave period.

1.3.3 Birth month and grade retention

As already mentioned, the Spanish educational legislation on school starting age establishes that one cohort must be composed of every student born in the natural year, and this is the only option open to parents[7]. Therefore, we have a 'natural experiment' because the birth month appears to be an exogenous variable with respect to the dependent variable considered in this research: the probability of having repeated a grade at the end of secondary education.

Instead of assuming a hypothetical year division like other researchers, we first set out to discover the shortest period of time, in months, that showed up statistically significant differences with respect to its nearest alternatives. In other words, how many months have to be aggregated to find statistical significant differences in grade retention? For this purpose, we explore several cross tabulations (one cross-tab for every possible child grouping by birth month: monthly, bimonthly, quarterly, and so on) in order to compare the percentages of repeaters among pupils born at different times of the same calendar year. If our hypothesis that birth month influences the probability of repeating a grade is true, there should be an upward trend in the percentages of repeaters as the year progresses, since children born earlier in the year are less likely to repeat a grade than students born later in the same year.

[6] We also carried out a one-way analysis of variance to find out if we could reject equal mean socioeconomic levels (the ESCS variable is defined in Section 3.4.) by birth month. The results of the pairwise comparison Bonferroni tests were not statistically significant at a 95% confidence level, further corroborating the conclusion that birth month is exogenously determined in Spain regardless of socioeconomic status.

[7] Legally, parents cannot keep their children at a pre-primary level for an additional year or postpone their children's entry to the first year of primary school.

At the same time, these percentages should be statistically and significantly different from each other.

In order to check their statistical significance, we use a chi-squared test with (r-1) (c-1) degrees of freedom (r denotes number of rows and c denotes number of columns) at a 95% confidence level, which is useful for testing the equality of proportions:

$$x^2_{(r-1)(c-1);\,0,05} = \sum_{ij} \frac{\left(f_{ij}-E_{ij}\right)^2}{E_{ij}} \tag{1}$$

where f_{ij} represents the observed frequency and E_{ij} is the expected frequency[8]. If this test rejects the null hypothesis (i.e. at least one proportion is different from any other), we run a chi-squared test by cell in order to determine where the differences are. On this occasion, we employ a chi-squared test with one degree of freedom again at a 95% confidence level. In this research, we are interested in only the shortest period of time that shows up statistically significant differences in every cell. In Spain, the shortest period of time that meets all the above requirements is the bimester, i.e. a bi-monthly aggregation of the births throughout the year. Results of repeaters by bimester are shown in Table 2.

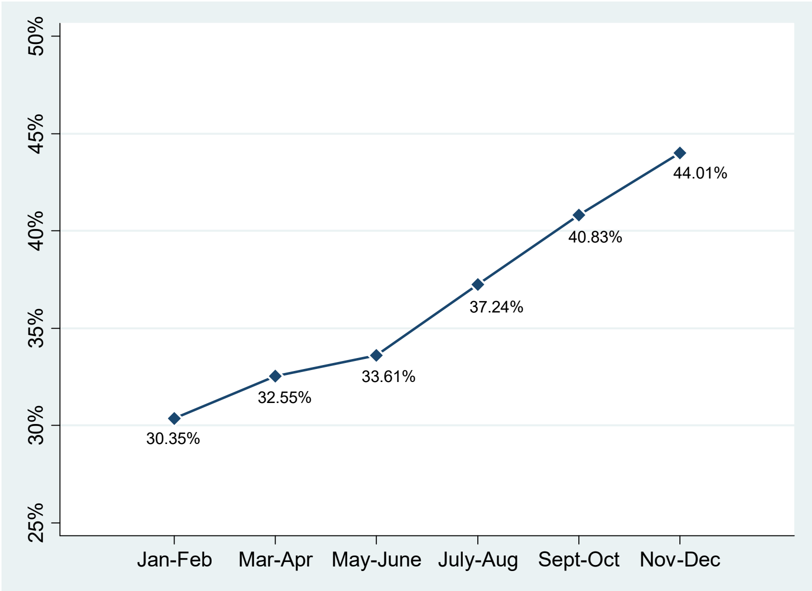
Table 2. Percentage of repeaters and non-repeaters by birth bimester

Bimester	**Repeater**	**Non Repeater**
January-February	30.35%	69.65%
March-April	32.55%	67.45%
May-June	33.61%	66.39%
July-August	37.24%	62.76%
September-October	40.83%	59.17%
November-December	44.01%	55.99%
Total	**36.56%**	**63.44%**

Source: Authors' own elaboration using data from PISA 2009 database (OECD, 2010).

[8] The expected frequency for the cell in the i^{th} row and the j^{th} column is the total number of subjects in row i by the total number of subjects in column j, divided by the total number of subjects in the whole table.

Figure 2. Distribution of repeater students by birth bimester



Source: Authors' own elaboration using data from PISA 2009 database (OECD, 2010).

On average, the percentage of repeaters increases significantly from one bimester to the next by 2.73 percentage points. Grade retention rates for pupils born in November and December are clearly 14 points higher than for pupils born in January and February. However, this is not a linear increment (the average increase in the first half of the year is 1.63 percentage points, rising to 3.39 percentage points in the last half). A preliminary conclusion related to this result is that the influence of the birth month becomes more pronounced as the year advances. Figure 2 illustrates the above phenomenon.

Note that all time periods longer than a bimester (quarter and semester) met the requirement as well. To the best of our knowledge, this is the first time that bi-monthly aggregation has been employed to analyse the influence of birth month on educational performance, since other authors always introduce this variable aggregated by longer time periods.

1.3.4 Variables

We have already analysed the influence of birth month on the probability of having repeated a grade by the age of 15 years, comparing percentages of repeaters and non-repeaters depending on their birth month. Nevertheless, this methodology can be extended to account for other control variables related to pupils, families, and schools, which may also have an impact on the dependent variable. For this purpose, we estimate a logistic regression including several control variables related to students' background apart from birth month. Logistic regression coefficients are especially useful for estimating odds ratios for each independent variable in the model. Odds ratios measure the probability of an event occurring over the probability of it not occurring. The regression we estimate is as follows:

$$Prob(R_{is} = 1) = \frac{e^{\alpha+\beta_1 BM_{is}+\beta_2 X_{is}}}{1+e^{\alpha+\beta_1 BM_{is}+\beta_2 X_{is}}} = \frac{1}{1+e^{-(\alpha+\beta_1 BM_{is}+\beta_2 X_{is})}}, \qquad (2)$$

where R_{is} denotes whether the student i in school s is a repeater (R_{is}=1) or not (R_{is}=0), BM_{is} is the student's birth month and X_{is} is the vector of control variables.

Regarding the variables, we use R_{is} as a dependent variable. At 15 years old, students are approaching the end of Spanish compulsory education and they should be in their 4th grade of ESO (*Enseñanza Secundaria Obligatoria*, compulsory secondary education in the Spanish education system), equivalent to 10th grade on the international scale. We consider that 15-year-old students who are not in 4th grade are repeaters.

The key variable referred to students' birth month, $BM_{is,}$ is aggregated bimonthly. In short, there are six categories classifying students according to the month in which they were born as follows: 'BM: January-February' represents students born in January and February, 'BM: March-April' represents students born in March and April, and so on. The remaining categories are 'BM: May-June', 'BM: July-August', 'BM: September-October', and 'BM: November-December', where the first bimester is the baseline category.

The set of control variables, which are exogenous with respect to the dependent variable and that will be introduced in the logistic regression model are:

Index of economic, social, and cultural status (ESCS): This is an index created by PISA from three variables related to the highest occupational status of parents, the highest educational level of parents in years of education according to ISCED, and educational possessions at home.

Gender: This variable will take the value 1 for boys and 0 for girls.

Immigrant status: This variable has the following categories: 'native students' are students born in Spain or students with at least one parent born in this country (which is the baseline category), 'second-generation immigrants' are students born in Spain but whose parents were born in another country, and 'first-generation immigrants' are foreign-born students whose parents are also foreign born. Students with missing responses for either their origin or their parents' origin have been saved in the category named 'uncertain origin', assuming that their refusal to answer these questions is because they have reasons for not wanting to disclose this information (Salinas and Santín, 2012).

Pre-primary attendance: All countries participating in PISA show a positive relationship between the proportion of students who received pre-primary education and average school system performance, even after accounting for the socioeconomic status (OECD, 2011). It is a four-category variable: 'pre-primary: non-attendance', students who report not having received pre-primary education; 'pre-primary: one year or less', students who attended pre-primary school for less than a year; and 'pre-primary: over a year', students who reported having attended pre-primary school for more than a year (which is the baseline category).

Family structure: Several studies highlight the influence of family situation on student performance. This variable is divided in three categories: 'single-parent family' composed of one parent and his/her children; 'mixed family' consisting of a couple in which one family member has children from a previous relationship; and 'nuclear family' composed of both parents and their children (which is the baseline category).

The following two variables concern parents' choice of private or public schools, and peer effects. They are possibly endogenous. This led us to run a second logistic regression model adding these variables related to parental choice in order to explore the influence of these variables on grade retention under the caveat that this model may show biased results.

School type: This variable is introduced in the regression in order to test whether school ownership influences the probability of enrolled students repeating a grade. It has three categories: 'public schools' (government funded and managed schools), 'private government-dependent schools' (government funded privately managed schools), and 'private schools' (independent schools). 'Public schools' will be the baseline category.

Peer effects: We adopt the common assumption that pupil learning is influenced by their classmates' characteristics. Although a random assignment of students across schools would be ideal, this is not the case in the Spanish education system where there is segregation depending on students' socioeconomic status. This variable is defined as the average ESCS of the pupils enrolled in the school where the student is assessed.

1.4 RESULTS

As mentioned in Section 3, we estimate two logistic regression models. In both models, the dependent variable is the repeater status at age 15 (R) and the key explanatory variable is the birth month: bimester (BM). In addition to this variable, we include the set of control variables defined in Section 1.3.4 related to individual, socioeconomic, and family characteristics (Model 1 in Table 3). Model 2 is an extension of Model 1, including variables related to school type and peer effects. Table 3 reports the parameters obtained when we estimate the models. The results for both models are quite similar. As the introduction of variables related to school ownership and peer effects does not produce major changes either to the sign and significance of the coefficients of the other explanatory variables or to the value of the associated odds ratios, we interpret both models jointly.

Concerning our main variable of interest, birth month appears to be a clearly significant factor and plays an important role in the probability of the assessed students being repeaters. The associated odds ratios are greater than one and higher as the year advances. This evidences the increased probability of repeating versus not repeating a grade. After the introduction of controls, there is still a difference between children born in January and February with respect to students born in March and April but it is no longer significant. From this point on, every student born in the following bimesters is significantly more likely to be a repeater at the age of 15 than their peers born in the first bimester. For example, children born in the third bimester (May and June) are 20 percentage points more likely to repeat a grade than children born in the first two months. More importantly, students born in the last two months of the year (November and December) have an 85% greater probability of repeating a year than their classmates born in January and February. The results obtained strongly indicate that being the older in the same cohort is a definite advantage in terms of repeating a grade. This result is consistent with the findings of Bedard and Dhuey (2006) for the United States, Calero (2006) for Spain, Crawford *et al.* (2011) for England, and Sprietsma (2010) for OECD countries, who all argue that being the oldest rather the youngest in the age cohort reduces the probability of grade retention.

Table 3. Models results

	MODEL 1			MODEL 2		
	Coefficients		Odds ratio	Coefficients		Odds ratio
Constant	-1.6479	***		-1.6059	***	
	(0.0729)			(0.0789)		
BM:Mar-Apr	0.0846		1.088	0.0776		1.081
	(0.0899)			(0.0905)		
BM:May-June	0.1790	**	1.196	0.1942	**	1.214
	(0.0899)			(0.0903)		
BM:July-Aug	0.2971	***	1.346	0.2955	***	1.344
	(0.0876)			(0.0879)		
BM:Sept-Oct	0.5023	***	1.653	0.4952	***	1.641
	(0.0862)			(0.0863)		
BM:Nov-Dec	0.6174	***	1.854	0.6201	***	1.859
	(0.0871)			(0.0875)		
ESCS	-0.6760	***	0.509	-0.5586	***	0.572
	(0.0254)			(0.0295)		
Gender: Boy	0.4880	***	1.629	0.4974	***	1.644
	(0.0504)			(0.0506)		
Immigrant Status: 2nd Gen	0.4696	**	1.599	0.4409	**	1.554
	(0.2084)			(0.2066)		
Immigrant Status: 1st Gen	1.2355	***	3.44	1.1967	***	3.309
	(0.0894)			(0.0891)		
Immigrant Status: Uncertain	0.7939	***	2.212	0.7570	***	2.132
	(0.1692)			(0.1733)		
Pre-primary: No	0.4439	***	1.559	0.4600	***	1.584
	(0.1222)			(0.1205)		
Pre-primary: ≤ 1 year	0.4207	***	1.523	0.4189	***	1.52
	(0.0890)			(0.0902)		
FamStruc: Single-parent	0.4703	***	1.6	0.5249	***	1.69
	(0.0699)			(0.0700)		
FamStruc: Mixed	1.1055	***	3.021	1.1067	***	3.024
	(0.2099)			(0.2178)		
School Type: GD School				-0.3597	***	0.698
				(0.0598)		
School Type: Private School				-0.5232	***	0.593
				(0.1219)		
Peer effects				-0.2841	***	0.753
				(0.0603)		

Note: Dependent variable: to be a repeater; SEs are presented in parentheses;
** significant at 95% level; *** significant at 99% level.
Source: Authors' own elaboration using data from PISA 2009 database (OECD, 2010).

Regarding control variables, findings are consistent with the results reported in the literature reviewed previously. Students' socioeconomic status has a negative and significant relationship with probability of grade retention, whereas boys are around 60% more likely to repeat a grade than girls. The immigrant status seems to be statistically significant and positively related to the dependent variable. However, the impact on the probability of grade retention decreases sharply for second-generation compared with first-generation immigrants. The value of the odds ratio associated with students classified as of 'uncertain origin' is positioned mid-way between the values for first- and second-generation immigrants, which could indicate that students from this category are members of the other two groups. Pre-primary school attendance is positively and significantly related to the probability of repeating a grade. Thus, children who received or one year's or less or no pre-primary education are more likely to repeat a grade than children who attended pre-primary school for more than a year. Noteworthy too is that family structure has a significant effect. Children living in mixed families are more likely to repeat a year than members of single-parent families, and both more than children living in a nuclear family. All in all, these outcomes for Spain are consistent with findings by Corman (2003, p. 417) for the United States, which he summarizes as follows: 'Boys are more likely to repeat a grade than girls. In addition, children who come from poorer households or who come from single-parent households are all at greater risk of failing in school'.

Model 2 highlights the influence of school type and peer effects. Children enrolled in private schools are 30% less likely to repeat a year than pupils enrolled in public schools. The impact is similar but lower comparing private government-dependent schools to public schools. These results suggest that such large differences in terms of grade retention probability depending on school ownership might be due not only to a different system of management but also to each school type establishing different internal requirements for deciding whether or not children should repeat a grade. Finally, the peer effect variable shows a negative and highly significant relationship with the dependent variable. This means that, even after accounting for students' socioeconomic background and the type of school that they attend, there is evidence that students enrolled in schools with a higher average socioeconomic status are less likely to repeat any grade.

1.5 CONCLUSIONS AND POLICY IMPLICATIONS

The phenomenon of grade retention is now a major problem in Spain. Grade retention rates, particularly during the last decade, have risen to over 30%. According to the OECD's PISA 2009 report, 36% of Spanish 15-year-old students had repeated at least one grade. These figures are a warning sign of school failure and early school dropout, whose percentages are consistent with grade retention rates. The consequences of this situation are negative for both individuals and the economy as a whole and even more so in the current economic crisis.

The aim of this chapter is to examine the possible influence of birth month on the likelihood of having repeated a year by the age of 15 in Spain. For this purpose, we use the data provided by the OECD's PISA 2009 report. Our first conclusion is that, at least for Spain, the birth month variable is exogenously distributed with respect to the probability of being a repeater. Regardless of socioeconomic status, Spanish parents do not plan the birth of their children at the beginning of the year based on expected educational outcomes, as evidenced by the fact that the observed distribution of births in 1993 (birth year of the assessed students) was statistically similar to the expected distribution, at a 95% confidence level.

Concerning the influence of birth month on grade retention, the first statistically significant differences appear with a bimonthly aggregation, where the retention rate of children born in last two months of the year is 14 percentage points higher than retention rate of students born in the first bimester of the same year. Note also that this is a non-linear increase, because the influence of the birth month becomes more pronounced as the year progresses, and particularly in the last semester.

In order to control by other variables that are also likely to influence the probability of repeating a year, we estimate two different logistic regression models. The findings show that birth month is statistically and significantly related to the dependent variable. The later students are born in the year, the greater is the increase in their probability of repeating versus not repeating a grade, which ranges from 19% (May-June bimester) to 85% (November-December bimester). No significant differences are found for children born between January and April after controlling for other variables. Other interesting results suggest that other factors related to the increased likelihood of repeating a grade are lower household educational level and income, male gender, immigrant status (above all first-generation immigrant students), not having received pre-primary schooling, living in a non-nuclear family, attending public school or belonging to a school with a lower average socioeconomic index.

Interestingly, maturity differences at early ages (due to birth month) are significant at the end of secondary education. This result implies that birth month has a sizeable and persistent effect on educational performance. Such findings suggest that there is a need for innovative educational strategies to solve this problem. Spain has settled for an inflexible admission rule, where children born between January 1st and December 31st of the same year have to enrol in the same grade at school. As there is no general public intervention for students born at the end of the year, there are maturity differences among children in one and the same class because of a near one-year age gap between children born at the beginning and end of the same year.

Authors like Strom (2004), Crawford *et al.* (2007) or Sprietsma (2010) advocate a more flexible rule, according to which parents should be able to choose when to enrol their children, especially if they were born at the end of the year. According to our results, an alternative policy for this issue would be to give parents of children born at the beginning and at the end of the year (i.e. the oldest and the youngest children in each cohort) the opportunity to decide whether their children should move up or down a year, respectively. The provision of additional tuition to offset the initial disadvantages of the youngest students in the class or the doubling of primary education classes with the aim of reducing age gaps from 12 to 6 months are other alternatives proposed by Gutiérrez-Domènech and Adserà (2012) and Ponzo and Scoppa (2014).

This research provides evidence that there is a clear problem that the educational authorities need to solve. The educational disadvantage incurred by the youngest pupils in their academic cohort should be viewed as a serious concern. On this ground, some public intervention is needed to ensure that individuals are not unfairly penalized by their birth month, as it is unacceptable in terms of efficiency and equal educational opportunities.

1.6 REFERENCES

Agasisti, T. and Cordero, J. (2013). Educational disparities across regions: A multilevel analysis for Italy and Spain. *Journal of Policy Modeling*, 35 (6), 1079-1102.

