

# **Bilingual Schooling and Earnings: Evidence from a Language-in-Education Reform\***

Lorenzo Cappellari

Università Cattolica Milano

Antonio Di Paolo

AQR-IREA, University of Barcelona

**February 2018**

## **Abstract**

We estimate the wage effects of bilingual education for the first time using a reform that introduced bilingualism in Catalan schools. Variation across years of schooling and birth cohorts provides identification. We find substantial effects of bilingual education, which increase baseline returns to education by about 20 percent. Robustness checks show that effects stem from exposure to the language-in-education reform and are neither a consequence of unobservable determinants of educational attainment nor an artefact of education-cohort specific trends in wages. The reform was mostly beneficial for individuals of non-Catalan background from low parental background, thus achieving its goal of levelling the playing field in a bilingual society.

**Keywords:** Bilingual education, returns to schooling, language-in-education reform, Catalonia

**JEL classifications:** J24, J31, I28

---

\*Funding from Càtedra Pasqual Maragall at Universitat de Barcelona, the MEC grant no. ECO2013-41022-R, the PRIN grant no. 2010T8XAXB and the RecerCaixa grant no. 2016ACUP00020 is gratefully acknowledged. Thanks are due to Frederic Udina, Josep Maria Martínez at IDESCAT staff for their assistance with the data. Ada Ferrer-i-Carbonell, Ramon Caminal, Albert Costa, Xavier Vila-i-Moreno, Germà Bel, Joan-Ramon Borrell and participants in seminars at EUR, URV, UdG, UB, IRVAPP-Trento, SFI Copenhagen and IDESCAT, and at EALE-SOLE, ESPE, IWAE, RES, SAEe and the Lisbon Workshop on Economics and Econometrics of Education provided useful comments. The usual disclaimers apply. Address for correspondence: Cappellari, Department of Economics and Finance, Università Cattolica Milano, Largo Gemelli 1, 20123 Milano, Italy, [lorenzo.cappellari@unicatt.it](mailto:lorenzo.cappellari@unicatt.it), Tel.: +39 02 72343010; Di Paolo, AQR-IREA, University of Barcelona, Avinguda Diagonal 690, 08034 Barcelona, Spain, [antonio.dipaolo@ub.edu](mailto:antonio.dipaolo@ub.edu).

## 1. Introduction

Language of instruction is a key input in the educational production function. While most students still receive education only in their native language, bilingual education is becoming increasingly popular among families and is attracting growing attention from educationalists. Bilingualism affects the quality of education, favours the development of cognitive skills and its effectiveness in fostering educational outcomes has been found to increase with early exposure (Bleakley and Chin, 2008). Evidence in cognitive psychology and related disciplines shows that bilingualism has positive impacts on several dimensions of cognition, including a reduction of the rate of skill depletion in the adult population (Adesope et al., 2010; Costa and Sebastian-Gallés, 2014). To the extent to which bilingualism affects cognitive skills, bilingual speakers would have higher earnings potential (Saiz and Zoido, 2005). Policies introducing bilingual education are also a mean for building a common identity in regions or countries where different cultural and national identities coexist (Aspachs-Bracons et al., 2008; Clots-Figueras and Masella, 2013). Despite the relevance of bilingual education for both school quality and skill formation, little is still known about its effects on economic outcomes. In this paper we contribute new evidence on this and provide the first quasi-experimental estimates of the earnings effects of bilingual schooling.

We estimate the labor market value of bilingual education using a reform of the language of instruction in the Spanish region of Catalonia, the so-called ‘Linguistic Normalization’. Until 1983, Spanish was the only official language in Catalan schools. Starting from that academic year, Catalan became a co-official language of instruction, which means that both Catalan and Spanish were used as medium of instruction. While other studies have investigated the effects of this reform on Catalan proficiency and on feelings of Catalan identity, we are the first to estimate its earnings effects.<sup>1</sup> Our empirical framework is built around a difference-in-difference (DiD) identification strategy akin to Angrist and Lavy (1997) and Angrist et al. (2008), which compares the returns to

---

<sup>1</sup> See Aspachs-Bracons et al. (2008) and Clots-Figueras and Masella (2013).

schooling across cohorts of students that were differentially exposed to bilingual education as a consequence of the reform.