Allen, J. and Barnsley, R. (1993). Streams and Tiers: The interaction of ability, maturity, and training in systems with age-dependent recursive selection. *Journal of Human Resources*, 28 (3), 649-659.

Asteriou, D. and Agiomirgianakis, G.M. (2001). Human capital and economic growth: Time series evidence from Greece. *Journal of Policy Modeling*, 23 (5), 481-489.

Bedard, K. and Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *Quarterly Journal of Economics*, 121, 1437-1472.

Benito, A. (2007). La LOE ante el fracaso, la repetición y el abandono escolar. *Revista Iberoamericana de Educación*, 43 (7), 1-11.

Calero, J. (2006). *Desigualdades tras la educación obligatoria: nuevas evidencias*. Working Paper 83/2006. Madrid: Fundación Alternativas.

Calero, J., Choi, A. and Waisgrais, S. (2010). Determinantes del riesgo de fracaso escolar en España: Una aproximación a través de un análisis multinivel aplicado a PISA 2006. *Revista de Educación,* 225-256, extra number.

Corman, H. (2003). The effects of state policies, individual characteristics, family characteristics, and neighbourhood characteristics on grade repetition in the United States. *Economic of Education Review*, 22 (4), 409-420.

Crawford, C., Dearden, L. and Greaves, E. (2011). *Does When You Are Born Matter? The Impact of Month of Birth on Children's Cognitive and Non-Cognitive Skills in England*. IFS Briefing Note 122. London: Institute for Fiscal Studies.

Crawford, C., Dearden, L. and Meghir, C. (2007). *When you are born matters: The impact of date of birth on child cognitive outcomes in England*. London: Centre for the Economics of Education, London School of Economics and Political Science.

Duval, R. and de la Maisonneuve, C. (2010). Long-run growth scenarios for the world economy. *Journal of Policy Modeling*, 32(1), 64-80.

Eide, E.R. and Showalter, M.H. (2001). The effect of grade retention on educational and labor market outcomes. *Economic of Education Review*, 20 (6), 563-576.

Gutiérrez-Domènech, M. and Adserà, A. (2012). Students performance in elementary schools. *Revista de Economía Aplicada*, 59 (20), 135-164.

Hanushek, E.A. and Kimko, D.D. (2000). Schooling, labour-force quality, and the growth of nations. *American Economic Review*, 90 (5), 1184-1208.

Jimerson, S.R., Anderson, G.E. and Whipple, A.D. (2002). Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the School*, 39 (4), 441-457.

Kawaguchi, D. (2011). Actual age at school entry, educational outcomes and earnings. *Journal of Japanese and International Economics*, 25 (2), 64-80.

Manacorda, M. (2012). The cost of grade retention. *The Review of Economic and Statistics*, 95 (2), 596-606.

OECD (2005). *PISA 2003 Data Analysis Manual. SPSS users*. Paris: OECD.

OECD (2009). *PISA Data analysis manual. SPSS SECOND EDITION*. Paris: OECD.

OECD (2010). *PISA 2009 Results: What Students Know and Can Do*. Paris: OECD.

OECD (2011). Does participation in pre-primary education translate into better learning outcomes at school. In *PISA in focus*. Paris: OECD.

Ponzo, M. and Scoppa, V. (2014). The long-lasting effects of school entry age: Evidence from Italian students. *Journal of Policy Modeling*, 36 (3), 578-599.

Psacharopoulos, G. (2007). *The costs of school failure: A feasibility study*. European Expert Network on Economics of Education, Monograph.

Rutkowski, L., González, E., Joncas, M. and Von Davier, M. (2010). International Large-Scale Assessment Data: Issues in Secondary Analysis and Reporting. *Educational Researcher*, 39 (2), 142-151.

Salinas, J. and Santín, D. (2012). Selección escolar y efectos de la inmigración sobre los resultados académicos españoles en PISA 2006. *Revista de Educación*, 358, 382-405.

Sprietsma, M. (2010). Effect of relative age in the first grade of primary school on long-term scholastic results: international comparative evidence using PISA 2003. *Education Economics*, 18 (1), 1-32.

Strom, B. (2004). *Student achievement and birthday effects*. Trondheim, Norway: Department of Economics, Norweigan University of Science and Technology, mimeo.

Turner, R. (2006). El Programa Internacional para la Evaluación de los Alumnos (PISA). Una perspectiva general. *Revista de Educación*, 45-74, extra number.

1.7 APPENDIX: LITERATURE REVIEW

Author	Objective	Database	Results
Allen & Barnsley (1993)	Analyse whether streaming at early ages has a long-term effect on the educational performance of Canadian and British students	Two specific surveys for Canadian Hockey League and for Canadian and British students	There are educational differences by birth quarter, and these differences persist and even increase in the long term
Bedard & Dhuey (2006)	Analyse the effect of age at school entry on educational outcomes for 9- and 13-year-old Canadian and American students	TIMNSS 1995 TIMNSS 1999	Relative age effects disappear over time, although retain a marginal effect into adolescence
Calero (2006)	Study the determinants of the low rate of individuals with post-secondary education finished	European Union Household Panel Data for Spain (PHOGUE)	Children born in the last quarter of the year have learning disabilities, so that their likelihood of repeating a grade increases
Calero *et al.* (2010)	Detect the determinants of school failure in Spain	PISA 2006	The grade retention policy is not an effective strategy for reducing the risk of school failure
Corman (2003)	Examine the effects of state education policies, and individual, family and neighbourhood characteristics on grade retention in USA	National Household Education Survey 1991, 1993, 1995, 1996	Children born in the fourth quarter of the year are more likely (around 6 percentage points) to repeat a grade than children born in the first quarter
Crawford *et al.* (2011)	Evaluate the impact of birth month on the academic performance of 5- and 8-year-old students and what its causes are	English National Pupil Database	Older students in each school cohort have higher average results and these differences remain in the long term
Eide & Showalter (2001)	Analyse the influence of grade retention on the probability of dropping out of high school and on labour market earnings	High School and Beyond from United States	Grade retention causes higher drop-out rates and lower future earnings

Gutiérrez-Domènech & Adserà (2012)	Analyse the effect of personal and socioeconomic characteristics on the academic achievement of 2nd-, 4th- and 6th-grade Catalonian students	Família i Educació a Catalunya 2005	Children born at the end of the year have lower academic achievement, and this disadvantage does not disappear over time
Kawaguchi (2011)	Analyse the effect of birth month on educational attainment and labour market outcomes for Japanese students	TIMSS 2003 Employment Status Survey 2002	Older children in same school cohort do better than younger ones in primary school and this initial advantage persist and develops into a difference in eventual educational attainment
Manacorda (2012)	Measure the effect of grade retention on students' subsequent school outcomes	Specific database from Uruguay	Negative impact of retention on educational performance
Ponzo & Scoppa (2014)	Evaluate the effect of age at school entry on school performance of 4^{th}-, 8^{th}- and 10^{th}-grade Italian students	PIRLS 2006 TIMSS 2007 PISA 2009	Younger children score lower than their older peers and that advantage remains into adolescence
Sprietsma (2010)	Analyse the effect of relative age on the academic results of 15-year-old students through an international comparison	PISA 2003	There is a long-term (non-linear) impact on academic results and on grade retention
Strom (2004)	Estimate the effect of age at school entry on school achievement for 15- to 16-year-old students in Norway	PISA 2000	Children born in the fourth quarter of the year have lower educational results and a higher likelihood of being held back or requiring special education

CAPÍTULO 2

THE IMPACT OF IMMIGRANT CONCENTRATION IN SCHOOLS ON GRADE RETENTION IN SPAIN: A DIFFERENCE-IN-DIFFERENCES APPROACH

2.1 INTRODUCTION

There has been a remarkable increase in the foreign population in Spain over the 2000s, with a constantly growing inflow that accounts for almost one third of the total immigrants received by the OECD (Cebrián *et al.*, 2010). This was the result of the expansion of the Spanish economy, motivated largely by the construction sector boom. These immigration rates have slowed down since 2009 and even declined slightly in absolute terms between 2010 and 2012, possibly due to the economic crisis (Sánchez, 2013). Throughout this period there has been a significant change in the composition of the immigrant population according to their countries of origin. In the early days most immigrants came mainly from Latin America, whereas the percentage of the immigrant population from other European countries, mainly European Union nonmembers, increased notably towards the end of this period (Rojas and Sánchez, 2011).

A direct consequence of this phenomenon is the higher proportion of immigrant students in the Spanish education system, rising from 1.5% in 2000 to 9.5% in 2011 with a 9.81% peak in 2009. Table 1 shows immigration figures in Spain from 2000 to 2011 and the evolution of the proportion of immigrant students in the Spanish education system.

Table 1. Data about immigrant population in Spain

Year	Immigrant Population	% Total Population	% Immigrant Students in the Education System
2000	923,879	2.3	1.5
2001	1,370,657	3.3	2.0
2002	1,977,946	4.7	2.9
2003	2,664,168	6.2	4.4
2004	3,034,326	7.0	5.7
2005	3,730,610	5.5	6.5
2006	4,144,166	9.3	7.4
2007	4,519,554	10.0	8.4
2008	5,220,600	11.3	9.4
2009	5,598,691	12.0	9.8
2010	5,747,734	12.2	9.7
2011	5,730,067	12.2	9.5

Source: Authors' own elaborations using data from the municipal register (National Institute of Statistics).

In most countries, immigrant students have lower educational outcomes, higher dropout rates and lower levels of noncompulsory education than native students (Driesen, 2000; Schnepf, 2008; Murat and Frederic, 2015). Studies focusing on average differences in educational outcomes between immigrant and native students from traditionally immigrant-receiving countries like Germany provide evidence that immigrant students are not able to definitively close the educational gap between themselves and their native classmates (Frick and Wagner, 2001; Ammermueller, 2007). In some other countries like Belgium and Canada, however, where native students continue to outperform their immigrant peers, the performance gap has narrowed despite the rising the percentage of immigrants (Entorf and Minoiu, 2005; OECD, 2011). Nevertheless, there are nonconclusive evidence about the impact of high rates of immigrant pupils on the achievement of natives. On the one hand, Brunello (2013), using cross-country data from nineteen different countries, and Contini (2013), analysing Italian education system, highlight that high proportions of immigrant students affect negatively the natives' learning outcomes, although the size of this effect is relatively small and weak. On the other hand, Ohinata and Van Ours (2013) and Geay *et al.* (2013) using data from primary education in Netherlands and England, respectively, do not find any negative impact of concentration of immigrant students on the performance of native ones. The latter, even detects a positive effect for Maths achievement of native pupils from England, but only when the impact of the increase of immigrant students is analysed within Catholic schools. The authors suppose that this positive impact is due to the fact that these immigrant students are from Eastern European families whose parents are highly educated. In Spain, recent articles have studied this phenomenon using different approaches: Calero and Waisgrais (2009) and Calero *et al.* (2009) compare the educational performance of immigrant students and their peers using multilevel regression techniques, concluding that the determinants of educational achievement affect native and immigrant students differently. Zinovyeva *et al.* (2014) perform Oaxaca-Blinder decomposition in order to analyse the educational gap between natives and immigrants and find that around half of this gap can be attributed to socioeconomic and family factors. Finally, Salinas and Santín (2012) employ a switching regression model to calculate the impact of immigration on the educational outcomes controlling for school type. They show that immigrant students have a higher probability of attending public schools and that the negative effect on native students produced by the concentration of immigrants is bigger in public schools than in private government-dependent schools.

Another relevant issue in Spain is the high rates of grade retention (around 30% of students), which is a warning sign of school failure and a good predictor of school dropouts. Several studies support the hypothesis that repeating a grade is often the main predictor of school failure (Roderick, 1994; Jimerson *et al.*, 2002; Benito, 2007). This has led us to study the effect of immigration from another perspective. We consider whether or not the increase in immigrant students recent years has had repercussions on grade retention rates particularly for native students. A similar approach is followed by Cristia *et al.* (2014) in order to test whether increased technology access in schools affect retention rates.

The present chapter uses an impact evaluation approach to study how the increase of the proportion of immigrant students in some schools can affect grade retention rates. For this purpose, we estimate the impact of the exogenous increase of immigrant students[1] in Spain from 2003 to 2009 using a Difference-in-Differences approach (DiD)[2]. Using this technique, we can determine whether the concentration of immigrants has a significant effect on student performance by comparing the percentages of students studying in the proper grade by age. A similar approach could be broadly applied to several developed countries, since nowadays the integration of immigrants into society becomes a main concern around the world. The same design could be also used to the so-called new immigration countries, those which transformed from immigrant-sending countries to immigrant-receiving ones. These countries such as Portugal, Italy or Greece, have increased considerably their foreign population in the past decades. In the same way, the traditionally immigrant-receiving countries, such as Germany, France or the UK, could test whether, even being in an advanced stage of the immigration process, this situation keeps having some impact on the educational achievement of the destination country.

The research reported here makes two contributions. Firs, we apply the DiD method to analyse the possible relationship between the increase of foreign students and grade retention rates. The idea behind this approach is that the treatment types could differ in some situations, depending in this case on the concentration of immigrants. On this ground, the treatment will be referred to as a *dose* treatment. Second, instead of applying this methodology to longitudinal data as is common practice in the previous literature, we use data from consecutive cross sections OECD's PISA reports for Spain. This approach of

[1] *Native students* are students born in the country of assessment or who have at least one parent who was born in that country. *Immigrant students* are students who are foreign-born and whose parents are also foreign-born or students who were born in the country of assessment but whose parents were not (OECD, 2010); i.e. without differentiating between first-generation immigrants and second-generation immigrants. We include them in the same general category (*Immigrant students*) due to the fact that most immigrant students evaluated in PISA 2003 were second-generation immigrants and the sample of first-generation immigrants would not be representative.

[2] We are aware that the residential choices made by immigrants, as well as their school district choices are non-random. However, it is a fact that immigrant population tends to set in lower socioeconomic areas and these areas do not change between the two periods analysed, as neither do the schools.

using data from different waves of PISA is similar to the one followed by Zinovyeva *et al.* (2014) who analyse the gap between immigrant student's achievement with respect to native students in Spain through Oaxaca-Blinder decompositions.

The chapter is structured as follows. Section 2.2 presents and justifies the applied methodology. In Section 2.3, we describe the dataset used and the selected variables included in the empirical analysis. Section 2.4 reports the results. We conclude in section 2.5 by discussing the implications of our findings for public policy.

2.2 METHODOLOGY

The goal of our research is to analyse the impact of the growth in immigrant students experienced by Spain over the last ten years on the average grade retention rates per school. According to the theory of impact evaluation, having foreign students enrolled at the school would be the *treatment,* and schools with immigrants would be the *treated* schools. Note, however, that this is a *dose treatment*, so we are not simply looking for the average effect of there being or not being foreign students at the school, but the effect of their concentration on the treated schools. Therefore, we have two groups. One group is composed of schools hosting the immigrants, considered as the *treated group.* These schools will have also received different treatments because the concentration of immigrants varies over time. The other group includes schools not hosting immigrants, known as the *nontreated* or *control group*.

The rate of nonrepeater students (who are in the correct grade) from 2003 to 2009 at the control schools will vary due to a number of possibly unknown factors. The variation of this rate at the treatment schools will be due to the same factors plus the variation in the component we are trying to evaluate, i.e. the arrival of immigrants. In order to estimate the impact of the exogenous increase in the number of immigrants, we use the DiD technique by means of which we can isolate the effect of immigrant arrival from the unknown factors. Although this technique requires panel data, it can also be estimated using cross-sectional databases, provided that they can be guaranteed to be consistently representative (Khandker *et al.*, 2010) and the samples are selected according to the same procedure throughout (Meyer, 1995). In this case, consecutive PISA reports (OECD, 2004; 2010) satisfy these requirements.

The DiD method calculates the average difference in outcomes separately for treatment and nontreatment groups over the period. Then, after taking an additional difference between the average changes in outcomes for these two groups, it is possible to identify the difference-in-differences impact, i.e. the estimated impact of the assessed issue. For our empirical educational model, let Y_t^T and Y_t^C denote the mean percentages of students in the proper grade for their age at treated and control schools, respectively, and *t* a dummy variable that can take two values: 2003 and 2009. The classical DiD technique estimates the average impact as follows:

$$DD = E(Y_{2009}^T - Y_{2003}^T) - E(Y_{2009}^C - Y_{2003}^C) \quad (1)$$

Note that if the treatment group differs from the control group in terms of observed and unobserved characteristics in addition to treatment, we need to assume that the differences between the two groups are time-invariant in order to obtain an unbiased difference-in-differences estimator. The DiD estimator can be solved using a regression. On the basis of the discussion in Ravallion (2008), the estimating equation would be as follows:

$$Y_t^D = \alpha + \beta Dt + \rho D + \gamma t + \varepsilon \quad (2)$$

where *D* is the treatment variable, *t* is the time dummy variable and the coefficient of the interaction of *D* and *t, β* represents the estimated impact of the treatment on outcome *Y*:

$$D = \begin{cases} 1\ if\ it\ belongs\ to\ the\ treatment\ group \\ 0\ if\ it\ belongs\ to\ the\ control\ group \end{cases}$$

$$t = \begin{cases} 1\ if\ year = 2009 \\ 0\ if\ year = 2003 \end{cases}$$

The coefficient of the interaction *β* indicates whether or not the increase in immigrant students has a significant impact on the dependent variable and how much impact it has. In addition to the interaction term, the variables time (*t*) and treatment (*D*) are also included in order to detect any isolated effects due to the time or to group membership.

As mentioned at the beginning of this section, we are not only interested in measuring the average effect of immigrant students on educational performance, but also the impact of their concentration. For this reason, we include what we call a *dose treatment* in our research, and these *doses* are the percentages of immigrants at each school belonging to the treated group, represented by the variable *Immig*[3]. Although dose treatments usually consider finite numbers of treatment levels (i.e. a discrete variable such as different cash transfer sums), this approach can also be applied to continuous treatments (Abadie, 2005), as is in this case. The explanatory variable *Immig* is added to a saturated model combined with time, treatment and the interaction of both variables. However, the saturated model cannot be estimated because of its perfect multicollinearity.

Since we are only interested in the term that contains the treatment dose ($\delta_2 ImmigDt$), the equation we finally estimate is as follows:

$$Y_t^D = \alpha + \beta Dt + \rho D + \gamma t + \delta ImmigDt + \varepsilon \tag{3}$$

The DiD estimator is now the result of adding two terms: the interaction coefficient β and the effect that contains the percentage of immigrants $\delta Immig$.

According to the specification, first we have the treatment (having immigrant students at school) which leads to the average impact of the increase of immigrant students between 2003 and 2009; and second, we include the dosage (percentage of immigrant students) which corrects the average effect since it allows that the concentration of immigrant has a different impact over the average one[4].

We can summarize our strategy as follows. In the first period, we have two groups: schools with and without immigrants. Across the two periods, we assume that immigrant students join the education system and enrol in the schools. This is equivalent to increasing the *dose* of immigrants in the education system, and we are interested in analysing the impact of this increase on grade retention. At the end of this period, we again have schools with no immigrant population (the control group) and schools with a higher mean percentage of immigrants (the treated group), although this mean is not uniformly distributed across schools. This implies that the dose received by each treated school is different[5].

---

[3] This idea is closely related to the approach developed by Abadie and Dermisi (2008).

[4] The same estimates can be done taking out the main treatment factor, letting alone the treatment intensity factor without significant changes in the results.

[5] It may be possible that some schools in the control group in the first period were classified in the treatment group during the second period, but the opposite is unlikely. This fact guarantees that control group samples are similar in terms of composition in both periods.

It is noteworthy that a basic assumption behind this technique is that the remaining covariates (*X*), which could affect both the treated and the control groups, must be unchanged over time. If this is not a valid assumption, the regression analysis should control those covariates in order to ensure a correct estimation as follows:

$$Y_t^D = \alpha + \beta Dt + \rho D + \gamma t + \delta ImmigDt + \eta X + \varepsilon \tag{4}$$

In this case, the regressions include five control variables. They are described in the following section. Furthermore, the trends of the treatment group and the control group are assumed to be equal in the absence of treatment, although this assumption cannot be tested. However, we performed a *placebo* test in order to check the validity of the DiD method. This test involves performing an additional DiD estimation using a *fake* treatment group (i.e. comparing two control groups) or a *fake* outcome (Gertler *et al.*, 2011). Because of the type of database, we chose the second option, using the *average percentage of girls per school* as our fake dependent variable uncorrelated with the treatment, as is also performed by Felfe *et al.* (2015).

Finally, the results section includes a simulation analysis of how the average promotion rates per school vary depending on the percentage of immigrant students enrolled in order to clarify our estimations.

2.3 DATA AND VARIABLES

2.3.1 The PISA Report

The data set used for the research comes from the PISA (*Programme for International Student Assessment*) survey, designed by the OECD in 1990s as a comparative, international, regular and continuous study on certain educational characteristics and skills of students worldwide (Turner, 2006). The PISA target population is composed of students who are aged between 15 and 16 years old at the time of the assessment, all of whom are born in the same year and who have completed at least six years of formal schooling. PISA measures their performance in math, reading and science. It also collects information about students' personal background and schools environment, for which purpose two questionnaires are administered, one addressed to school principals and another to students[6]. These surveys have taken place every three years since the year 2000 focusing on one of the above three areas each time.

An important aspect that is to be taken into account in an empirical analysis using PISA data is that the data are gathered by means of a two-stage sampling procedure. First, a sample of schools is selected in every country from the full list of schools containing the total student population. Then, a sample of 35 students is randomly selected within each school. As a result, statistical analyses have to consider sampling weights in order to ensure that sampled students adequately represent the analysed total population (Rutkowski *et al.*, 2010)[7].

2.3.2 Sample, variables and the identification strategy

Although the DiD method usually uses panel data, repeated cross-sectional data from the same areas has also been used in the literature (Chaudhury and Parajuli, 2010; Felfe *et al.*, 2015[8]). We use data from two different waves, 2003 and 2009, which provide information useful for interpreting average results concerning the 2002/03 and 2008/09 academic years. The chosen unit of analysis is the school and, therefore, the data is aggregated at school level. PISA samples are composed of different school types that can be divided into three groups according to their ownership: public (government managed and funded schools), private (privately managed and funded schools) and private government dependent (privately managed and government funded schools). In our research, we focus on schools that are comparable in terms of public funding and also share the same admission criteria[9], i.e. public and private government-dependent schools. The sample is composed of 336 schools (199

[6] Parents complete a third questionnaire. However, this information is only available for a limited number of countries and, unfortunately, Spain is not one of them.

[7] These weights include adjustments for nonresponse by some schools and students within schools and weight cutting to prevent a small set of schools or students having undue influences. These processes are based on intensive calculation methods, known as resampling methods, which consist of taking multiple samples from the original sample. Specifically, PISA uses the Balanced Repeated Replication (BRR) with 80 replicates. For an extensive description of this procedure, see OECD (2005; 2009a).

[8] This study specifically applies the same methodological approach to the same database that us (different waves of PISA) to test the impact of a substantially public childcare expansion in Spain.

[9] Public-funded schools cannot reject immigrant students that ask for a position in the school. This fact prevent the model of potential bias results occurring if we had that schools with no immigrants were the result of selection. On the other hand, note that immigrant students attending private schools are a minority that can afford an expensive education, and they are not supposed to generate any educational problem.

public schools and 137 private government-dependent schools) in 2003 and 806 schools (512 public schools and 294 private government-dependent schools) in 2009[10].

Regarding the variables, we use the *percentage of students who are in their correct grade* (without repeating any year) and the *percentage of native students who are in their correct grade* as dependent variables[11]. Since PISA assesses 15-year-old students, we consider that 4th-grade ESO students (the so-called *Enseñanza Secundaria Obligatoria*, i.e. compulsory secondary education in the Spanish system, equivalent to 10th grade on the international scale) are in their correct year[12]. We differentiate between these two dependent variables in order to distinguish how the concentration of immigrant students in schools affects grade retention and native grade retention, in particular.

In our analysis, the treated schools are schools that have immigrant students. As the distribution of immigrant students is not uniform across the education system, the concentration of these students differs from one school to another. As we described in the methodology section, the aim of introducing this issue in our econometric models, we consider a *dose treatment*. In this way, we include the *percentage of immigrants* (*Immig*) in the base model (2), defined as the ratio between immigrant students and the total number of students sampled by school in order to capture the potential effects of a higher presence of immigrants in schools (3).

The school distribution by control and treated groups, and the different treatment doses are shown in Tables 2 and 3, respectively.

---

[10] The difference in sample size between the two periods is due to the fact that PISA 2009 covered more regions with an extended sample than PISA 2003 (14 regions in 2009 and 3 regions in 2003). However, both samples can be used to obtain general conclusions for Spain due to the fact that both PISA 2003 and PISA 2003 are nationally representative.

[11] Grade retention is chosen as dependent variable instead of student test scores due to the fact that the values of test scores in PISA are rescaled each wave (OECD average equals 500 and SD 100). Therefore, it is impossible to make comparisons of one country performance over time because normalization avoids concluding if the output is really increasing or decreasing over time. For more detail see PISA 2006 Technical Report (OECD, 2009b).

[12] We can do this assumption due to the fact that in Spain there are legal constraints preventing parents from choosing their children’s enrolment cohort. Parents cannot postpone their children's entry to the first year of primary school with the aim of their children being more mature and performing better at school.

Table 2. School distribution by groups

	2003		2009	
	Schools	**%**	**Schools**	**%**
Control Schools	154	45.8	168	20.8
Treated Schools	182	54.2	638	79.2
N	336	100	806	100

Source: Authors' own elaborations using data from PISA (OECD, 2004; 2010).

Table 3. Different treatment doses within treated schools

	2003		2009	
Treated Schools: *Immig Dose*	**Schools**	**%**	**Schools**	**%**
< 5%	81	44.50	136	21.32
5% - 10%	54	29.67	161	25.24
10% - 15%	27	14.83	119	18.65
15% - 20%	7	3.85	79	12.38
20% - 25%	7	3.85	49	7.68
> 25%	6	3.30	94	14.73
Total	182	100	638	100

Source: Authors' own elaborations using data from PISA (OECD, 2004; 2010).

From Table 2 we conclude that the percentage of schools with immigrants grew significantly from 2003 (54.17% of total) to 2009 (79.17% of total). Additionally, Table 3 shows that around 11% of schools had an immigrant student population of more than 15% in 2003, whereas this percentage multiplied by more than three in 2009 reaching 34.79%.

Moreover, as we explained above, we select a set of control variables to be introduced in the model (names in brackets denote variable names in the results tables).

Concerning parental background, we included the *Index of parental occupational status* (Parental Occupation) that represents the index of highest occupational status of parents according to the *International Socio-Economic Index of Occupational Status* (ISEI, Ganzeboom *et al.*, 1992). We built a variable that represents the average value of this index for each school. We assume that the higher the average parental occupational status, the greater their income, whereby students enrolled at this school will have higher average socioeconomic status, and the *Parental educational level* (Parental Education), an index of highest educational level of parents in years of education according to the *International*

Standard Classification of Education (ISCED, OECD, 1999). Again, we construct a variable that represents the average value of this index for each school.

As regards school characteristics, we selected the *Type of School* (School Type) as a *dummy* variable that takes value 1 if the school is a private government-dependent school and 0 for a public school and the *Quality of school resources* (School Resources) which is a continuous variable based on the school principal's responses to seven questions available from PISA 2003 and PISA 2009 databases related to the availability of computers for educational purposes, educational software, calculators, books, audiovisual resources and laboratory equipment.

In order to control for the school location, we introduced four dummy variables related to the town population: *Village*, *Small town*, *City* (taken as the baseline category) and *Large City*. Each dummy variable takes value 1 is the school is located in a town with an amount of population within the bounds specified (in Table 5, it can be checked the different bounds for every dummy variable). Tables 4 and 5 report the main descriptive statistics for the variables considered in our analysis and the distribution of control and treatment schools within the different population sizes.

Table 4. Descriptive statistics

	2003				**2009**			
	Control		**Treated**		**Control**		**Treated**	
Schools	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Dependent variables								
% Students in the correct year	0.7342	0.1666	0.7073	0.1847	0.7310	0.19482	0.6508	0.1764
% Native students in the correct year	0.7342	0.1666	0.6634	0.1852	0.7310	0.19482	0.5995	0.1869
Independent variables								
% Immigrant students (Immig)	0.0000	0.0000	0.0835	0.1044	0.0000	0.0000	0.1394	0.1273
Parental Occupation	43.2738	8.4051	43.7276	7.7022	47.4312	10.1658	44.6906	7.6220
Parental Education	11.4560	1.6249	11.2753	1.6100	12.6508	2.0495	12.1989	1.6387
School Type	0.4400	0.4980	0.3800	0.4870	0.5200	0.5010	0.3200	0.4680
School Resources	-0.0393	0.9982	-0.0932	1.0074	0.0332	0.7855	-0.0156	0.8472

Source: Authors' own elaborations using data from PISA (OECD, 2004; 2010).

Table 5. Distribution of schools within the different population sizes

	2003		2009	
Regions	**Control**	**Treated**	**Control**	**Treated**
Village (Pop.<15,000)	50	54	71	202
Small Town (Pop. 15,000-100,000)	52	53	42	217
City (Pop. 100,000-1,000,000)	49	66	53	198
Large City (Pop. > 1,000,000)	3	9	2	21

Source: Authors' own elaborations using data from PISA (OECD, 2004; 2010).

It is well-known in the literature the existence of two factors that must be taken into account when the educational achievement of immigrant students is being analysed: their country of origin and whether they are familiar with the language spoken in the country of destination. However, these variables could not be included in the current empirical analysis as control variables due to the lack of information in the PISA 2003 and PISA 2009 surveys, at least in the Spanish case.

According to Table 4, regarding the dependent variables, it can be seen that both the *percentage of students in the correct year* and the *percentage of native students in the correct year* in the control group are quite similar comparing the year 2003 to the year 2009. On the other hand, the rates in both dependent variables have experienced a decrease between the two periods of time assessed in the treated group. This decline is higher in case of the second dependent variable, what suggests us that the percentage of immigrant students in the correct year (variable not presented in Table 4, but can be obtained by subtracting the second dependent variable from the former one) has slightly increased. This result might be due to higher amount of immigrant students in the year 2009 compared to year 2003 and the improvement of Spanish education system in terms of integration.

In relation to the independent variables, our *dosage* (percentage of immigrant students enrolled in the school) has considerably risen between the first and the second year. Comparing the evolution of the indexes of *parental occupation status* and *parental education level* in both groups between the year 2003 and year 2009 it can be said that these variables increased over the two years both in control group and the treatment group, suggesting that the average levels of parental occupation and parental education have improved throughout this period of time in Spain coinciding with the economic boom. It can be claimed the same for the school resources index. Focusing on the school type variable, it can be noted that there is a rise of private government-dependent schools without immigrant students during the considered time (in relative terms). In our opinion it is because the combination of two different factors. First, the number of private government-dependent schools who provide compulsory secondary education increased in greater extent than the number of public schools between the two years. Secondly, it can suggest the existence of some degree of segregation of immigrant students and this intuition can be, somehow, right due to the fact that private government-dependent schools often locate in larger cities and neighbourhoods and on average, in higher socioeconomic status areas within the cities, where some immigrant families can be found living around but not the most of them.

2.4 ANALYSIS OF RESULTS

2.4.1 Results

This section presents the results for the models described in the methodology. Specifically, we estimate three different models for each dependent variable: *percentage of students in their correct grade* (Students) and *percentage of native students in their correct grade* (NStudents). Model 1 is the basic difference-in-differences model estimation (2). Model 2 is equivalent to the basic model plus the treatment dose (3) captured through the percentage of immigrants at the school combined with the interaction term ($\delta Immig$). Finally, Model 3 estimates equation 4 as an extension of Model 2, in which control variables are also introduced in order to single out the net effect of treatment. By including these variables, we can test whether or not they have a separate effect on the outcome.

Table 6 reports the model estimation parameters, showing variable coefficients, standard errors and statistical significance in each column. At this point, all effects will be quantified on the average percentage of students who are in the correct grade for their age and, therefore, have not repeated any year.

First, regarding estimates of the *percentage of students in their correct grade* (Model 1) shows that, taken separately, neither the time variable nor group membership has a significant effect on the dependent variable. With respect to the coefficient associated with the interaction term *(β),* i.e. the difference-in-differences estimator, we observe no significant difference between treated (schools with immigrants enrolled) and control group (schools without immigrants enrolled) throughout the evaluated period. The information provided by the interaction term is the average effect of an increase of immigrants. Thus, given that PISA evaluated schools have few immigrants on average, it is reasonable to assume that, on average, promotion rates at schools with an average number (few) of foreign students do not decrease significantly compared to 2003 with respect to control schools. This result appears to suggest that schools with low mean values have adapted well to this new situation (slight increase of immigrant student enrolment). The addition of the dose treatment in Model 2 discloses similar results related to the above variables. However, the coefficient associated with the interaction term by the percentage of immigrants ***(δ)***, i.e. the difference-in-differences *dose* estimator turns out to be statistically significant and is negatively related to the dependent variable. This implies that the concentration of immigrant students has a negative impact on grade retention for all students (immigrant and native students) with respect to the control group.

Table 6. Difference-in-differences estimations for all students

All Students	**Model 1**		**Model 2**		**Model 3**	
Constant	0.6924	***	0.6924	***	0.0961	
	(0.0280)		(0.0281)		(0.0679)	
Year (t)	-0.0579		-0.0579		-0.0992	**
	(0.0462)		(0.0462)		(0.0388)	
Treatment (T)	0.0003		0.0003		-0.0067	
	(0.0342)		(0.0342)		(0.0262)	
Interaction	0.0026		0.0767		0.0645	
	(0.0510)		(0.0519)		(0.0397)	
Immig (interact)			-0.5499	***	-0.3235	***
			(0.0705)		(0.0658)	
Parental Occupation					0.0067	***
					(0.0016)	
Parental Education					0.0180	**
					(0.0077)	
School Type					0.0806	***
					(0.0208)	
School Resources					0.0134	
					(0.0088)	
Village					0.0226	
					(0.0196)	
Small Town					0.0176	
					(0.0193)	
Large City					-0.0267	
					(0.0249)	

Notes: SEs are presented in parentheses.
Source: Authors' own elaborations using data from PISA (OECD, 2004; 2010).

Model 3 parameters illustrated in Table 6 can be interpreted similarly. The only notable difference is that the effect of immigrant concentration persists and is significant, albeit to a lower extent, despite control based on the variables related to school type, school resources, school location and school average socioeconomic status, through indexes that represent the level of parental education and parental occupation. With respect to the control variables introduced in the model, variables representing the educational level and occupational status of parents and the type of school are statistically significant.

Table 7 illustrates the three model estimation parameters for the percentage of native students in their correct grade only.

Table 7. Difference-in-differences estimations for native students

Native Students	**Model 1**		**Model 2**		**Model 3**	
Constant	0.6924	***	0.6924	***	0.1115	*
	(0.0280)		(0.0281)		(0.0675)	
Year (t)	-0.0579		-0.0579		-0.0996	**
	(0.0462)		(0.0462)		(0.0388)	
Treatment (T)	-0.0394		-0.0394		-0.0460	*
	(0.0339)		(0.0339)		(0.0260)	
Interaction	-0.0102		0.1105	**	0.0987	**
	(0.0509)		(0.0514)		(0.0391)	
Immig (interact)			-0.8959	***	-0.6749	***
			(0.0567)		(0.0532)	
Parental Occupation					0.0061	***
					(0.0017)	
Parental Education					0.0192	**
					(0.0078)	
School Type					0.0789	***
					(0.0204)	
School Resources					0.0138	
					(0.0087)	
Village					0.0224	
					(0.0193)	
Small Town					0.0134	
					(0.0188)	
Large City					0.1115	*
					(0.0675)	

Notes: SEs are presented in parentheses.
Source: Author's calculations using data from PISA (OECD, 2004; 2010).

According to Table 7, the estimation of the *percentage of native students in their correct grade* (dependent variable) shows only one relevant difference with respect to the previous model. In this case, the last two models report a statistically significant interaction coefficient *(β)* with a positive correlation with the dependent variable. Hence, it can be argued that, when the percentage of immigrants enrolled is introduced (treatment dose), native students benefit on average from having a small number of immigrant students in the classroom. We believe that this effect may be due to the fact that immigrants are more susceptible to suffer grade retention. It is worth to note that this result could also be due to, for instance, the improvement of non-cognitive skills of native students because of sharing classroom with their immigrant peers. However, it would be a topic of discussion for a different paper, and it is beyond the scope of the current one[13]. Nevertheless, this slight advantage is offset and, finally, even cancelled out by the *dose* coefficient.

2.4.2 Simulation

To clarify the above results, Table 8 is a simulation of how the average promotion rates vary in schools based on the percentage of enrolled immigrant pupils[14]. Any percentage of enrolled immigrants has negative effects on the percentage of nonrepeaters for all students, although these effects are significant when the proportion of immigrants in the classroom is above 10%. For example, schools with a 10% concentration of immigrant students have around three immigrant pupils per classroom (for a 30-student classroom), which results in a decrease of from one to two nonrepeater pupils. In the case of native students, however, concentrations of immigrant students of under 15% have neither negative nor positive effects. Teachers appear to substitute potential native repeaters by these immigrant students when there are not many immigrant students in the class (fewer than four to five students) and the percentage of nonrepeating native students decreases.