Catalan has been spoken in Catalonia since the Middle Ages, but was progressively dismissed with the Spanish domination of the eighteenth century and eventually banned during Franco's dictatorship (1939–1975), when Spanish became the only legal language. Both Catalan and Spanish are romantic languages, but Catalan is closer to English compared with Spanish.<sup>2</sup> The return of democracy in the late 1970s provided the political framework for a reintroduction of Catalan, also as a means of promoting the political autonomy of the region and its internal social cohesion. Increasing Catalan proficiency in the population became an important target for the new regional government, because a substantial share of the population was non-Catalan after the internal migration flows of the 1950s and 1960s, especially in the more industrialised areas of the region. More in general, after forty years of prohibition, proficiency in written Catalan was problematic among prime age and young individuals, irrespective of regional origins. Catalan became a compulsory subject in schools immediately after the end of dictatorship, but it was only with the 1983 law that a broader reform of the educational system towards bilingualism was put in place.

The Language Normalization Act (LNA) of 1983 established Catalan as a medium of instruction in primary and compulsory secondary schools alongside Spanish, making the education system effectively bilingual. This language-in-education reform affected virtually all students enrolled at the time of implementation, irrespective of their birth cohort or school of enrolment. The comprehensiveness of the reform rules out any possibility of students' selection into bilingual schools, which naturally occurs whenever two systems (monolingual and bilingual) co-exist and parents are free to choose the language of education of their children (as, for example, in the Basque Country, see Aspachs-Bracons et al., 2008).

---

<sup>2</sup> According to the novel measure of linguistic distance “Automated Similarity Judgment Program Database” (ASJP, described in Ispording and Otten, 2014), the distance between English and Catalan (Spanish) is 86.26 (93.34) on a (roughly) 1-100 scale. Catalan and Spanish are sufficiently different to generate the so-called “cognitive advantage” for Catalan-Spanish bilinguals (compared to Spanish monolinguals) in non-linguistic cognitive tasks (Garbin et al., 2010).

From a theoretical point of view, there are two main channels through which exposure to bilingualism at school might be rewarded in the labor market. First, bilingualism can favour the development of cognitive skills, in this way increasing the stock of human capital accumulated per year of education – resulting in higher wage returns. Related to this, it may increase the quality of education if more skilled teachers are needed for bilingual teaching, inducing higher returns to schooling. Second, bilingual knowledge might increase search efficiency and increase the rate of arrival of job offers in a local labor market in which Catalan-speaking employers represent a relevant share of overall labor demand, leading to better job matches and higher wages. Effects through better search might also derive from better access to Catalan social networks, which could be particularly relevant for individuals with non-Catalan origins that increasingly perceived themselves as Catalans after the reform (Aspachs-Bracons et al., 2008; Clots-Figueras and Masella, 2013).

Our analytical approach is in the spirit of Angrist and Lavy (1997), who studied the wage effects of the 1983 ‘Arabization’ policy in Morocco, which replaced French with Arabic as the language of instruction of post-primary education for all birth cohorts beginning post-primary school in that year or later. They develop a DiD estimator across levels of education and birth cohorts and conclude that returns to schooling were substantially lowered for individuals exposed to the reform, a likely consequence of the loss of skills in French induced by ‘Arabization’. A similar strategy is used by Angrist et al. (2008), who evaluated the effect of English-intensive instruction on English skills in Puerto Rico, exploiting the reform that in 1949 substituted English with Spanish as the language of instruction in all grades irrespective of year of birth, finding no effect on English skills. In both cases, the policy effect is captured by the differential return to each additional year of schooling (in terms of either wages or language skills) between individuals that were differently exposed to the reforms.

The Catalan reform differs in contents from the ones studied by Angrist and Lavy (1997) and Angrist et al. (2008) because it introduced a second official language of instruction rather than

switching it from one to another. Still, their methodological framework represents the starting point of our empirical analysis. In particular, we implement a DiD estimator with partial exposure as in Angrist et al. (2008), and perform several falsification exercises to address the validity of the underlying common trend assumption.