However, when immigrant concentrations climb to over 15% (more than five immigrants per class), we start to detect a significant negative impact on natives' results compared with natives in the control group. According to the summary statistics presented in Table 3, this negative effect of immigrant concentration will impact on 34% of schools (those with immigrant concentration above 15% in 2009). In this case, the presence of six immigrant students per classroom (equivalent to an immigrant concentration of around 20%) leads to a reduction of from two to three individuals in the rate of nonrepeating native students. This finding, which is similar to previous findings reported in the literature (Calero and Waisgrais, 2009), provides empirical evidence demonstrating that there is a clear negative peer effect related to a high concentration of immigrant students in some schools.

[13] See Heckman *et al.* (2006), Lleras (2008) or Levin (2012) for a discussion on non-cognitive education measures and the necessity of taking into account that dimension when educational achievement or labour market outcomes are being analysed.
[14] Simulations are based on the estimations from Models 2 (equation 3) and 3 (equation 4) contained in Tables 6 and 7. It makes no sense to run a simulation based on Model 1 because this model does not include the percentage of immigrants.

Table 8. Simulation of results for different percentages of immigrant students

% Immig	All Students		Native Students	
	MODEL 2	MODEL 3	MODEL 2	MODEL 3
1	-0.01	-0.01	0.10	0.09
5	-0.03	-0.02	0.07	0.06
10	-0.06	-0.03	0.02	0.03
15	-0.08	-0.05	-0.02	-0.00
20	-0.11	-0.06	-0.07	-0.04
25	-0.14	-0.08	-0.11	-0.07
30	-0.17	-0.10	-0.16	-0.10
35	-0.19	-0.11	-0.20	-0.14
40	-0.22	-0.13	-0.25	-0.17
45	-0.25	-0.15	-0.29	-0.21
50	-0.28	-0.16	-0.34	-0.24

Source: Authors' own elaborations.

2.4.3 Placebo test

As mentioned in the methodology section, one assumption of the DiD method is that the trends of the treatment and control groups would be equal in the absence of the treatment, i.e. both groups are similar in all variables but the treatment. Because we cannot prove this assumption, we perform a *placebo* tests in order to check whether the identified effects are due to such treatment and endorse the correct selection of the control and treatment groups (Gertler *et al.*, 2011).

In our research, we apply the *placebo* test using a fake dependent variable -*average percentage of girls at school*-, knowing that it should not be affected by the increase of immigrant students in classrooms, but at the same time it seems to be correlated to grade retention, as girls are less likely to repeat a grade than boys (Corman, 2003). Table 9 summarizes the results which corroborate our hypothesis: the DiD estimator (coefficient associated with the interaction term) and the DiD *dose* estimator (coefficient associated with immigrant concentration) are not statistically significant in any of the models.

Table 9. Placebo test: Difference in differences models using *percentage of girls at school* as a fake dependent variable

Percentage of Girls	Model 1		Model 2		Model 3	
Constant	48.3916	***	48.3916	***	54.4187	***
	(1.5588)		(1.5596)		(4.7314)	
Year (t)	1.2782		1.2782		1.9073	
	(1.6757)		(1.6765)		(1.6171)	
Treatment (T)	0.6165		0.6165		0.3852	
	(1.9767)		(1.9776)		(1.8682)	
Interaction	-0.1767		-0.6931		-0.4445	
	(2.1176)		(2.1648)		(2.0038)	
Immig (interact)			3.8517		-0.5403	
			(3.1009)		(4.2394)	
Parental Occupation					0.0483	
					(0.1259)	
Parental Education					-0.5341	
					(0.3351)	
School Type					-3.1862	***
					(1.0709)	
School Resources					0.6730	
					(0.4642)	
Village					-1.3859	
					(1.5927)	
Small Town					-1.6750	
					(1.2487)	
Large City					1.5648	
					(4.7314)	

Notes: SEs are presented in parentheses.
Source: Authors' own elaborations using data from PISA (OECD, 2004; 2010).

2.5 CONCLUSIONS

Since the late 1990s, there has been a constantly growing inflow of immigrants, leading to a remarkable increase in the foreign population in Spain. This has affected the percentage of immigrant students who have joined the Spanish education system and account for around 9.5% of the school population for the year 2011 up from 1.5% in 2000. At the same time, Spain is feeling the effect of other relevant issues like consistently very high grade retention rates of around 30%.

Given this background, the aim of this chapter is to estimate the impact of the exogenous increase of immigrant students from 2003 to 2009 using a DiD approach, which would reveal whether immigrant concentration had a significant effect on the percentage of nonrepeater students. We use the data provided by consecutive OECD's PISA reports.

In our identification strategy, schools with foreign students enrolled constitute our *treatment* group, whereas schools composed of only native students define our *control* group. On top of the traditional mean effect estimations, however, we analyse the impact of the concentration of immigrants in classrooms in this article. For this reason, we refer to a *dose treatment* (Abadie and Dermisi, 2008), where the dose is the percentage of immigrant students and hence, the DiD estimator is the sum of the terms related to interaction and the percentage of immigrants ($DD = \beta + \delta Immig$).

Since we are interested in evaluating the effect of the immigration phenomenon on students and native students, in particular, we have two dependent variables: *percentage of students who are in their correct grade* and *percentage of native students who are in their correct grade.* For each dependent variable, we estimate three models: the basic DiD model (Model 1), an equivalent model introducing the treatment dose (Model 2) and an extension of the previous models that includes a set of control covariates (Model 3). Moreover, we develop a *placebo* test to check the validity and the robustness of the approach.

Analysing the effect on all students, we find that the interaction coefficient *(β)* (DiD basic impact estimator) appears not to be statistically significant; however, the term associated with the dose of immigrants *(δ)* (percentage of immigrant students) has a negative and statistically significant relationship with the percentage of students who are in their correct grade. The impact on native students is different, as the interaction coefficient *(β)* in the DiD dose estimator is statistically significant and positive, but this small advantage is offset and finally cancelled out by the *dose* term *(δ)* when the concentration of immigrants is above 15%.

In conclusion, immigrant students joining the Spanish education system does not, on average, decrease school promotion rates with respect to 2003. This situation is even beneficial to native students because foreign students are more greatly affected by grade retention. Taking into account the *dose* (percentage of immigrants enrolled per school), however, we find that the concentration of immigrant students has a negative impact on promotion rates. In other words, the average percentage of repeaters, and, in particular, the average percentage of native repeaters, has increased in 2009 with respect to 2003 as a consequence of higher immigrant concentrations in some schools. However, native students are only affected by higher concentrations of immigrant students (above 15%).

The key question is why the addition of immigrant students had such an impact on the education system. A potential reason for this result is that immigrant students have a language deficit and lower educational level when they join the Spanish education system. Therefore, when the number of immigrant students per classroom grows, the average educational level of the students in these classrooms drops and more students fail to reach the educational level for promotion. Some possible educational strategies to manage this situation would be to regulate the maximum percentage of immigrants per school in order to avoid high concentrations. Nevertheless, once high concentrations of immigrant pupils is a fact in some schools, policy-makers should contemplate specific strategies in order to avoid the negative effects of large concentration of immigrant students and at the same time, fostering the improvement of immigrant students' educational attainment. Policies such as the provision of more resources for specific language and skills training in those schools with high concentration of immigrants enrolled could solve problems of adaptation to the new education system. Those resources could be employed for hiring specialized teachers who focus on immigrant students and their progress or for reducing size of classrooms so the concentration of immigrant pupils would be lower. Moreover, it is widely known in the literature of economics of education that both family background and home environment are key variables to children' learning process; hence, policies aimed to improve the integration of those immigrant families would be extremely useful. Strategies driven to reduce their labour market insertion problems or special instructions for acquisition of the new language would impact positively on immigrant socioeconomic status and as a consequence, on their offspring's educational performance.

2.6 REFERENCES

Abadie, A. (2005). Semiparametric Difference-in-Differences estimators. *The Review of Economics Studies*, 72 (1), 1-19.

Abadie, A. and Dermisi, S. (2008). Is terrorism eroding agglomeration economies in Central Business Districts? Lessons from the office real estate market in downtown Chicago. *Journal of Urban Economics*, 64 (2), 451-463.

Ammermuller, A. (2007). Poor background or low returns? Why immigrant students in Germany perform so poorly in the Programme for International Student Assessment. *Education Economics*, 15 (2), 215-230.

Benito, A. (2007). La LOE ante el fracaso, la repetición y el abandono escolar. *Revista Iberoamericana de Educación*, 43 (7), 1-11.

Brunello, G. and Rocco, L. (2013). The effect of immigration on the school performance of natives: cross country evidence using PISA test scores. *Economic of Education Review*, 32, 234-246.

Calero, J., Choi, A. and Waisgrais, S. (2009). Determinantes del rendimiento educativo del alumnado de origen nacional e inmigrante. *Cuadernos Económicos del ICE*, 78, 281-311.

Calero, J. and Waisgrais, S. (2009). Rendimientos educativos de los alumnos inmigrantes: identificación de la incidencia de la condición de inmigrante y de los peer-effects In *Proceedings of the XVI Meeting of Public Economic*, Granada (Spain).

Cebrián, J.A., Bodega, M.I., Martín-Lou, M.A. and Guajardo, F. (2010). La crisis económica internacional y sus repercusiones en España y en su población inmigrante. *Estudios Geográficos*, 71 (268), 67-101.

Chaudhury, N. and Parajuli, D. (2010). Conditional cash transfers and female schooling: the impact of the female school stipend programme on public school enrolments in Punjab, Pakistan. *Applied Economics*, 42 (28), 3565-3583.

Contini, D. (2013). Immigrant background peer effects in Italian schools. *Social Science Research*, 42 (4), 1122-1142.

Corman, H. (2003). The effects of state policies, individual characteristics, family characteristics, and neighbourhood characteristics on grade repetition in the United States. *Economic of Education Review*, 22 (4), 409-420.

Cristia, J., Czerwonko, A. and Garofalo, P. (2014). Does technology in schools affects repetition, dropout and enrollment? Evidence from Peru. *Journal of Applied Economics*, 17 (1), 89-112.

Driesen, G.W.J.M. (2000). Ethnicity, forms of capital and educational achievement. *International Review of Education*, 47 (6), 513-538.

Entorf, H. and Minoiu, N. (2005). What a difference immigration policy makes: A comparison of PISA scores in Europe and traditional countries of immigration. *German Economic Review*, 6 (3), 355-376.

Felfe, C., Nollenberger, N. and Rodríguez-Planas, N. (2015). Can't buy mommy's love? Universal childcare and children's long-term cognitive development. *Journal of Population Economics*, 28 (2), 393-422.

Frick, J.R. and Wagner, G.G. (2001). Economic and social perspectives of immigrant children in Germany. Discussion Paper 301, IZA.

Ganzeboom, H.B.G., De Graaf P.P.M., Treiman, D.J. and De Leeuw, J. (1992). A standard international socio-economic index of occupational status. *Social Sciences Research*, 21 (1), 1-56.

Geay, C., McNally, S. and Telhaj, S. (2013). Non-native Speakers of English in the Classroom: What Are the Effects on Pupil Performance? *The Economic Journal*, 123 (570), 281-307.

Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B. and Vermeersch, C.M.J (2011). *Impact evaluation in practice.* Washington, DC: World Bank.

Heckman, J.J., Stixrud, J. and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behaviour. Working paper 12006, National Bureau of Economic Research.

Jimerson, S.R., Anderson, G.E. and Whipple, A.D. (2002). Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the Schools*, 39 (4), 441-457.

Khandker, S.R., Koolwal, G.B. and Samad, H.A. (2010). *Handbook on impact evaluation: quantitative methods and practices.* Washington, DC: World Bank.

Levin, H.M. (2012). More than just test scores. *Prospects*, 42 (3), 269-284.

Lleras, C. (2008). Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings. *Social Science Research,* 37 (3), 888-902.

Meyer, B.D. (1995). Natural and Quasi-Experiments in Economics. *Journal of Business & Economic Statistics*, 13 (2), 151-161.

Murat, M. and Frederic, P. (2015). Institutions, culture and background: the school performance of immigrant students. *Education Economics,* 23 (3), 612-630.

OECD (1999). *Classifying educational programmes: Manual for ISCED-97 implementation in OECD countries*. Paris: OECD.

OECD (2004). *Learning for tomorrow's world. First results from PISA 2003*. Paris: OECD.

OECD (2005). *PISA 2003 Data analysis manual. SPSS users*. Paris: OECD.

OECD (2009a). *PISA 2006 Technical Report*. Paris: OECD.

OECD (2009b). *PISA Data analysis manual. SPSS SECOND EDITION*. Paris: OECD.

OECD (2010). *PISA 2009 Results: What students know and can do*. Paris: OECD.

OECD (2011). *How are school systems adapting to increasing numbers of immigrant students? PISA IN FOCUS*. Paris: OECD.

Ohinata, A. and Van Ours, J.C. (2013). How immigrant children affect the academic achievement of native Dutch children. *The Economic Journal*, 123 (570), 308-331.

Ravallion, M. (2008). Evaluating anti-poverty programs. In *Handbook of Development Economics,* (Eds) T. P. Schultz and J. Strauss. Amsterdam, North-Holland.

Roderick, M. (1994). Grade retention and school dropout: Investigating the association. *American Educational Research Journal*, 31 (4), 729-759.

Rojas, J. and Sánchez, P. (2011). Evolución reciente y proyecciones de la población en España. *Boletín Económico del Banco de España*, 6, 61-68.

Rutkowski, L., González, E., Joncas, M. and Von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, 39 (2), 142-151.

Salinas, J. and Santín, D. (2012). Selección escolar y efectos de la inmigración sobre los resultados académicos españoles en PISA 2006. *Revista de Educación*, 358, 382-405.

Sánchez, P. (2013). Evolución reciente de la población en España y proyecciones a corto y largo plazo. *Boletín Económico del Banco de España*, 1, 71-80.

Schnepf, S.V. (2008). Inequality of learning amongst immigrant children in industrialized countries. Discussion Paper 3337, IZA.

Turner, R. (2006). El Programa Internacional para la Evaluación de los Alumnos (PISA). Una perspectiva general. *Revista de Educación*, 45-74, extra number.

Zinovyeva, N., Felgueroso, F. and Vázquez, P. (2014). Immigration and student achievement in Spain: evidence from PISA. *SERIEs*, 5 (1), 25-60.

CAPÍTULO 3

ASSESSMENT OF *PROGRAMA 18-25* USING A FUZZY RDD: HOW DO CONDITONAL CASH TRANSFERS IMPACT ON EDUCATIONAL ACHIEVEMENT?

3.1 INTRODUCTION

The Europe 2020 strategy has set, within their targets related to employment, that labour force must be qualified according to labour market demands. To fulfil with this aim, it is essential to encourage lifelong learning, and raise the overall quality of education at all levels. The *early school leaving*[1] figures for Spain reveal that, despite its decreasing over last years, early school leaving rates have decreased from 30% in 2010 to about 22% in 2014 (MECD, 2013; 2016), such figures are still far from the 10% goal set by Europe 2020 strategy, though. The economic theory notes that failing to complete secondary education comes at a great cost to both the individual, due to the major labour insertion problems, and a higher risk of social exclusion (Oreopoulos, 2007; Oreopoulos and Salvanes, 2011), and the economy as a whole, because of the lower economic growth prospects (Barro, 2001; Psacharopoulos, 2007). Given the Spanish economic situation, where unemployment rates are over 20% and youth unemployment rates are close to 45%, reducing school dropout rates and fostering a proper qualification of the labour force, objectives that should be a priority in the design of the national and regional economic policy.

Governments of most developed countries have implemented measures to address the problems found in the labour market in order to reduce the mismatch between the labour supply and demand. These measures can be grouped into the so-called *Active Labour Market Programmes* (ALMP). Such action consists, at international level, of several programmes such as job search assistance, public employments, wage subsidies policies to employers, or educational and training for unemployed people. The reports of Dar and Tzannatos (1999) and Betcherman *et al.* (2004) collect the systematic evaluations of ALMPs carried out in different countries in recent years[2]. Malo and Cueto (2016) show a similar review for the Spanish case.

In the region of Extremadura and according to the Employment Observatory, around 41% of people registered as unemployed in the regional Public Employment Service (*Servicio Extremeño Público de Empleo*, hereafter SEXPE) had not completed the secondary education in 2012, figures that have remained stable in recent years. In this context, the

[1] The percentage of people aged 18-24 who have only completed lower secondary education or less and are no longer in education or training.

[2] In addition, aiming at completing this information, a literature review has been carried out from the date of the aforementioned reports to the present. A summary table of the different assessments found can be seen in Appendix I. This table includes contributions related to the evaluations of ALMPs, training programmes for people with labour insertion problems and adult education.

Government of Extremadura signed an agreement with SEXPE with the aim of developing specific programmes targeted to unemployed people. Its goals were to provide education and training for people with labour insertion problems, particularly young people aged between 16 and 25 years, long-term unemployed women, disabled people, or people at risk of social exclusion.

The policy evaluated in this study, the so-called *Programa 18-25*, belongs to the training programmes mentioned above. The *Programa 18-25* was launched by the Government of Extremadura in November 2012 aiming at reducing the number of unemployed and uneducated people. The target population was unemployed people aged between 18 and 25 years who had not completed the compulsory secondary education (the so-called *Educación Secundaria Obligatoria* (ESO) in the Spanish education system). This programme was meant, through an economic incentive of 1,000 euros, to motivate these people to get back into education system and obtain the diploma from upper secondary school in the shortest possible period of time.

In Spain, the only programme with similar features is the scholarship called *2ª Oportunidad* from Andalusia region, which entered into force in the 2011/12 academic year and remains in force. This scholarship is targeted at 18-24 year old people who enrol in compulsory secondary education, vocational training programmes or Spanish baccalaureate. It consists on a monthly cash transfer (around 75% of Spanish Public Income Index (IPREM[3])) conditioned to attend school regularly, do homework and pass the different exams, being the maximum amount received per academic year 4,000 euros.

After *Programa 18-25* was implemented, two similar education policies were launched. First, the Government of Castile-La Mancha introduced an identical programme to the Extremadura one in the 2013/14 academic year, which was in effect for the two following academic years. Second, in January 2015, the programme called *Gradua2* was launched by the Government of Castile-León targeted to people over 18 years old without the compulsory secondary education diploma. It consisted on a 6-months free training course in order to pass the secondary education certificate exam. The beneficiaries could enjoy a scholarship for transport, accommodation and/or meals depending on the average family income. In addition, those students who obtained the Diploma and enrolled in vocational

[3] *Indicador Público de Rentas de Efectos Múltiples (IPREM*): index designed as a wage indicator or reference aimed at assisting in the determination of amounts of certain scholarships, grants or unemployment benefits. It was created in 2004 to be employed in substitution of the minimum wage.

training or Spanish baccalaureate in next academic year received an economic incentive of 500 euros.