One potential shortcoming of DiD estimators that exploit differential exposure to policy changes across birth cohorts and schooling levels stems from the endogeneity of schooling decisions. More specifically, this is an issue if the policy changed the unobservables driving selection into education. To dispel these doubts about the validity of identification, we repeat the analysis relying on policy exposure only during compulsory education (which is not chosen), a strategy proposed by Clots-Figueras and Masella (2013) who analysed the effects of the Catalan reform on national identity.

We use data from the Survey on the Living Conditions and Habits of the Catalan Population for 2006 and 2011. We find that one year of bilingual education generates on average an extra earnings return of 1.4 percent, on top of the yearly return to education of 6.5 percent. Conversely, we do not find any significant effect of the reform on employment, working hours or occupation.

We subject the estimated wage effects to a number of sensitivity checks. We control for selection into employment and for differences in life-cycle earnings trajectories across birth cohorts. To address the validity of the common trend assumption, we perform placebo analyses using both individuals belonging to the same birth cohorts from different non-bilingual Spanish regions and Catalan individuals belonging to older birth cohorts not affected by the reform. Evidence from all of these sensitivity checks points towards the robustness of our findings. We alternatively implement the estimator à la Clots-Figueras and Masella (2013) based only on variation in exposure at compulsory school. The results are in line with those obtained exploiting variation in exposure throughout the whole schooling process, indicating that our evidence is not driven by endogenous selection into higher schooling levels.

In the last part of the paper we explore heterogeneities in the estimated impacts of the reform and find that it was effective mostly for individuals of non-Catalan background (in terms of either origins or language) and from low parental background. While the former effect is consistent with growing language proficiency, the latter may reflect a broad increase in educational quality brought about by the reform. Overall, these results suggest that the language-in-education reform achieved its goal of levelling the playing field in a bilingual society.

The rest of the paper proceeds as follows. In Section 2 we review the literature on language reforms, bilingual education and the 1983 reform of the Catalan schooling system. Section 3 provides a detailed account of the institutional setting, which is key in the construction of the treatment variable for the analysis. Section 4 describes the data, while Section 5 lays out the empirical framework. Section 6 reports the main results, together with a number of sensitivity checks, while Section 7 investigates the heterogeneous effects of the reform. We conclude in Section 8.

## **2. Related literature**

There exists a growing literature in economics that studies bilingualism and its impacts on socio-economic outcomes. Chin et al. (2013) analyse the impact of Bilingual Education Programmes for primary schools in Texas, which affected all school districts where the enrolment of Limited English Proficiency (LEP) Spanish-speaking students was above a certain threshold. Exploiting this feature of the policy in a quasi-experimental setting, they find no effect of bilingual schooling on test scores among LEP students, while some effect is found among non-LEP students. Anghel et al. (2016) evaluate the effects of a bilingual education policy on academic performance in some primary schools in the Madrid region. This is a *language-in-education* policy, i.e. not only was the language of instruction changed by the policy, but also other aspects of the educational process were affected such as training and wage incentives for teachers. The policy introduced English as medium of instruction in subjects such as Science, History and Geography. Results point to

negative effects on test score performance of low-background students, but only in the subject that was taught in English. A similar negative effect has been found by Ivles and King (2014) for a *language-of-instruction* reform (i.e. limited to the medium of instruction) that took place in Russian secondary schools in Latvia.

There is also a large literature from cognitive psychology, which analyses the effects of bilingualisms on different aspects of language learning and cognition (see Bialystok et al., 2009 for an overview). The evidence from this field indicates that, in terms of verbal performance, bilingual and monolingual children experience the same timing for language learning, but the former tend to produce a smaller vocabulary in each language. The gap in verbal tasks is reduced with the level of language proficiency, and is more likely to reflect different amounts of language exposure rather than differences in the ability to learn a given language. Indeed, bilingual individuals are advantaged in learning additional (foreign) languages during childhood and adulthood. Moreover, and most importantly, exposure to bilingualism – especially early in life – has a positive impact on cognitive development, mostly working through improved executive functions and their corresponding brain structures (Costa and Sebastián-Gallés, 2014). Bilingual individuals outperform their monolingual counterparts in several aspects of cognitive control, showing greater ability to switch between tasks, maintain attention, and to inhibit distracting information. Additionally, there is growing evidence pointing to additional beneficial effects of bilingualisms in later life, as it serves to protect against some aspects of age-related cognitive loss. This suggests that there might be some longer-term benefits of exposure to bilingualism at school, which could remain hidden when focusing only on early outcomes in schools. However, the analysis of longer-term outcome measures, such as labor market performance, may tell a different story. To the best of our knowledge, there is no evidence on the labor market effects of bilingual education, a gap we try to fill with this paper.