In this chapter, the following two objectives are considered. The first aim is to check whether some ‘pull effect’ took place caused by the *Programa 18-25* introduction; that is, whether the number of people enrolled in secondary education increased within potential programme's beneficiaries (individuals aged between 18 and 25 years) with respect to those who were not. On the other hand, we also try to evaluate whether enjoying this programme increased the probability of obtaining the upper secondary education diploma or, at least, the likelihood of passing the total amount of modules enrolled.

To do this, we have employed administrative data provided by Regional Department of Education and Culture in Extremadura and data obtained through surveying the students enrolled in secondary education for adults during the 2013/14 academic year. Due to the fact that *Programa 18-25* distinguishes between males and females and it is not possible to evaluate the impact over the total population, in this chapter we look at males only. We consider that the present study is relevant due to a two-fold reason. To our knowledge, this is the first time this kind of programme[4] is evaluated in Spain (or at regional level). Secondly, given that other regions' government are currently running or considering this type of programme, we believe that the results obtained in this study are likely to be of interest to them.

The chapter is structured as follows. In the following section, the *Programa 18-25* features are described in depth. Section 3.3 presents and justifies the research design and the applied methodology. In Section 3.4, the dataset used and the selected variables included in the empirical analysis are explained. Finally, the chapter winds up with the report of the results and the main conclusions.

[4] The *Programa 18-25* could be classified as an Active Labour Market Programme (ALMP) due to the fact that it is targeted to improve unemployed population's job opportunities. Even so, this programme is also an educational policy, belonging to adult education, which aims to increase the average qualification of the labour force in Extremadura.

3.2 DESCRIPTION OF *PROGRAMA 18-25*

As we mentioned previously, the *Programa 18-25* was part of an agreement signed by SEXPE and the Regional Department of Education and Culture, with the aim of developing specific training programmes targeted to unemployed people. This agreement invoked Act 4/2011, March 7th, of Education of Extremadura [5], which states that the Regional Government has the power to promote policies targeted to ensure the right to lifelong learning of those who left the education system, encouraging them to get back to study in order to improve their personal and professional life. The *Programa 18-25* was adopted in November 2012 for the 2012/13 academic year[6]. From then on, the programme was running for the two following academic years, 2013/14 and 2014/15[7]. When the political party in charge of the regional government changed, in May 2015, after the regional polls the programme was cancelled. In this study, we assess *Programa 18-25* during the 2013/14 academic year.

The aim of this programme was to reduce the amount of unemployed people who did not have the compulsory secondary education in Extremadura, in the spirit of improving their job opportunities, through an economic incentive. The potential beneficiaries were those unemployed aged between 18 and 25 years people without the upper secondary education diploma. Moreover, long-term unemployed[8] women could also benefit without age limit. The programme management, both regarding education and administrative concerns, conformed to the provisions of secondary education for adults (*Educación Secundaria para Personas Adultas*, ESPA) of Extremadura region[9]. Thus, when the student applied for admission in adult education should express his/her intention to benefit from *Programa 18-25*, enrolling in remaining modules up to six per academic year.

The economic incentive was conditioned to academic achievement, so that the programme beneficiaries must attend school regularly, pass regular exams of each module enrolled (as every student enrolled in secondary education for adults) and, moreover, pass

[5] DOE (*Extremadura Official Journal*) No. 47, March 9th 2011. The Extremadura Official Journal (hereafter DOE) is a daily written publication used by the regional government of Extremadura to publish public or legal notices, such as decrees, acts, agreements, etc.
[6] DOE No. 228, November 26th 2012.
[7] DOE No. 173, September 6th 2013; DOE No. 173, September 9th 2014.
8 It is considered long-term unemployed people those who have been unemployed for 12 months or more.
[9] For further information, see *Order August 1, 2008, which regulates Compulsory Secondary Education for Adults in the Autonomous Community of Extremadura* (DOE No. 158, August 18th 2008).

specific tests for those modules. Only those students who met the requirements mentioned above would receive the cash transfer, which consisted on two independent payments of 500 euros, at the end of each semester, upon request of beneficiaries.

3.3 IDENTIFICATION STRATEGY

3.3.1 Research design

In order to assess the impact of *Programa 18-25*, the degree of academic success of those who benefit from the programme along the 2013/14 academic year is compared to those who enrolled in secondary education for adults (in the same academic year) but they didn't benefit from the programme (due to the fact that they did not meet the age or employment status requirements).

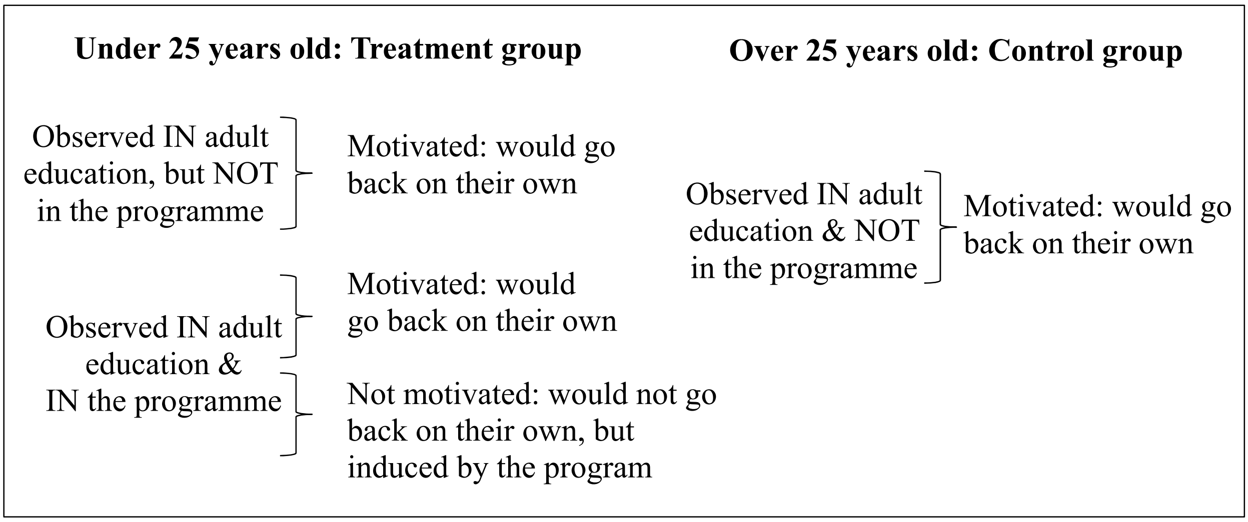
For this purpose, we followed the next steps. First, the principals in all high schools where adult education was taught were informed about the evaluation project asking for their cooperation. Second, we sent out a survey[10] to every student enrolled in secondary education for adults throughout the first weeks of the 2013/14 academic year. At the same time, the office in charge of adult education in the government of Extremadura provided us with essential administrative information for running the assessment.

A previous analysis of the students' evaluated behaviour revealed essential issues to determine the methodology to apply, how to define the treated and control groups and the measures of interest to estimate the programme impact. Accordingly, it is essential to identify the typology of the students that take part in the programme. First, there were some individuals who, choosing to enrol in the programme, obtained the upper secondary education diploma but however, they did not received the economic incentive because they did not attend the specific tests either they did not request the payment at the end of the semester. Second, we detect that some individuals got the payment of 500 euros per semester (which entailed meeting the requirements of attendance and passing regular exams and specific tests) and, even so, they did not get the upper secondary education diploma. This was due to the fact that they still had to pass some remaining modules; i.e. the remaining

[10] The survey can be found in Appendix II.

modules in order to get their certificate were more than six[11]. Finally, as *Programa 18-25* was not a mandatory public policy, some potential beneficiaries, despite they enrolled in secondary education for adults, they did not take up the policy. As one requirement of the programme was to enrol in all the remaining modules up to six to complete the upper secondary education, these potential beneficiaries chose to enrol in fewer modules and therefore did not access to benefit from the programme.

Figure 1. Treatment and Control groups



Source: Authors' own elaboration.

Figure 1 shows how the treatment and control group were composed. The treatment group includes the males enrolled in secondary education for adults who are 25 years or younger[12], at December 31st 2013. At the same time, this group is composed of those who chose not to enrol in the programme and those who did. In the latter, it is possible to distinguish between individuals who are intrinsically motivated, i.e. they would enrol in secondary education for adults regardless of the programme; and individuals who are extrinsically motivated, those who get back into education induced by *Programa 18-25*. Concerning the control group, this consists on those who enrolled in adult education but could not benefit from the programme due to the fact that they were older than 25 years (intrinsically motivated). In order to carry out a proper assessment of the programme, we should compare treated and non-treated individuals who are intrinsically motivated. However, it is not possible to distinguish between these two kind of motivated individuals

[11] This is the maximum number of modules enrolled allowed per academic year, both in secondary education for adults and in *Programa 18-25.*

[12] The minimum age in this group is 18 years, since this is the minimum age required by law to enrol in adult education.

within the treatment group. Although we guess that the *Programa 18-25* did not induce males who were 25 years old or younger to enrol in adult education to a greater extent we analyse whether or not some 'pull effect' resulting from the policy took place.

3.3.2 Methodology

First, for the purpose of checking whether there were some 'pull effect' caused by the introduction of *Programa 18-25*, we plot the number of students enrolled in secondary education for adults by age range. This allows us to visually detect whether there were a boost in enrolment due to the treatment; i.e. whether the number of students enrolled in secondary education for adults who were programme's beneficiaries increased considerably with respect to those who were not.

In addition and for the same purpose, we estimate the following regression:

$$Y_{is} = \alpha + \lambda Treatment + \beta Years_{is} + \beta Modules_{is} + \varepsilon_{is} \quad (1)$$

where Y_{it} corresponds to the frequency of individuals who are i years old in school s; $Treatment$ is a binary program indicator that equals 1 if i belongs to the 20-25 interval and equals 0 otherwise; the variable $Years$ denotes ages i and $Modules$ indicates the average number of modules enrolled by school s and ages i. Finally, the term ε_{is} represents the error term. This equation (1) is estimated by ordinary least squares with school fixed effects. The coefficient λ of this regression estimates the average bump of individuals whose enrolment in secondary education for adults was induced by *Programa 18-25*.

Second, we examine the impact of *Programa 18-25* on the variables of interest. Due to the programme nature, individuals are ranked based on a running variable (*Age*), being able to benefit from the policy only those who do not exceed a certain threshold (25 years), we decide to use the regression discontinuity design (RDD). This approach allows us to compare individuals in a narrow band above and below the cutoff point with similar characteristics, making easier to identify the causal effect of the programme. RDD was introduced in the evaluation literature by Thistlewaite and Campbell (1960) when tried to study the effect of a scholarship only granted to those students who obtained an specific test scores above a threshold[13]. From then on, this method has been applied in order to evaluate education issues as diverse as the effect of class size on students' performance (Angrist and Lavy, 1999), the impact of university financial aid awards on college enrollment (Van del Klaauw, 2002), the influence of grade retention on educational attainment (Jacob and Lefgren, 2004) or the impact of the month of birth on cognitive and non-cognitive skills (Crawford *et al.,* 2014).

The RDD is adequate for programs or policies that have a continuous eligibility index, X_i, with a strictly defined cut-off point, $\bar{x}$, to determine who is eligible and who is not. Then, if D_i denotes the treatment so:

$$D_i \begin{cases} 1\ if\ X_i \leq \bar{x} \rightarrow Treated \\ 0\ if\ X_i > \bar{x} \rightarrow Non-treated \end{cases}$$

The main advantage of regression discontinuity approach is that comparing the results obtained by units in a close neighbourhood (above and below) around the eligibility cut-off, our comparison will be as good as if we had a randomized trial, and therefore, differences in outcomes can be entirely attributed to the intervention itself (Gertler *et al*., 2011).

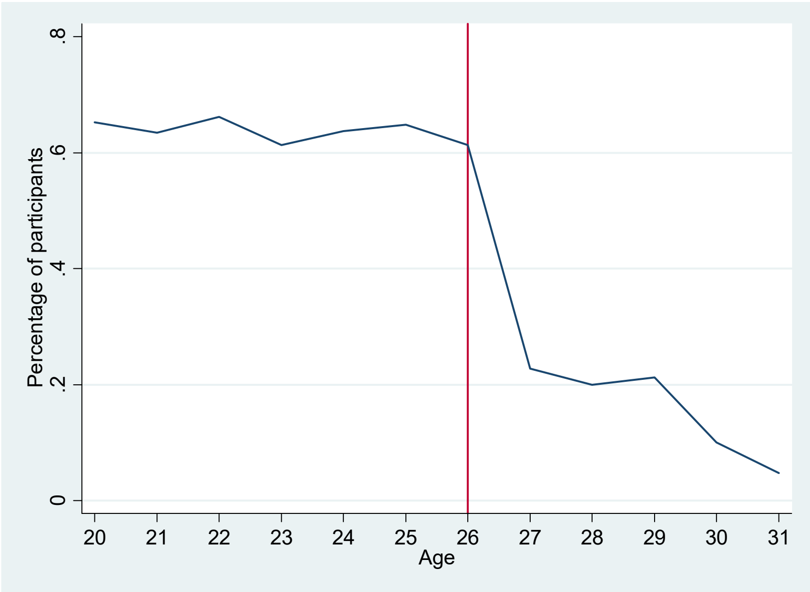
There are two main general settings within the RDD. The *sharp regression discontinuity design* is applied when the running variable defines precisely treatment and control group. In the *fuzzy regression discontinuity design*, instead, the running variable does not perfectly determine treatment group but creates a discontinuity in the probability of receiving the treatment (Schlotter *et al.,* 2011). Fuzzy RDD happens when the eligibility rules are not strictly adhered due to the fact that some unobserved variables rule the assignment to treatment (Hahn *et al.,* 2001).

Our case is the latter, where some eligible individuals did not self-select into *Programa 18-25* (known in the literature as *never-takers*), while some individuals found a way to enrol in the programme even when they did not meet the requirement of age (*always-takers*). As it can be seen in Figure 2, the percentage of participants in *Programa 18-25* is less than one to the left of the cut-off point and greater than zero to the right of the cut-off

[13] The history of the RDD within impact evaluation theory can be found in Cook (2008). Moreover, see Van der Klaauw (2008) for a review of researches in the field of economics where this method is applied.

point. It means that the running variable (*Age*) did not perfectly match the treatment, and therefore, a fuzzy RDD arises here.

Figure 2. Percentage of participants in *Program 18-25* by age



Source: Authors' own elaboration using data from the administrative sample.

Fuzzy RDD can be analysed in an Instrumental Variables framework, defining a simple dummy variable, denoted by I_i, for whether the running variable X_i is below or above the cut-off point and using it as an instrument for treatment variable D_i in the estimation of the outcome equation (Angrist and Pischke, 2008). Therefore, fuzzy RDD is estimated through the following equations:

First stage or treatment equation: $$D_i = \gamma_0 + \gamma_1 I_i + \gamma_2 X_i + \varepsilon_1 \quad (2)$$

Second stage or outcome equation: $$Y_i = \beta_0 + \beta_1 \widehat{D}_i + \beta_2 X_i + \varepsilon_2 \quad (3)$$

where D_i indicates if the individual belongs to the treated group or the control group; I_i is the instrument created from the running variable, which is represented by X_i; $\widehat{D}_i$ is the estimated treatment variable in the first stage and denotes the probability of receiving the treatment. Finally, Y_i corresponds to the measures of interest chosen to evaluate the programme impact.

There are several concerns to take into account when RDD is applied. First, the running variable should not be manipulated in order to ensure assignment to treatment. In the present study, the running variable (the age) meet this requirement (Imbens and Wooldridge, 2009). Second, the specification may be sensitive to the functional form used in modelling the relationship between the assignment variable and the outcome variable (Gertler *et al,* 2011). On the other hand, RDD produces local average treatment effects that cannot necessarily be generalized to units far away from the cut-off point (Kandher *et al.,* 2010). Finally, it is not always possible to find enough observations available close enough to the threshold. In order to solve the problem related to the limited sample size, the interval around the cut-off point can be increased, but as we move further from the eligibility threshold, eligible and ineligible units will become more different, what can bias the comparison (Schlotter *et al.,* 2011). Including covariates may eliminate some bias resulting of the higher bandwidths (Imbens and Lumieux, 2008). In this research, the bandwidth was extended due to the limited sample size available, thus, a set of control variables were included in the empirical model addressing to avoid bias.

The equations we finally estimate are as follows:

First stage or treatment equation:

$$D_i = \gamma_0 + \gamma_1 CutOffAge_i + \gamma_2 Age_i + \gamma_3 Z_i + \varepsilon_1 \tag{4}$$

Second stage or outcome equation:

$$Y_i = \beta_0 + \beta_1 \widehat{D}_i + \beta_2 Age_i + \beta_3 Z_i + \varepsilon_2 \tag{5}$$

where Y_i corresponds the measures of interest of the programme impact; D_i indicates the real treatment; $\widehat{D}_i$ denotes the estimated treatment, estimated using the instrument that is represented by the *CutOffAge* variable[14]; Age_i is the running variable and the set of covariates is denoted by Z_i.

3.4 DATA AND VARIABLES

3.4.1 Data

The information we are using comes from two sources. On the one hand, we have an *Administrative sample* created from administrative records, provided by the regional government, of all students enrolled in secondary education for adults in Extremadura in the

[14] The definition of the *CutOffAge* variable is explained in next section.

2013/14 academic year. This sample is composed of 5,485 observations, males and females, aged between 18 and 60 years. This data includes information about the students' gender, their age at December 31st 2013, if they benefited from *Programa 18-25*, the number of modules enrolled, the school where they enrolled and their results per semester.

On the other hand, we have a second sample collected through about 1,555 surveys sent-out to the schools where secondary education for adults in Extremadura were provided during the academic year evaluated[15]. After linking both samples, the administrative information and the surveys, we gather 1,232 observations (males and females). This final sample (from now on *Survey sample*) includes variables about students' socioeconomic status, in addition to the information mentioned above.

Prior to the empirical analysis, both samples are restricted following the next criteria. First, as it was mentioned before, the evaluation is focused on males. Second, due to the wide disparity in the *Age* variable, both samples are reduced to those individuals aged between 20 and 31 years old; that is, 6 years above and below the cut-off point set by the programme (25 years). Moreover, the number of modules enrolled by students could be from one to six. Most of participants enrolled four or more modules and it is clear that the effort that entailed passing the modules is obviously higher as long as more modules are enrolled. Consequently, aiming to reduce the heterogeneity among the data caused by these concerns, we focus the study on individuals enrolled in 4 modules or more. So the final samples are composed of 1,412 observations in the *Administrative sample* and 299 ones in the *Survey sample*.