While the labor market effects of bilingual education are still unexplored, there exists evidence on the effect of policies that changed the language used in schools from one to another. In

many cases, these changes occurred in former colonies switching from colonial to own language. This literature has an important methodological content, which informs the empirical strategy of our paper. A seminal contribution is provided by Angrist and Lavy (1997), who analyse a 1983 language-of-instruction reform in Morocco. The reform switched the language of instruction at secondary schools from French to Arabic and affected all incoming students beginning secondary education in 1983, while leaving incumbent students unaffected. This variation is the basis of a DiD estimator of the wage effect of the policy using years of education in the new regime as the running variable, which compares wages across educational levels for students enrolled before and after the policy took place. Their results point towards a negative effect on earnings of changing the language of instruction from French to Arabic. Using the reform as an instrumental variable, they estimate significant positive earnings effects of French proficiency.

The effect of a switch in the language of instruction from English to Spanish in Puerto Rican schools is the topic of Angrist et al. (2008), who consider English proficiency as the main outcome of interest. Differently from the Moroccan case, the reform in Puerto Rico affected all individuals in school at the time of the reform, and not only incoming students, which generates partial exposure to treatment. They apply a logic similar to Angrist and Lavy (1997) and use years of education in the new regime to derive a DiD estimator, but allow for partial exposure to the treatment exploiting variation across birth cohorts and educational levels, which is also the strategy that we adopt in the present paper. Focusing on English proficiency as the outcome of interest, they find a negative effect of the language reform, which, however, is not robust to allowance for flexible time trends by birth cohorts and educational levels.

Lleras-Muney and Shertzer (2015) estimate the effect of English-only schooling for migrants in the US within the so-called Americanization Process (1910–1930), exploiting variation in the timing of schooling legislations across states. They consider several short- and long-term outcomes, such as literacy, English proficiency schooling, employment, income and social assimilation. Their findings indicate only a modest positive effect of English-only schooling on literacy among low-



background foreign-born children, but the reform seemed to be ineffective in terms of labor market and social integration outcomes.

Chakraborty and Bakshi (2016) study the wage effect of the 1983 suppression of English instruction as a subject in primary schools of Indian State of West Bengala. Given the absence of information on actual policy exposure, the authors rely on measures of potential exposure to public schools (the only ones affected by the policy) at the district/region level at the time of implementation and a DiD estimator. Results indicate the existence of a significant wage penalty for individuals who were potentially exposed to the abolition of English as a subject.

While there is no study of the labor market effects of bilingual education there exists evidence on civic outcomes, which have been investigated focusing on the same Catalan reform that we exploit in this paper for labor market outcomes. Aspachs et al. (2008) consider the Catalan reform in conjunction with a similar reform implemented in the Basque Country in the same year (1983). These reforms affected all individuals in school at the time of the reform and the authors exploit a research design that is similar to that of Angrist et al. (2008), i.e. a DiD across schooling levels and years of birth with partial exposure to treatment. They find an impact only in the case of Catalonia, because in the Basque Country two different language regimes (Basque-Spanish and Spanish-only) were allowed to coexist and parents were free to choose the language of instruction for their children, while in Catalonia the bilingual regime was compulsory in all primary and secondary schools. Clots-Figueras and Masella (2013) analyse the effect of the Catalan reform on feelings of Catalan identity and use only years of exposure at compulsory school to cope with the fact that total years of education in the new regime partly reflect possibly endogenous choices of educational attainment. In this way, they exploit only variation across cohorts before and after the reform and not across educational levels, as years of compulsory education in the new regime vary only across birth cohorts and not within them. They confirm the findings of Aspach et al. (2008) that substituting Spanish-only with Spanish-Catalan education increased feelings of Catalan identity.