Following the literature recommendations, we look for a wide enough bandwidth in order to have a proper sample size for the empirical analysis and, at the same time, narrow enough in a manner in which the individuals included are alike. To this end, as robustness tests, the analysis is replicated using different bandwidth sizes. The widest bandwidth is composed of males aged between 20 and 31 years (6 years to the left and to the right of the threshold). Its main problem is that, due to its extent, the sample could include males who are unlikely to be comparable (in terms of motivational, personal, social and economic characteristics). Narrowing the bandwidth down to 4 years above and below the cut-off point, we obtained a second group composed of males aged 22-29 years. Finally, the

[15] We obtained replies from 59 out of 62 schools where secondary education for adults is provided in Extremadura (one school from province of Cáceres and two ones from province of Badajoz did not replied).

narrowest bandwidth consists on 24-27 year-old males (2 years around the eligibility cut-off). Additionally, for each bandwidth mentioned, three groups depending on the number of modules enrolled are distinguished. Table 1 presents the different sample sizes according to the age bandwidth and the modules enrolled, both for *Administrative sample* and for *Survey sample*.

Table 1. Sample size by age bandwidth and modules enrolled

	Administrative Sample			
Age Bandwidth	**Modules**	**From 4 to 6 mod.**	**From 5 to 6 mod.**	**6 mod.**
	24 - 27 years	433	363	336
	22 - 29 years	907	770	686
	20 - 31 years	1,412	1,197	1,049
	Survey Sample			
Age Bandwidth	**Modules**	**From 4 to 6 mod.**	**From 5 to 6 mod.**	**6 mod.**
	24 - 27 years	108	94	89
	22 - 29 years	206	183	171
	20 - 31 years	299	267	245

Source: Authors' own elaboration.

3.4.2 Variables

The impact of *Programa 18-25* is measured providing two different outcomes. The first dependent variable, *Diploma*, is a binary variable that takes value 1 if the student obtained the diploma for upper secondary education at the end of the academic year and 0 otherwise. The second dependent variable, *Success*, is a binary variable that takes value 1 if the student passed every module enrolled in that academic year and 0 if he/she did not. With this dependent variable we try to capture those individuals who got the payment per semester but they did not obtain the diploma, due to the fact that they still had to pass some remaining modules in the following academic years.

The running variable settled by *Programa 18-25* is the student's age at December 31st 2013 (*Age*). The real treatment is represented by the variable called *ProgrammeSexpe*, a dummy variable that equals 1 if the student belongs to the treated group and 0 if he/she does not. The classification of individuals into the treated or control group should be equal regardless of which of the previous variables is employed; however, as it was mentioned in Section 3.3.1, we detected the presence of *always-takers* and *never-takers*, motivating the building of an instrument, *CutOffAge*, that takes value 1 if the individual is 25 years old or younger and 0 otherwise.

An initial statistical analysis showed that, even in the narrowest bandwidth, the treated and control group averages differ (statistically and significantly) in some variables. For this reason, a set of covariates were included in the empirical model in order to avoid selection bias. Concerning students' socioeconomic status, we selected, on the one hand, the *Employment Status*, a dummy variable that equals 1 if the individual is not currently working and 0 otherwise. On the other hand, the average family income is introduced through three *dummy* variables: monthly average family income above 1,500 euros (*Income>1,500*), family income between 1,000 and 1,500 euros per month (*Income 1,000-1,500*) and the baseline category, when the student claims that the monthly income is less than 1,000 euros (*Income<1,000*). As regards student environment, two different variables are included. *Rural*, defined as 1 if the student lived in a rural environment and 0 otherwise; and *Region*, binary variable that takes value 1 if the student lived in the province of Badajoz and 0 if he lived in the province of Cáceres. Finally, the number of modules enrolled per student in the academic year evaluated is introduced as a discrete variable (*Modules enrolled*).

Tables 2 and 3 report the main descriptive statistics, by bandwidth and number of modules enrolled, for the *Administrative sample* and the *Survey sample*, respectively. Focusing on *Administrative sample*, students who place above and below the cut-off point differ in terms of covariates like number of modules enrolled, region in which they live and by construction, age. Regarding *Survey sample*, treated and non-treated students display also statistically significant differences in variables related to employment status and average family incomes, disappearing the latter difference when the bandwidth is narrowed.

Table 2. Descriptive statistics from Administrative sample

Number of modules	Variables	Age Bandwidth								
		20-31 years			22-29 years			24-27 years		
		Control	Treated	Difference	Control	Treated	Difference	Control	Treated	Difference
From 4 to 6	Age	25.0707	22.8642	2.2065 ***	25.4266	24.0278	1.3988 ***	25.6790	25.0258	0.6532 ***
	Rural	0.3397	0.3101	0.2956	0.3207	0.3210	-0.0003	0.2840	0.3137	-0.0297
	Region	0.6069	0.7584	-0.1515 ***	0.6332	0.7477	-0.1145 ***	0.6049	0.7823	-0.1773 ***
	Modules Enrolled	5.5483	5.6202	-0.0719 *	5.5353	5.6531	-0.1177 **	5.5062	5.6790	-0.1728 **
From 5 to 6	Age	25.2386	22.8490	2.3896 ***	25.5000	24.0000	1.5000 ***	25.7154	25.0258	0.6896 ***
	Rural	0.3361	0.3119	0.0242	0.3191	0.3283	-0.0092	0.2615	0.3090	-0.0475
	Region	0.5913	0.7580	-0.1668 ***	0.6217	0.7554	-0.1337 ***	0.6077	0.7983	-0.1906 ***
	Modules Enrolled	5.8631	5.8853	-0.0222	5.8586	5.9120	-0.0535 **	5.8769	5.9528	-0.0759 **
6	Age	25.3510	22.9605	2.3905 ***	25.5671	24.0518	1.5153 ***	25.6754	25.0405	0.6349 ***
	Rural	0.3389	0.3144	0.0246	0.3218	0.3271	-0.0052	0.2456	0.3108	-0.0652
	Region	0.5889	0.7615	-0.1725 ***	0.6245	0.7553	-0.1308 ***	0.6053	0.7973	-0.1920 ***

Note: Difference in means significant at: *** 99%, ** 95%, * 90%.
Source: Authors' own calculations.

Table 3. Descriptive statistics from Survey sample

Number of modules	Variables	Age Bandwidth								
		20-31 years			22-29 years			24-27 years		
		Control	Treated	Difference	Control	Treated	Difference	Control	Treated	Difference
From 4 to 6	Age	25.1188	22.5808	2.5380 ***	25.4800	23.6260	1.8540 ***	25.8444	24.6191	1.2254 ***
	Employment status	0.8119	0.9949	-0.1831 ***	0.8000	0.9924	-0.1924 ***	0.8444	0.9841	-0.1397 **
	Income <1,000	0.6436	0.7475	-0.1039 *	0.7067	0.7786	-0.0720	0.7111	0.7778	-0.0667
	Income 1,000-1,500	0.2574	0.1313	0.1261 **	0.2000	0.1145	0.0855	0.2000	0.0952	0.1048
	Income >1,500	0.0594	0.0455	0.0140	0.0533	0.0382	0.0152	0.0444	0.0476	-0.0032
	Rural	0.3267	0.3990	-0.0726	0.3200	0.4122	-0.0922	0.3111	0.4762	-0.1651 *
	Region	0.3663	0.7525	-0.3862 ***	0.3600	0.7481	-0.3881 ***	0.3778	0.7460	-0.3683 ***
	Modules Enrolled	5.6337	5.7525	-0.1189	5.6000	5.7863	-0.1863 *	5.5778	5.7778	-0.2000
From 5 to 6	Age	25.2046	22.5754	2.6291 ***	25.6563	23.6219	2.0344 ***	26.0526	24.6429	1.4098 ***
	Employment status	0.8295	0.9944	-0.1649 ***	0.8125	0.9916	-0.1791 ***	0.8947	0.9821	-0.0874
	Income <1,000	0.6364	0.7374	-0.1011 *	0.7031	0.7563	-0.0532	0.6842	0.7500	-0.0658
	Income 1,000-1,500	0.2614	0.1397	0.1217 **	0.2031	0.1261	0.0771	0.2105	0.1071	0.1034
	Income >1,500	0.0568	0.0447	0.0121	0.0469	0.0420	0.0049	0.0526	0.0536	-0.0009
	Rural	0.2955	0.3966	-0.1012	0.2969	0.4118	-0.1149	0.2632	0.4643	-0.2011 **
	Region	0.3750	0.7598	-0.3848 ***	0.3594	0.7479	-0.8852 ***	0.3684	0.7321	-0.3637 ***
	Modules Enrolled	5.8750	5.9385	-0.0635	5.8750	5.9664	-0.0914 *	5.8684	6.0000	-0.1316 **
6	Age	25.3507	22.6548	2.6959 ***	25.6607	23.6609	1.9998 ***	26.0000	24.6429	1.3571 ***
	Employment status	0.8312	0.9940	-0.1629 ***	0.8036	0.4348	-0.1877 ***	0.9091	0.9821	-0.0731
	Income <1,000	0.6234	0.7381	-0.1147 *	0.6786	0.9913	-0.0693	0.6667	0.7500	-0.0833
	Income 1,000-1,500	0.2727	0.1488	0.1239 **	0.2143	0.7478	0.0839	0.2121	0.1071	0.1050
	Income >1,500	0.0649	0.0417	0.0233	0.0536	0.1304	0.0101	0.0606	0.0536	0.0070
	Rural	0.3117	0.3988	-0.0871	0.3036	0.0435	-0.1138	0.2727	0.4643	-0.1916 *
	Region	0.3896	0.7679	-0.3782 ***	0.3750	0.7565	-0.3815 ***	0.3636	0.7321	-0.3685 ***

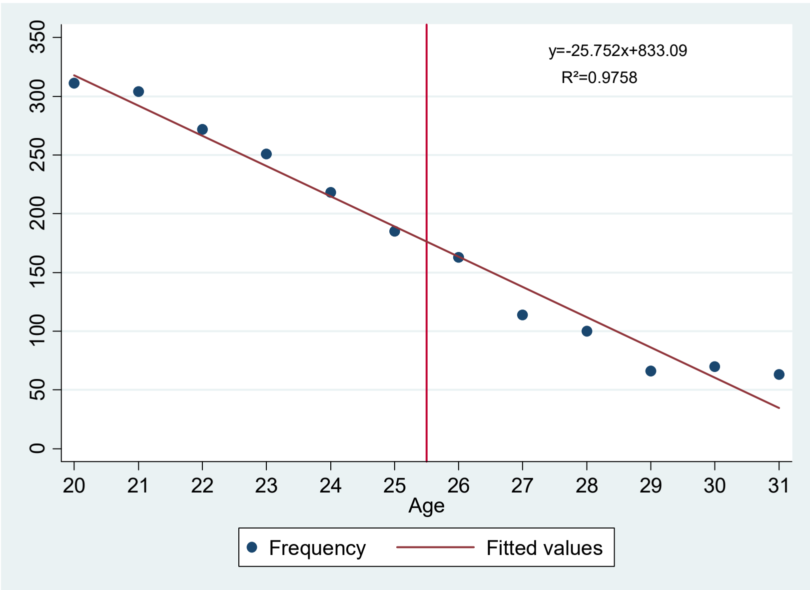
Note: Difference in means significant at: *** 99%, ** 95%, * 90%.
Source: Authors' own calculations.

3.5 RESULTS

3.5.1 'Pull Effect'

Figure 3 plots the number of students enrolled in secondary education for adults by age between ages 20 and 31. A visual inspection does not seem to reveal any evidence about the existence of a jump in enrolment around the cut-off point, only a negative linear relationship between students' age and enrolment rates is observed. This finding is consistent with the theory of human capital that states that younger people are more likely to invest in education due to their lower opportunity cost and their longer working life for a return of the investment (Becker, 1962).

Figure 3. Number of persons enrolled in adult education by age between ages 20 and 31



Source: Authors' own elaboration using data from the administrative sample.

Since it is not possible to compare enrolment rates of males in secondary education for adults in Extremadura with the same rates in previous academic years due to the lack of data, we use the national evolution of enrolment rates as a proxy. The figures of enrolment rates of males in secondary education for adults in Spain in recent academic years are shown in Table 4. The negative relationship between students' age and enrolment rates observed in the region is similar to the trend shown at national level in academic years available. Moreover, note that while *Programa 18-25* was in effect in Extremadura, the national enrolment rates of males in secondary education for adults increased considerably too.

Hence, if regional enrolment rates in this kind of education increased in those years, it was not caused by the programme introduction but by the economic crisis situation in Spain, making people of all ages to get back into education.

Table 4. Figures about males' enrolment rates in secondary education for adults by age groups

Academic year	Age Groups			
	20-24 years	**25-29 years**	**30-39 years**	**40-49 years**
2010-2011	2,330	1,046	841	512
2011-2012	2,288	848	890	530
2012-2013	2,651	1,204	1,197	747
2013-2014	2,286	1,116	1,205	697

Source: Official Educational Figures in Spain (MECD, 2013, 2014, 2015 and 2016)

Table 5 shows the results for 'pull effect' estimates by OLS with school fixed effects. The treatment coefficient is not statistically significant, thus it can be claimed that the policy evaluated did not result in a statistically significant increase in the number of students enrolled in secondary education for adults. Consistent with Table 4 and Figure 3, the *Age* variable seems to be negative and statistically significant, so that the inverse relationship between students' age and enrolment rates is reinforced[40].

Table 5. 'Pull effect' estimates by OLS with school fixed effects

Enrolled students by age and by school	Coefficient	
Constant	12.38	***
	(2.186)	
Treatment	0.441	
	(0.396)	
Years	-0.384	***
	(0.072)	
Average modules enrolled	0.041	
	(0.115)	
R-Squared	0.781	
N	708	

Note: SEs are presented in parentheses. ***significant at 99%.
Source: Authors' own calculations.

All in all we cannot conclude that there was a 'pull effect' induced by the *Programa 18-25* introduction. Therefore, we consider that there are not clear subgroups of students intrinsically motivated and students extrinsically motivated within the treated group, but

[40] It have been estimated other specifications including higher order polynomials on age and its interactions, resulting the same findings. They are available upon request.

most of them were enough motivated on their own to get back into education regardless the programme, and, hence, individuals within the treated group are completely comparable with individuals within the control group.

3.5.2 Impact of *Programa 18-25*

This section discusses estimates of the impact of *Programa 18-25* on two dependent variables: *Diploma* and *Success*. As a first exploration for a possible effect of the programme on outcomes, we plot the average value of each dependent variable as a function of age in Figures 4 and 5, respectively. Both figures show that there is no clear evidence of a significant jump in the average outcomes near the cut-off point. Thus, although it is possible to infer from this graphical analysis that the programme did not have any impact we check this previous result through a fuzzy RDD.

Figure 4. Average of variable *Diploma* against age

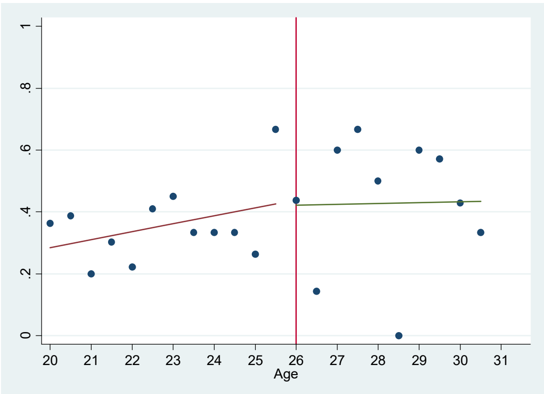
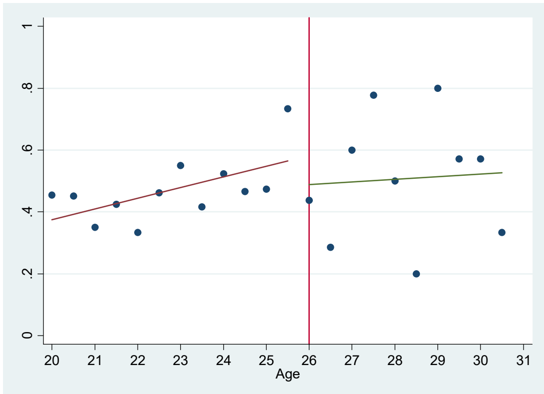


Figure 5. Average of variable *Success* against age



Note: The circles are the average outcomes for student with a given age. The fitted lines are predicted probabilities from a linear probability model (estimated separately on either side of the threshold).
Source: Authors' own elaboration.

For this purpose, two models for each dependent variable are estimated, replicating estimates as many times as needed depending on the combination of bandwidth and number of modules enrolled. Model 1 is the straight fuzzy RDD model estimation without including the set of covariates. In Model 2, control variables are introduced in order to avoid potential bias resulting from using a broad range of data. As the dependent variables employed in both stages are binary variables, RDD is estimated applying logistic regressions[41].

[41] As a robustness check, several alternative specifications have been estimated (quadratic, cubic and higher-order polynomial on age, different specifications of the functional form at both sides of the discontinuity) resulting that our findings are not sensitive to functional-form assumptions. These results are available upon request.

Table 6 reports the model estimation parameters using the *Sample survey*. This table shows coefficients, standard errors and statistical significance of *Treatment* and running variable (*Age*) for the narrowest bandwidth (24-27 years) and the intermediate one (22-29 years). Results, obtained using the widest bandwidth (20-31 years), are presented in Appendix III.

First, regarding estimates of the probability of obtaining upper secondary education diploma at the end of the academic year (*Diploma*), Model 1 shows that treatment does not result statistically significant in any of the combination of bandwidth and number of modules enrolled. Similar findings are obtained in Model 2, concluding that, even introducing covariates in order to control for potential bias, the treatment is not statistically significant. Turning to the second dependent variable, *Success*, neither of the two models reports a statistically significant effect of the programme, whatever the combination of bandwidth and modules is. These results suggest that benefit from *Programa 18-25* does not increase the likelihood of passing every module enrolled in the 2013/14 academic year.

In order to check whether the lack of programme impact is due to not having enough statistical power because of the limited sample size, we replicate the estimates using *Administrative sample*. Table 7 reports the results obtained. Again, the policy has not any effect on any of the dependent variables. This is consistent with the findings obtained using the *Sample survey*, concluding that the *Programa 18-25* effect does not depend on the number of observations employed in the evaluation[42].

From the results shown in tables above, it is worth to note that the *Age* variable behaviour is homogeneous across all estimates, since it appears as statistical and significantly related to dependent variables only in one of the specifications; which justifies the broadening of the age window.

[42] The results (in both samples) are robust to the choice of the age window. See Appendix III.

Table 6. Impact of *Program 18-25* on dependent variables. *Survey sample*.

Number of modules	Dependent variable	Age Bandwidth							
		24-27				22-29			
		Model 1		Model 2		Model 1		Model 2	
		Diploma	Success	Diploma	Success	Diploma	Success	Diploma	Success
4-6	Treatment	2.1477	2.7274	0.4614	1.9483	-0.2008	1.0075	-0.4872	0.1426
		0.3519	(2.0874)	(2.0804)	(2.0146)	(1.0690)	(1.0323)	(1.0115)	(0.9739)
	Age	0.9270 *	0.8292	0.6140	0.6930	0.0967	0.2244	0.0520	0.1339
		(0.5619)	(0.5461)	(0.5085)	(0.4918)	(0.1424)	(0.1370)	(0.1213)	(0.1162)
	Covariates	No	No	Yes	Yes	No	No	Yes	Yes
	N	108				206			
5-6	Treatment	2.1290	2.7495	0.7726	2.0047	-0.3979	0.6769	-0.3483	0.0880
		(2.2677)	(2.1900)	(2.0497)	(1.9855)	(1.0399)	(0.9896)	(1.0137)	(0.9702)
	Age	1.0146	0.9485	0.5308	0.5921	0.0512	0.1677	0.0174	0.0810
		(0.6662)	(0.6338)	(0.5490)	(0.5276)	(0.1452)	(0.1376)	(0.1267)	(0.1210)
	Covariates	No	No	Yes	Yes	No	No	Yes	Yes
	N	94				183			
6	Treatment	1.7054	2.3703	0.7726	2.0047	-0.1394	0.8819	-0.1866	0.2816
		(2.1027)	(2.0236)	(2.0497)	(1.9855)	(1.0746)	(1.0180)	(1.0282)	(0.9916)
	Age	0.7853	0.7454	0.5308	0.5921	0.0788	0.1862	0.0350	0.1069
		(0.6080)	(0.5770)	(0.5490)	(0.5276)	(0.1476)	(0.1391)	(0.1278)	(0.1235)
	Covariates	No	No	Yes	Yes	No	No	Yes	Yes
	N	89				171			

Note: SEs are presented in parentheses. ***significant at 99%; **at 95%; * at 90*. Covariates included in Model 2: *Employment Status*, *Income between 1,000-1,500€*, *Income >1500€*, *Rural*, *Region* and *Modules enrolled*.
Source: Authors' own calculations.

Table 7. Impact of *Program 18-25* on dependent variables. *Administrative sample*.

Number of modules	Dependent variable	Age Bandwidth							
		24-27				24-27			
		Model 1		Model 2		Model 1		Model 2	
		Diploma	Success	Diploma	Success	Diploma	Success	Diploma	Success
4-6	Treatment	2.8915	0.9103	3.0060	1.9196	0.1680	2.2257	-0.7340	2.2953
		(3.4003)	(3.0009)	(3.2810)	(2.8630)	(3.4887)	(3.0774)	(3.7455)	(3.2614)
	Age	0.3742	0.1687	0.3313	0.2745	0.1033	0.2876	0.0332	0.2898
		(0.4584)	(0.4073)	(0.4265)	(0.3699)	(0.2653)	(0.2344)	(0.2828)	(0.2461)
	Covariates	No	No	Yes	Yes	No	No	Yes	Yes
	N	433				907			
5-6	Treatment	2.5308	1.7043	3.3508	3.1519	1.7813	3.1691	0.4544	2.9810
		(3.8720)	(3.3048)	(3.4714)	(2.9452)	(3.7166)	(3.2377)	(4.1948)	(3.6290)
	Age	0.2390	0.2030	0.2771	0.3535	0.2125	0.3487	0.0999	03289
		(0.5235)	(0.4500)	(0.4437)	(0.3737)	(0.2971)	(0.2591)	(0.3326)	(0.2879)
	Covariates	No	No	Yes	Yes	No	No	Yes	Yes
	N	363				770			
6	Treatment	2.5693	2.0656	2.4112	2.5501	1.6330	3.8231	-0.6584	3.5918
		(3.3927)	(2.9002)	(3.5178)	(2.9972)	(5.0684)	(4.3675)	(5.1355)	(4.3905)
	Age	0.3116	0.2867	0.2459	0.3212	0.2202	0.4230	0.02798	0.3975
		(0.4331)	(0.3723)	(0.4299)	(0.3615)	(0.4121)	(0.3556)	(0.4113)	(0.3517)
	Covariates	No	No	Yes	Yes	No	No	Yes	Yes
	N	336				686			

Note: SEs are presented in parentheses. ***significant at 99%; **at 95%; * at 90*. Covariates included in Model 2: *Rural*, *Region* and *Modules enrolled*.
Source: Authors' own calculations.

In short, and consistent with the visual evidence (Figure 4 and 5), the programme does not show any statistically significant impact when a fuzzy regression discontinuity design is applied. The findings reveal that both the probability of getting the diploma of secondary education and the probability of passing all modules enrolled during the 2013/14 academic year, were not statistically and significantly different for treated and control group. Hence, the fact of benefiting from *Programa 18-25* did not involve, at least for males evaluated, clear effects on expected results.

3.6 CONCLUSIONS

The dramatic figures of early school leaving and youth unemployment for Spain point out the need of developing specific policies addressed to increase the average education level of Spanish population, with the aim of increasing job opportunities and, from a general point of view, improving the economic growth prospects.

In that sense, in the region of Extremadura, over 40% of unemployed people registered in the regional Public Employment Service had not completed the compulsory secondary education in 2012. In this context, the regional Government adopted *Programa 18-25* targeted to unemployed people aged between 18 and 25 years (no age limit for long-term unemployed women), without the upper secondary education diploma. The aim of this policy was to motivate these people to get back into education system, through an economic incentive of 1,000 euros if, at the end of the academic year, they obtained the certification.

The aim of this chapter is to evaluate the *Programa 18-25* impact on males in the 2013/14 academic year. First, we analyse whether this policy caused some 'pull effect', i.e. increasing the enrolment rates of males under 25 years old to a greater extent than the enrolment rates of males ineligible for the policy. Once this hypothesis was rejected, we tried to assess the effect of the programme on the probability of obtaining upper secondary education diploma (first dependent variable: *Diploma*) and on the probability of passing every module enrolled (second dependent variable: *Success*). The fuzzy regression discontinuity design was selected among quasi-experimental methods, due to the fact that it was the most suitable one according to the nature of the programme, since excluding unemployed people older than 25 years causes a jump or discontinuity in treatment but also takes into account the observed presence of always-takers and never-takers. To carry out the empirical analysis, we employ administrative data provided by the Regional Department of Education and Culture and data obtained from surveys sent out to every student enrolled in secondary education for adults during the academic year evaluated.

The findings show that, at least for males evaluated, the likelihood of obtaining the diploma or passing every module enrolled did not depend on being a beneficiary from *Programa 18-25*. Furthermore, these results are robust to alternative regression specifications and a variety of bandwidth.

In this research we would like to highlight, first, the need to carry out more formal impact evaluations that allow to distinguish causation from accidental associations or correlations. These evaluation should be used by policy makers as a worth feedback tool for taking rational decisions about continuing or nor with public programs at work. In this sense, future public programs should be designed, before starting to apply the program, to enhance the evaluation results. Second, the null effect of *Programa 18-25*: the economic incentive offered did not increase neither the probability of completing successfully the academic year nor the likelihood of obtaining upper secondary education diploma. Finally, it is worth considering that results for Extremadura are likely to be relevant to other Spanish regions, given that their governments are currently running or considering this type of policies.

3.7 REFERENCES

Angrist, J.D. and Lavy, V. (1999). Using maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics*, 144 (2), 533-575.

Angrist, J.D. and Pischke, J.S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Barro, R.J. (2001). Human capital and growth. *The American Economic Review*, 91(2), 12-17.

Battistin, E. and Rettore, E. (2002). Testing for programme effects in a regression discontinuity design with imperfect compliance. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 165(1), 39-57.

Becker, G.S. (1962). Investment in human capital: A theoretical analysis. *The Journal of Political Economy*, 9-49.

Betcherman, G., Dar, A. and Olivas, K. (2004). *Impacts of active labor market programs: New evidence from evaluations with particular attention to developing and transition countries*. Social Protection Discussion Paper, Washington DC: World Bank.

Blundell, R., Meghir, C., Dias, M.C. and Reenen, J.V. (2004). Evaluating the employment impact of a mandatory job search program. *Journal of the European Economic Association*, 2(4), 569-606.

Cook, T.D. (2008). Waiting for life to arrive: a history of the regression-discontinuity design in psychology, statistics and economics. *Journal of Econometrics*, 142(2), 636-654.

Crawford, C., Dearden, L. and Greaves, E. (2014). The drivers of month-of-birth differences in children's cognitive and non-cognitive skills. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 177(4), 829-860.

Dar, A. and Tzannatos, Z. (1999). *Active labor market programs: A review of the evidence from evaluations*. Social Protection Discussion Paper, Washington DC: World Bank.

De Giorgi, G. (2005). *Long-term effects of a mandatory multistage program: the New Deal for Young People in the UK* (No. 05/08). IFS Working Papers, Institute for Fiscal Studies (IFS).

Fitzenberger, B. and Speckesser, S. (2007). Employment effects of the provision of specific professional skills and techniques in Germany. *Empirical Economics*, 32(2-3), 529-573.

Gerfin, M. and Lechner, M. (2002). A microeconometric evaluation of the active labour market policy in Switzerland. *The Economic Journal*, 112(482), 854-893.

Gertler, P., Martinez, S., Premand, P., Rawlings, L.B. and Vermeersch, C.M.J. (2011). *Impact Evaluation in Practice.* Washington DC: World Bank.

Hahn, J., Todd, P. and Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1), 201-209.

Imbens, G.W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2), 615-635.

Imbens, G.W. and Wooldridge, J.M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47 (1), 5-86.

Jacob, B.A. and Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *Review of Economics and Statistics*, 86(1), 226-244.

Khandker, S.R., Koolwal, G.B. and Samad, H.A. (2010). *Handbook on impact evaluation: quantitative methods and practices*. Washington DC: World Bank.

Kluve, J., Lehmann, H. and Schmidt, C.M. (2008). Disentangling treatment effects of Active Labor Market Policies: The role of labor force status sequences. *Labour Economics*, 15(6), 1270-1295.

Larsson, L. (2003). Evaluation of Swedish youth labor market programs. *Journal of Human Resources*, 38(4), 891-927.

Lemieux, T. and Milligan, K. (2008). Incentive effects of social assistance: A regression discontinuity approach. *Journal of Econometrics*, 142(2), 807-828.

Lindley, J., Mcintosh, S., Roberts, J., Murray, C.C. and Edlin, R. (2015). Policy evaluation via a statistical control: A non-parametric evaluation of the 'Want2Work'active labour market policy. *Economic Modelling*, 51, 635-645.

Malo, M.A. and Cueto, B. (2016). El impacto de las políticas activas de mercado de trabajo en España. *Proceedings of the XXIII Meeting of Public Economics*, Ourense, Spain.

Mauldon, J., Malvin, J., Stiles, J., Nicosia, N. and Seto, E.Y. (2000). *The impact of California's Cal-Learn demonstration project, final report.* UC Data Archive & Technical Assistance.

Ministerio de Educación, Cultura y Deporte (2013): *Las cifras de la educación en España. Curso 2010-2011*. Madrid: MECD.

Ministerio de Educación, Cultura y Deporte (2014): *Las cifras de la educación en España. Curso 2011-2012*. Madrid: MECD.

Ministerio de Educación, Cultura y Deporte (2015): *Las cifras de la educación en España. Curso 2012-2013*. Madrid: MECD.

Ministerio de Educación, Cultura y Deporte (2016): *Las cifras de la educación en España. Curso 2013-2014.* Madrid: MECD.

Oreopoulos, P. (2007). Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics*, 91(11-12), 2213-2229.

Oreopoulos, P. and Salvanes, K.G. (2011). Priceless: The nonpecuniary benefits of schooling. *The Journal of Economic Perspectives*, 25(1), 159-184.

Psacharopoulos, G. (2007). *The costs of school failure a feasibility study*. European Expert Network on Economics of Education (EENEE).

Rodríguez-Planas, N. and Jacob, B. (2010). Evaluating active labor market programs in Romania. *Empirical Economics*, 38(1), 65-84.

Schlotter, M., Schwerdt, G. and Woessmann, L. (2011). Econometric methods for causal evaluation of education policies and practices: a non-technical guide. *Education Economics*, 19(2), 109-137.

Schwerdt, G., Messer, D., Woessmann, L. and Wolter, S.C. (2012). The impact of an adult education voucher program: Evidence from a randomized field experiment. *Journal of Public Economics*, 96(7), 569-583.

Thistlewaite, D. and Campbell, D. (1960). Regression-discontinuity analysis: an alternative to the ex-post facto experiment. *Journal of Educational Psychology* 51(6), 309–317.

Van der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: A Regression–Discontinuity Approach. *International Economic Review*, 43(4), 1249-1287.

Van der Klaauw, W. (2008). Regression–discontinuity analysis: a survey of recent developments in economics. *Labour*, 22(2), 219-245.

3.8 LEGAL TEXTS

- Order August 1, 2008, which regulates *Compulsory Secondary Education for Adults in the Autonomous Community of Extremadura* (DOE No. 158, August 18th 2008).
- Announcement of November 15, 2012 by publicly disclosing the Plan of the Public Employment Service of Extremadura and the Regional Department of Education and Culture for the development of a specific training program aimed at obtaining the diploma of compulsory secondary education for unemployed adults aged between 18 and 25 years (DOE No. 228, November 26th 2012).
- Act 4/2011, of March 7, Extremadura Education (DOE No. 47, March 9th 2011).
- Resolution of September 2nd, 2013, of the Regional Department of Employment, Industry and Innovation and the Regional Minister of Education and Culture, for which the publication of the Plan of Public Employment Service of Extremadura and the Regional Department of Education and Culture is available for the development of specific training programmes for unemployed people to obtain qualifications of initial vocational training, secondary education diploma and technical and advanced vocational training, of August 6th, 2013 (DOE No. 173, September 6th 2013).
- Resolution of August 28th, 2014, of the Regional Minister of Employment, Women and Social Policy and the Regional Minister of Education and Culture, for which the publication of the Plan of Public Employment Service of Extremadura and the Regional Department of Education and Culture is available for the development of specific training programmes for unemployed people to obtain qualifications of compulsory secondary education, completion diploma on key competences, initial vocational training diploma, and technical and advanced vocational training diploma (DOE No. 173, September 9th 2014).

3.9 APPENDIX I: LITERATURE REVIEW

Author	Policy evaluated	Place	Dates	Data	Evaluation approach	Results
Battistin & Rettore (2002)	Vocational training programme targeted at unemployed people: an intensive course to improve participants' computer skills	Torino (Italy)	From October 1995 to June 1996	Data obtained by telephone interviewing 17 months after the programme completion	Regression discontinuity design	Attending the programme did not increase the probability of being employed 17 months after completing the training
Blundell *et al.* (2004)	*New Deal* for young people aged between 18 and 24 years old who had been claiming unemployment insurance for six months (pilot experience). The programme combined job search assistance, wage subsidies to employers and education and training	United Kingdom	From January to March, 1998	Data provided by the pilot areas before the National Roll Out of the programme and on data available following the National Roll Out	Difference in differences combined with matching	The programme increased the probability of young men of finding a job in the first four months of the *New Deal*
De Giorgi (2005)	*New Deal* for young people (National Roll Out): mandatory multistage policy targeted at 18-24 year old unemployed	United Kingdom	From April 1998 to December 2001	Administrative dataset (New Deal Evaluation Database) and longitudinal sample of UK unemployed (JUVOS)	Regression discontinuity design	The employment probability within 18 months since starting the programme increased for male participants
Fitzenberger & Speckesser (2007)	Specific professional skills and techniques programme targeted at unemployed persons or persons at risk of becoming unemployed	Germany	From 1993 to the end of 1997	Administrative data about (un)employment and participation in active labour market programmes generated by the Federal Employment Office	Propensity score matching	Negative lock-in effects were found shortly after the beginning of the programme. However, later, the effects turned positive and persisted until the end of the evaluation period

Gerfin & Lechner (2002)	Active Labour Market Programs: It consisted on vocational training programmes, employment programmes and temporary wage subsidy.	Switzerland	Programs starting after 1st January 1998	Database created from administrative records provided by the unemployment insurance system and by the social security system	Matching estimator	Employment programmes performed poorly, Vocational training ones showed mixed effects depending on the sub-population considered, and Temporary wage subsidy resulted to be effective in terms of the chances to find a job
Kluve *et al.* (2008)	Two labour market policies: a training programme and *Intervention works* (a wage subsidy scheme)	Poland	From January 1992 to August 1996	Data from the 18th wave of the Polish Labour Force Survey	Difference in differences combined with matching	The training programme increased its participants' employment rate while *Interventions works* policy affected negatively to the employment prospects of treated
Larsson (2003)	Two active labour programmes for youth people (aged 20-24 years). *Youth practice*: a subsidize work programme for unemployed with a high school diploma. *Labour market training*: aimed at improve the skills of unemployed with low education.	Sweden	Between years 1992 and 1995	Data collected from Swedish National Labour Market Board and Statistics Sweden	Matching estimator	*Youth practice* and *Labour market training* resulted to have negative short-term effects on earnings and employment. However, in the long-term, the effects appeared to be statistically insignificant

Lemieux and Milligan (2008)	Effects of social assistance benefits on the labour market behaviour of men without children	Quebec (Canada)	Between years1985 and 1986	Data from the 1986 and 1991 Censuses and from the Labour Force Survey (LFS)	Regression discontinuity design	More generous social assistance benefits reduce the probability of employment of less-educated men without children. Also, the employment rate of all men (independently if they have children or not) drops to a lesser extent in response to the higher benefits
Lindley *et al.* (2015)	*Want2Work* (pilot scheme): Active labour market policy aimed to improve the re-employment chances of benefit recipients, primarily those receiving Incapacity Benefits (people unable to work due to health problems).	Particular areas of Wales	From September 2004 to March 2008	Data on *Want2Work* participants collected by the Welsh Assembly Government and data drawn from the Quarterly Labour Force Survey (QLFS) for Great Britain	Propensity score matching	*Want2Work* participants were more likely to move into employment. Besides, restricting the analysis to the original target (incapacity benefit participants), the employment impact of the programme increased around to 13 percentage points
Mauldon *et al.* (2000)	*Cal Learn*: programme designed to help parenting teenagers (under 19 years old) on welfare increase their graduation rates (High School Diploma (HSD) or General Education Development certificate (GED)) through a mixture of cash conditional transfers and case management services	Four counties in California (Alameda, Los Angles, San Bernardino and San Joaquin)	From October 1994 to October 1999	Combination of survey data, county administrative data, high school records and official GED statistics	Randomization design	The programme substantially increased graduation rates, although these increases were concentrated exclusively in GEDs rather than in HSDs. There were no significant benefits of the programme in terms of employment, earnings or reduced use of welfare within the period of time evaluated.

Rodríguez-Planas & Jacob (2010)	Four active labour market programmes: training and retraining, self-employment assistance, public employment and relocation services	Romania	The year 1999	Follow-up survey designed and collected for the evaluation	Propensity score matching	Training and relocation services programme had a positive impact both on employment rates and on average earnings. Self-employment assistance programme increased the probability of being employed but non-effects were found regarding earnings. However, the effect of public employment programme on its participants' employment prospects was negative.
Schwerdt *et al.* (2012)	A general voucher programme for adult education (individuals aged between 20 to 60; no limitation with respect to employment status)	Switzerland	The year 2006	Data from the Swiss Labour Force Survey (SLFS)	Randomization design	No statistically significant effects on earnings or employment probability 1 year after treatment were found. A sub-group analysis revealed that individuals with low level of education were more likely to benefit from participating in adult education; however highly-educated individuals were the ones who took advantages of the voucher.