### **3. Institutional background**

Catalan belongs to the family of romance languages (together with French, Italian, Occitan, Portuguese and Spanish) and has been the local language of the Spanish region of Catalonia since the early eleventh century. Starting with the War of Spanish Succession (1701–1714) and the subsequent incorporation of Catalonia within the Spanish Crown, the use of Catalan was progressively limited to domestic use and the language lost much of its social relevance. During the dictatorial regime established in the 1930s, Catalan was banned in the public milieu, its private use was prosecuted and Spanish became the only official language. After Franco's death in 1975 the country went through the democratic transition. The decentralization process that took place with the Democratic Constitution of 1978 recognized the co-officiality of Spanish and local languages in bilingual regions, which were allowed to stimulate the public and private use of their own languages. The Catalan government pursued the use of Catalan through language policies, culminating with the Language Normalization Act (LNA) of 1983. This law represented a sharp change of language policy that we detail later in this section, and enjoyed full support from all political parties and society at large. It was conceived to provide the institutional and legal basis for a complete transition towards a bilingual society, favouring the return of Catalan as the co-official language of the region.

The socio-demographic landscape of Catalonia after the Franco regime represented the main challenge in pursuing the aims of the reform. Due to mass migration from the Spanish-speaking areas of the country towards Catalonia since the 1950s, a substantial share of Catalan residents (i.e. migrants and their offspring) were Spanish native speakers, with limited knowledge of Catalan, particularly in the periphery of the city of Barcelona, where most migrants were located. Instead, in Catalan-speaking families Catalan represented the native language even for new generations born during the dictatorship. This means that individuals of Catalan origin were fluent, at least orally, in their native language. It was against this background of linguistic segmentation that the local

government used language-in-education policies as the main instrument of ‘language normalization’ for the new generations.

Immediately after the 1978 Constitution came into effect, Catalan language became a compulsory subject (for at least three hours per week) in non-tertiary education. With the LNA of 1983, Catalan became a medium of instruction in primary and compulsory secondary schools, alongside Spanish, making the education system effectively bilingual. The reform established that, by the end of compulsory school, all pupils must have achieved complete proficiency in the four basic competences (understanding, speaking, reading and writing) in both Catalan and Spanish. Under the new system, the two official languages were taught as subjects in a similar number of hours. Catalan had to be used as language of instruction in at least one area of study (over eight, including languages) from grade 3 to 5, and in two areas from grade 6, while Spanish had to be used as language of instruction in at least one area throughout the course of studies. Beyond these minimum mandatory requirements, the exact amount of teaching in each of the two co-official languages was determined by students’ composition in terms of language background and by teachers’ language skills – not all teachers were initially proficient in Catalan.

The 1983 reform introduced language immersion programmes in primary and pre-primary education (Arnau and Vila, 2013; Artigal, 1997). These were targeted to schools whose students predominantly (more than 70 percent) came from Spanish-speaking families, which tended to be settled in neighborhoods where Catalan had very little presence. Schools in immersion programmes used Catalan as the only language of instruction during the first years of education and followed a specific methodology to stimulate second language (L2, i.e. Catalan) acquisition. Spanish was introduced as additional language of instruction only at a later stage (normally grade 3).

The LNA also regulated the language regime in post-compulsory secondary school. However, the introduction of Catalan as medium of instruction in upper-secondary education was less intense and more gradual compared to compulsory education. The reform established that upper-secondary schools, besides mandatory language courses in Spanish and Catalan, had to employ Catalan as

medium of instruction in at least two subjects. In practice, however, the choice between Catalan and Spanish was left to teachers and taken on the basis of the linguistic composition in the classroom.

The language regime of universities was not explicitly regulated by the LNA, as universities were already endowed with special autonomy when the language policy was implemented. The law only established the right to use any of the two official languages and required Catalan universities to offer Catalan courses to students and teachers with limited knowledge of Catalan.<sup>3</sup>

Although with varying intensity, all the levels of the educational system became bilingual in the year of the reform. This means that while cohorts starting primary school after 1983 were fully affected by bilingual education, cohorts already in school in 1983 were differentially affected by bilingual schooling depending upon the grade attended in that year. This creates a partial exposure framework similar to the one exploited in Angrist et al. (2008) for the case of Puerto Rico. We come back in Section 4 on the practical implementation of the estimator based on partial exposure.