3.10 APPENDIX II:

SURVEY DISTRIBUITED AMONG STUDENTS ENROLLED IN SECONDARY EDUCATION FOR ADULTS

PRESENTATION: In order to evaluate the development of adult education programmes, we would appreciate you to answer the following questions (the ID field is necessary in order to record data, but in any case the data will be treated in a completely anonymous way).

ID: …………………………

1.- RESPONDENT GENDER

a) Male b) Female

2.- DATE OF BIRTH (dd/mm/yyyy) …… / …… / …………

3.- BIRTH COUNTRY

a) Spain b) Other (Indicate) ……………………….

4.- WHAT WAS YOUR EMPLOYMENT SITUATION WHEN YOU ENROLLED IN ADULT EDUCATION?

a) Unemployed b) Employed c) Self-Employed

5.- IF YOU WERE UNEMPLOYED AT THE TIME OF ENROLLING IN ADULT EDUCATION, HOW LONG WERE YOU UNEMPLOYED BEFORE STARTING THE COURSE?

a) Less than 3 months
b) Between 3 and 6 months
c) Between 6 and 12 months
d) More than 12 months
e) I have never worked before

6.- HAVE YOU APPLIED FOR THE €1000 SCHOLARSHIP IF YOU MEET THE SET REQUIREMENTS OF ATTENDANCE AND ACADEMIC RESULTS?

a) Yes b) No

7.- ¿DO YOU HAVE FAMILY RESPONSIBILITIES (CHILDREN, ADULTS, SICK OR DISABLED…)?

a) Yes b) No

8.- IF SO, DO YOU BENEFIT FROM THE POSSIBILITY OF HIRING SOMEONE WHO MAY TAKE CARE OF THEM DURING YOUR ATTENDANCE TO THE TRAINING CLASSES?

a) Yes b) No

9.- WHAT IS YOUR FATHER'S BIRTH COUNTRY?

a) Spain b) Other (indicate) ……………………….

10.- WHAT IS YOUR MOTHER'S BIRTH COUNTRY?

a) Spain b) Other (indicate) ……………………….

11.- WHAT IS YOUR FATHER'S PROFESSION? (if retired, unemployed or passed away, indicate his previous activity):

a) Household or without profession defined.
b) Worker in construction, general industry or agriculture, low-level official, civil servant or similar
c) Professor, entrepreneur (with employees in charge), medium-high level civil servant, private company manager, liberal professional (doctor, lawyer, ...), or similar.

12.- WHAT IS YOUR MOTHER PROFESSION? (if retired, unemployed or passed away, indicate his previous activity):

a) Household or without profession defined.

b) Worker in construction, general industry or agriculture, low-level official, civil servant or similar

c) Professor, entrepreneur (with employees in charge), medium-high level civil servant, private company manager, liberal professional (doctor, lawyer, ...), or similar.

13.- WHAT IS THE LEVEL OF STUDIES ACHIEVED BY YOUR FATHER?

a) Without studies
b) Secondary education
c) Spanish baccalaureate or Vocational training
d) Degree-qualified
e) DK/DA

14.- WHAT IS THE LEVEL OF STUDIES ACHIEVED BY YOUR MOTHER?

a) Without studies
b) Secondary education
c) Spanish baccalaureate or Vocational training
d) Degree-qualified
e) DK/DA

15.- MONTHLY TOTAL INCOME OF THE FAMILY YOU LIVE WITH IS AROUND:

a) Less than 1,000 euros
b) Between 1,000 and 1,500 euros
c) More than 1,500 euros

16.- ONCE YOU HAVE REACHED THE DIPLOMA OF UPPER SECONDARY EDUCATION, WOULD YOU LIKE TO START A HIGHER DEGREE?

a) Yes, baccalaureate
b) Yes, vocational training
c) Yes, college
d) No

17.- WHERE ARE YOU ATTENDING ADULTS EDUCATION TRAINING?

- CENTER NAME: ……………………………………………
- LOCATION: …………………………………………

18.- DO YOU NEED TO MOVE TO ATTEND CLASSES OF ADULTS EDUCATION TRAINING?

a) No, the center is in my area
b) Yes, less than 10 Km
c) Yes, between 10 and 25 Km
d) Yes, more than 25 Km

3.11 APPENDIX III:

Table III.1. Impact of *Programa 18-25* on dependent variables: Diploma and Success, bandwidth 20-31 years, both samples

Number of modules	Dependent variables	Survey sample				Administrative sample			
		Model 1		Model 2		Model 1		Model 2	
		Diploma	Success	Diploma	Success	Diploma	Success	Diploma	Success
4-6	Treatment	-0.5140	0.3356	-0.7740	-0.2177	-0.5346	-0.3964	-0.4590	-0.2342
		(0.8549)	(0.8144)	(0.8580)	(0.8193)	(1.4596)	(1.2834)	(1.5421)	(1.3464)
	Age	0.0267	0.0900	0.0147	0.0594	0.0333	0.0567	0.0367	0.0650
		(0.0834)	(0.0787)	(0.0756)	(0.0711)	(0.0864)	(0.0760)	(0.0903)	(0.0790)
	Covariates	No	No	Yes	Yes	No	No	Yes	Yes
	N	299				1,412			
5-6	Treatment	-0.5487	0.1844	-0.6012	-0.0984	0.4370	0.4351	0.1321	0.2242
		(0.8248)	(0.7772)	(0.8627)	(0.8232)	(1.5994)	(1.3840)	(1.7378)	(1.4946)
	Age	0.0195	0.0701	0.0149	0.0520	0.0811	0.1011	0.0587	0.0854
		(0.0825)	(0.0770)	(0.0785)	(0.0740)	(0.0999)	(0.0867)	(0.1081)	(0.0932)
	Covariates	No	No	Yes	Yes	No	No	Yes	Yes
	N	267				1,197			
6	Treatment	-0.5596	0.3159	-0.5479	0.0008	-0.2689	-0.0943	-0.7645	-0.3125
		(0.8681)	(0.8161)	(0.8898)	(0.8475)	(1.7545)	(1.5064)	(1.8339)	(1.5672)
	Age	-0.0019	0.0733	0.0039	0.0592	0.0380	0.0745	0.0015	0.0586
		(0.0862)	(0.0805)	(0.0793)	(0.0752)	(0.1089)	(0.0937)	(0.1118)	(0.0958)
	Covariates	No	No	Yes	Yes	No	No	Yes	Yes
	N	245				1,049			

Note: SEs are presented in parentheses. ***significant at 99%; **at 95%; * at 90*. Covariates included in Model 2 in Sample survey: *Employment Status*, *Income between 1,000-1,500€*, *Income >1500€*, *Rural*, *Region* and *Modules enrolled*. Covariates included in Model 2 in Administrative sample: *Rural*, *Region* and *Modules enrolled*.
Source: Authors' own calculations.

CONCLUSIONES GENERALES

Con la presente tesis doctoral hemos querido contribuir a que en nuestro país se vaya desarrollando una cultura evaluadora de las políticas públicas y más concretamente en el ámbito de la educación. La investigación consta de tres ensayos que están relacionados aunque son independientes. Puesto que cada uno de ellos cuenta con un apartado específico de conclusiones, en esta sección nos limitaremos a presentar los principales resultados obtenidos y algunas recomendaciones de política educativa derivadas de la investigación.

En el primer capítulo partimos de un experimento natural, el mes de nacimiento, y evaluamos su efecto sobre la probabilidad de repetir curso de los alumnos españoles a la edad de 15 años. La principal conclusión es que el impacto del mes de nacimiento sobre el rendimiento educativo es relevante y persistente en el tiempo. El que los alumnos nacidos al final del año se vean penalizados frente a los que lo hicieron a principios del mismo, no parece una situación aceptable ni en términos de eficiencia ni de equidad entendida como igualdad de oportunidades educativas, y lo cierto es que el sistema educativo español no cuenta con ningún instrumento específico para solventar o mitigar este problema.

Entre las posibles medidas de política educativa, la literatura sugiere, por un lado, flexibilizar las reglas de incorporación al sistema educativo de tal forma que los padres puedan elegir el momento de escolarizar a sus hijos, sobre todo en el caso de los nacidos en los últimos meses del año; por otra parte, también se plantea proporcionar tutorías compensatorias para los más jóvenes del aula o duplicar el número de cursos en la formación básica, con el fin de reducir a la mitad la diferencia máxima de 12 meses de edad entre los matriculados en un mismo curso académico.

Los resultados del segundo capítulo, en el que estudiamos el impacto del incremento exógeno del porcentaje de alumnos inmigrantes en España sobre la repetición de curso, muestran que la incorporación a nuestro sistema educativo de este tipo de alumnado no disminuyó, en media, las tasas de promoción de curso de las escuelas españolas. De hecho, la situación ha sido incluso beneficiosa para los alumnos nativos puesto que los estudiantes inmigrantes se veían afectados en mayor medida por la repetición de curso. Sin embargo, se comprobó que la acumulación de estudiantes extranjeros incrementó el porcentaje medio de repetidores, tanto inmigrantes como nativos, siendo necesarias, en este último caso, mayores concentraciones de alumnado inmigrante para alcanzar el mismo impacto en los resultados.

La pregunta relevante es por qué el incremento de alumnos inmigrantes produce tales efectos en el sistema educativo y a través de qué canales. Esto podría deberse a que los alumnos inmigrantes se incorporan a nuestro sistema educativo con un dominio del idioma insuficiente y un menor nivel educativo que provocaría que, al aumentar el número de alumnos inmigrantes por clase, se reduzca el nivel educativo medio del aula. Algunas estrategias educativas para 'solucionar' esta situación podrían ser redistribuir el número de alumnos inmigrantes entre las escuelas (evitando concentraciones demasiado elevadas) y/o proporcionar clases adicionales para compensar las dificultades de adaptación al nuevo idioma, sobre todo en las edades más tempranas.

Por último, el tercer capítulo evalúa el *Programa 18-25*, una medida de política educativa implantada por el Gobierno de Extremadura, concluyendo que, al menos para la población evaluada, el programa no resultó eficaz al no conseguir sus objetivos a corto plazo. En primer lugar, la medida no generó el efecto llamada esperado, manteniéndose las tasas de matriculación en educación para adultos de los menores de 25 años, en niveles similares a los de los cursos anteriores a su implantación. Además, el incentivo monetario no supuso un incremento ni de la probabilidad de obtener el título de la ESO, ni de finalizar el curso académico superando todos los módulos matriculados. Del análisis realizado en este ensayo destacamos, por un lado, la importancia de llevar a cabo una correcta evaluación que permita conocer el grado de cumplimiento de los objetivos iniciales de una determinada política y por otro, su utilidad para aquellas comunidades autónomas que han puesto en marcha programas similares y siempre con la salvedad de que los resultados pudieran ser diferentes en otros contextos.

Para concluir la presente tesis doctoral, nos gustaría transmitir la necesidad de que las decisiones de política educativa en España se basen en análisis experimentales o cuasi-experimentales que permitan medir el efecto de tales medidas. Evidentemente, este tipo de análisis no pretende sustituir la labor del político que es a quien corresponde tomar las decisiones, pero la política real precisa evaluaciones basadas en evidencias que mejoren el diseño y la efectividad de aquella. En ese sentido, cualquier política o intervención pública debería ser diseñada desde sus orígenes con el objetivo de ser evaluada y, en consecuencia, contar con bases de datos longitudinales que permitieran un seguimiento del posible impacto de la misma más allá del corto plazo.

CONCLUSIONS

The aim of the current PhD dissertation is to contribute to develop an evaluation culture of public policies, and particularly in the field of education in Spain. This research consists of three essays that are related to each other, but independent at the same time. Since each of them has a specific section of conclusions, in this section we will only present the main results and some educational policy recommendations derived from the research.

In the first chapter we start from a natural experiment framework, the month of birth, and we evaluate its effect on the probability of 15-year-old Spanish students repeating a grade. The main conclusion is that the impact of month of birth on educational performance is relevant and persistent over time. The fact that the pupils born at the end of the year are penalized by the education system compared to those born at the beginning does not seem acceptable either in terms of efficiency or equity, understood as equal educational opportunities, and the truth is that the Spanish education system does not seem to have any specific instrument to solve or mitigate this problem.

Amongst the possible proposals of educational policy, the literature suggests, on the one hand, more flexible rules, so that parents can choose when to enrol their children, especially in the case of those born in the last months of the year; on the other hand, the provision of additional tuition for the younger students, or doubling the number of primary education classes, in order to halve the maximum difference of 12 months among those enrolled in the same academic year.

The results of the second chapter, in which the impact of the exogenous increase of the percentage of immigrant students in Spain on grade retention is studied, show that immigrant students joining Spanish education system did not, on average, decrease promotion rates of Spanish schools. In fact, the situation has been even beneficial to native students because immigrant students were more affected by grade retention. However, it was found that the concentration of foreign students increased the average percentage of repeaters, both immigrants and natives, being a higher concentration of immigrant students still necessary in the latter case, to achieve the same impact on the results.

The question of why the increase of immigrant students produces such an effect on the education system and the channels it uses, there could be several answers. A possible reason would be a language deficit and lower educational level of immigrant students joining the Spanish education system causing that, increasing the number of immigrant students per classroom, the average educational level is reduced in terms of classroom. Some educational strategies to address this situation would be to redistribute the number of immigrant students per school (avoiding high concentrations) and/or provide additional training to compensate for difficulties in adapting to the new language, especially at early ages.

Finally, the third chapter assesses the *Programa 18-25*, an educational policy implemented by the Government of Extremadura, concluding that, at least for people under evaluation, the programme was not effective by not getting its short-term goals. First, the policy did not result in the expected *'pull effect'*, keeping enrolment rates in adult education of individuals under 25 years similar to the rates registered in academic years prior to the programme introduction. In addition, the economic incentive did not involve an increase or the probability of obtaining the diploma of upper secondary education, or to finish the academic year passing all modules enrolled. From the analysis performed in this essay we can highlight, on the one hand, the importance of carrying out a proper evaluation to determine the degree of compliance with the initial objectives of a policy and on the other, its usefulness for those regions that have implemented similar programmes and always with the caveat that the results could be different in other contexts.

To conclude this PhD dissertation we would like to convey the need for decisions on educational policy in Spain that are based on experimental or quasi-experimental designs that allow measuring the impact of such interventions. Obviously this type of analysis does not intend to replace the role of policy-makers, who are the last responsible for making decisions, but real policies requires assessment based on empirical evidence to improve the design and effectiveness of such policies. In that sense, any public intervention should be designed from its origin with the goal of being evaluated and therefore should have longitudinal databases that allow monitoring the possible impact of it beyond the short term.

RESUMEN / ABSTRACT

RESUMEN

El objetivo de la presente tesis doctoral es aportar varias experiencias de evaluación en el ámbito educativo español, con el fin de ir enriqueciendo la escasa cultura evaluadora existente y la necesidad de la rendición de cuentas del sector público en nuestro país. La tesis se compone de tres ensayos que están relacionados aunque son independientes. Los dos primeros se centran en cuestiones que han caracterizado al sistema educativo español en los últimos años como es la repetición de curso o la llegada de alumnos inmigrantes y el tercero evalúa una política educativa específica puesta en marcha por el Gobierno de la Comunidad Autónoma de Extremadura.

En el primer capítulo partimos de un experimento natural, el mes de nacimiento, y evaluamos su efecto sobre la probabilidad de repetir curso de los alumnos españoles a la edad de 15 años. La principal conclusión es que el impacto del mes de nacimiento sobre el rendimiento educativo es relevante y persistente en el tiempo. De hecho, los estudiantes nacidos en el último bimestre del año tienen una probabilidad de repetir un 80% superior a la de sus compañeros de clase nacidos en el primer bimestre.

El objetivo del segundo capítulo es evaluar el impacto del incremento exógeno del porcentaje de alumnos inmigrantes en España sobre la repetición de curso. La conclusión fundamental es que la llegada de alumnos inmigrantes no da lugar, en media, a una reducción de las tasas de promoción de curso en las escuelas, sino que incluso es beneficioso para los nativos. Sin embargo, la acumulación de inmigrantes sí tiene un impacto negativo sobre los porcentajes de repetidores tanto de inmigrantes como de nativos, aunque en este último caso es necesaria una acumulación mayor para poder observar un efecto similar.

Por último, el tercer capítulo evalúa el *Programa 18-25*, una medida de política educativa implantada por el Gobierno de Extremadura. En este tercer capítulo se trata de comprobar el efecto de dicho programa sobre la población masculina durante el curso académico 2013/2014. Los resultados obtenidos ponen de manifiesto que en primer lugar, la medida no generó el efecto llamada esperado, manteniéndose las tasas de matriculación en educación para adultos de los menores de 25 años, en niveles similares a los de los cursos anteriores a su implantación. Además, el incentivo monetario no supuso un incremento ni de la probabilidad de obtener el título de la ESO, ni de finalizar el curso académico superando todos los módulos matriculados.

ABSTRACT

The aim of the current PhD dissertation is to provide several evaluation experiences about the Spanish education system, in order to enrich the scarce evaluation culture and the need for accountability public sector in Spain. This PhD dissertation is composed of three essays that are related to each other, but independent at the same time. The first two focus on issues that have defined the Spanish education system in recent years and the third one assesses a specific educational policy implemented by the government of Extremadura.

In the first chapter we start from a natural experiment framework, the month of birth, and we evaluate its effect on the probability of 15-year-old Spanish students repeating a grade. The main conclusion is that the impact of month of birth on educational performance is relevant and persistent over time. Furthermore, a bimonthly impact on grade retention can also been identified since the students who were born in the last two months of the year present a probability to repeat a grade 80% higher than their classmates who were born in the first two months of the year.

The goal of the second chapter is to assess the impact of the exogenous increase of the percentage of immigrant students in Spain on grade retention. The main conclusion derived from empirical analysis is that the arrival of immigrant students does not on average decrease school promotion rates and is even beneficial to native students. However, the concentration of immigrant students at the same school over a threshold does have a negative impact on the percentage of repeaters, for both immigrants and natives, although in the latter case greater concentrations of immigrant students are required to observe similar effects.

Finally, the third chapter assesses the *Programa 18-25*, an educational policy implemented by the Government of Extremadura. In this chapter the impact of this programme on the male population during its second call (the 2013/14 academic year) is measured. The results show that first, the policy did not result in the expected *'pull effect'*, keeping enrolment rates in adult education of individuals under 25 years similar to the rates registered in academic years prior to the programme introduction. In addition, the economic incentive did not involve an increase or the probability of obtaining the diploma of upper secondary education, or to finish the academic year passing all modules enrolled