The LNA promoted bilingualism among teaching staff (Arenas, 1990). Several ‘specialized’ teachers were hired to guarantee the minimum level of staff proficiency required by the law. These new teachers were allocated on the basis of the linguistic composition of the school in terms of both teachers and students. At the same time, less Catalan-proficient teachers were given the opportunity to develop adequate skills through special courses promoted by the SEDEC (*Servei d’Ensenyament del Català*, Catalan Teaching Service) and the Institutes of Education of the three Catalan public universities of the time. While in 1978–79 only 52 percent of teachers in Catalonia were able to speak and write in Catalan, this percentage rose to 87 percent in 1986–87 (Arenas and Muset, 2007).

#### **4. Data and descriptive statistics**

---

<sup>3</sup> According to anecdotal evidence, the use of Catalan in tertiary education was already quite widespread during the early 1980s as almost half of university courses were taught in Catalan.

We use data from the Survey on the Living Conditions and Habits of the Catalan Population (*‘Enquesta de Condicions de Vida i Hàbits de la Població’*, ECVHP), waves 2006 and 2011. The ECVHP provides information on labor market outcomes (net monthly earnings, hours of work and occupation), socio-demographic characteristics (province of birth of individuals and their parents, and parental education), educational attainment, language use (Catalan, Spanish, both Catalan and Spanish, other languages) and self-reported Catalan proficiency (both oral and written). ECVHP data are representative of the Catalan population and the only one containing the information needed for the analysis of this paper.

Our sample consists of individuals who completed education, were born either in Catalonia or in other Spanish regions but migrated to Catalonia when they were aged at most 6, thus excluding individuals who were, at least in part, educated outside Catalonia. We begin by selecting individuals born between 1965 and 1977 because they were aged 6 to 18 and thence potentially attending either primary or secondary school when the reform came into effect in 1983. We complement these cohorts by extending the birth year limit backward and forward, including in the sample individuals that were subject to the same compulsory schooling law as those born between 1965 and 1977. The first and last years of birth satisfying this requirement are 1961 and 1982, which therefore define the birth year limits for inclusion in the estimating sample.

Apart from the year of birth, the main variables in our analysis are years of schooling, years of bilingual schooling and earnings. Years of schooling are imputed from the very detailed information about completed levels of education available in the database. We impute years of exposure to bilingual education on the basis of years of schooling and year of birth. The exact imputation of bilingual education is reported in Table 1. Individuals born in 1977 or after received all their schooling under the bilingual system and for them exposure coincides with years of schooling, while those born between 1966 and 1976 were partially exposed to the reform and individuals born before 1966 were not affected by reform. Bilingualism was not explicitly regulated at college and whether one should consider college years as years of exposure to bilingual education is inherently

undetermined. To get around the indeterminacy, we consider years of college education as years of exposure only for cohorts that have not yet started college by 1983 (cohorts 1966 and onwards). Our treatment variable, therefore, would deliver a lower bound of the true effect if individuals that had already completed secondary school by 1983 (cohorts born in 1965 or earlier) actually received some bilingual education during college years. In summary, the treatment variable illustrated in Table 1 is obtained by variation in exposure to bilingual education across birth cohorts and educational levels, generating a DiD framework similar to the one of Angrist et al. (2008).

We include in the analysis all individuals who are regularly employed at the time of the survey.<sup>4</sup> As the effects of interest are identified by earnings variation across years of education and birth years, our model will control for year of birth fixed effects. In order to avoid confounding the latter with life-cycle effects, we use repeated cross-sections (2006 and 2011) and control for quadratic experience profiles in the baseline specification. Clearly, the business cycle was very different between 2006 and 2011, and we will provide sensitivity checks on whether changing selection into employment between the two years affects our results.

The ECVHP collects information about net monthly earnings in both waves, but in 2006 information is reported in brackets rather than continuously. We harmonise the earnings information across waves by deflating 2011 earnings to 2006 prices and by imputing the mid-point of reported earnings brackets in 2006.<sup>5</sup> We analyse the resulting variable by means of linear regression models. Detailed descriptive information about net monthly earnings is displayed in Table A1 in the Appendix, while remaining descriptive statistics on the estimating sample are reported in Table A2 of the Appendix.

In order to run falsification exercises on the validity of the common trend assumption underlying our DiD strategy, we complement the main ECVHP sample with two auxiliary sources of information that we use to conduct placebo analyses. First, to test for the existence of

---

<sup>4</sup> We run sensitivity checks excluding from the analysis self-employed and entrepreneurs, finding that results were robust to this sample selection.

<sup>5</sup> Alternatively, we discretized 2011 earnings into the same intervals that categorize earnings in 2006, and applied interval regressions. Results were pretty much unaffected by changing the harmonization procedure.

contemporaneous cohort-education trends in earnings, we use data from the Spanish component of the European Survey on Income and Living Conditions, (EU-SILC, waves 2006 and 2011) and retain observations of individuals born in Spain and residing in non-bilingual regions. Second, we use older cohorts from the ECVHP, namely individuals born between 1945 and 1960 who had already completed education before the introduction of the reform and, therefore, were not affected by bilingual education. This ‘placebo cohort’ enables considering the presence of pre-existing trends in the Catalan labor market. Descriptive statistics on the placebo samples are reported in Appendix Table A3.

## 5. Empirical strategy

We identify the labor market returns to bilingual education by exploiting variation in exposure across years of birth and years of schooling. As shown in Table 1, the LNA reform generates full, partial or null exposure to the new language regime at school depending on the year of birth and the level of education, inducing differential treatment intensity. The treatment is described by the following variable:

$$e_i = \begin{cases} s_i & \text{if } b_i \geq 1977 \\ \max\{0; s_i - (1977 - b_i)\} & \text{if } 1965 < b_i < 1977 \\ 0 & \text{if } b_i \leq 1965 \end{cases} \quad (1)$$

where  $e_i$  are years of *potential* exposure of person  $i$ ,  $s_i$  are years of schooling, and  $b_i$  is the year of birth. The treatment variable  $e_i$  may differ from *actual* exposure because of grade retention and differential intensity in the use of Catalan as medium of instruction across schools. Therefore, our treatment variable can be interpreted as capturing an Intention to Treat (ITT) effect.

### 5.1 Baseline model

Following Angrist et al. (2008), our baseline specification for estimating the effect of the reform on labor market outcomes is:

$$w_i = \alpha + \beta'x_i + \gamma s_i + \delta e_i + \theta_b + \varepsilon_i \quad (2)$$

where  $w_i$  denotes labor market outcomes,  $x_i$  is a vector of controls (year, gender, potential experience and its square) and  $\theta_b$  is a year of birth fixed effect. Our main outcome of interest is monthly net earnings from employment, but we also consider other outcomes such as employment, hours and occupation.<sup>6</sup> The coefficient  $\delta$  is the additional return per year of education generated by the reform, and parameterises a DiD estimator – across birth years and schooling levels – that allows for partial exposure. Equation (2) provides consistent estimates of the effects of the reform as long as any other trends in labor market outcomes (besides the ones induced by the reform) were common across birth years and schooling levels.<sup>7</sup>

## 5.2 Accounting for potential confounders

Life-cycle earnings trends may confound the effects of the reform because treated and non-treated cohorts are observed at different stages of their labor market trajectories, a form of life-cycle bias. Life-cycle effects are controlled for in our baseline model through year of birth fixed effects. Moreover, we exploit the availability of two cross-sections and further saturate the specification with age dummies, omitting two of these –rather than one—from the regression to avoid collinearity. The use of two cross-sections comes with a cost, since the second one (2011) refers to a period of slack in the Spanish and Catalan economies, which has substantially increased unemployment, especially among the young. The potential change in selection into employment might affect the estimates; to check if this is an issue, as a sensitivity we estimate a joint model for earnings and employment and correct for selection.<sup>8</sup>

---

<sup>6</sup> Besides providing evidence on labour market outcomes other than earnings, looking at the effect of the reform on working hours is particularly relevant in our case because 2006 earnings are reported in brackets, preventing us from deriving hourly earnings measures.

<sup>7</sup> Equation (2) imposes linearity of the return to bilingual schooling. Alternatively, we allowed for non-linearities through a specification with dummies for years of exposure or through a spline function. Results from these models do not indicate any significant variation of the effect by years of exposure.

<sup>8</sup> We estimate an endogenous selection model using as exclusion restriction the unemployment rate at the province level when the individual was 16, the legal minimum working age. We also show that the estimated effects are not very different if we estimate the earnings equation separately by year (see Table A4 in the Appendix).



An additional confounding factor that we take into account is the contemporaneous expansion of education that took place in Spain over the period in which the LNA reform was implemented. To account for it, following Angrist et al. (2008), we augment the specification with a measure of the educational cumulative density function (CDF) and its square. For each survey respondent, we use data from the Spanish Census of 2011 and define the educational CDF as the fraction of individuals with a lower level of education who were born in the same year of birth. The CDF would capture the changing selectivity into schooling levels due to education expansion, which might have reduced the average ability of individuals with higher schooling attainments.

### 5.3 Placebo experiments

We pay special attention to the existence of spurious relationships between the measure of exposure and the outcome due to education-cohort specific trends. The main assumption upon which validity of the identification strategy rests is the absence of differential trends in labor market outcomes across years of birth and years of schooling. We address these concerns by means of two different falsification exercises. First, we run a placebo analysis that uses data on cohorts of individuals born in the same years as those in the baseline estimation sample (1961–1982), but not residing in Catalonia or in any other bilingual region of Spain. We use the Spanish component of EU-SILC (waves 2006 and 2011) and impute reform exposure ‘as if’ it had been implemented also in those regions.

Second, we go back to ECHVP data and consider cohorts of Catalans born between 1945 and 1965 (which were never actually exposed to bilingual education), to whom we impute placebo exposure to bilingual schooling ( $\pi^*_i$ ), pretending that the reform was rolled out in year  $t^* \in [1963, 1970]$  rather than in 1983, thus generating partial placebo exposure for birth cohorts from 1946 to  $(t^* - 7)$  and full placebo exposure for birth cohorts from  $(t^* - 6)$  to 1965:

$$\pi^*_i = \begin{cases} s_i & \text{if } (t^* - 6) \leq b_i < 1966 \\ \max\{0; s_i - ((t^* - 6) - b_i)\} & \text{if } 1946 \leq b_i < (t^* - 6) \\ 0 & \text{if } b_i < 1946 \end{cases} \quad (3)$$

We estimate the baseline model on these cohorts using  $\pi^*_i$  as the treatment variable. Since these cohorts were not exposed to the true reform, obtaining a significant effect of the placebo variable would suggest that our treatment variable (1) is capturing pre-existing education-cohort specific trends in earnings rather than the impact of bilingual schooling.

#### 5.4 Selection into education

There is a final major concern with the derivation of the treatment variable, namely that it is constructed on the basis of completed schooling, which is a choice variable.<sup>9</sup> This would not be an issue if the unobservables driving selection into education did not change after the implementation of the LNA reform. Conversely, if before the reform some individuals with intrinsically high ability were not enrolling into post-compulsory education because school programmes were taught only in Spanish, and the introduction of bilingualism induced their younger counterparts to continue education, then we would be erroneously attributing to the reform what in effect is a change in the unobservables.

In order to gauge the relevance of this potential issue, we apply the strategy of Clots-Figueras and Masella (2013) and define the treatment as years of potential exposure at *compulsory* education:

$$\tilde{e}_i = \begin{cases} 8 & \text{if } b_i \geq 1977 \\ \max\{0; 8 - (1977 - b_i)\} & \text{if } 1970 < b_i < 1977 \\ 0 & \text{if } b_i \leq 1970 \end{cases} \quad (4)$$

As suggested by Clots-Figueras and Masella (2013), using compulsory education means that the corresponding exposure variable only depends on year of birth, which would free from any selection bias the estimated effect of exposure to bilingual schooling. Following their intuition, as

---

<sup>9</sup> Additional threats to identification may come from selective migration in-and-out of Catalonia because of the reform. Using data from the Census of 1991, we found that reform exposure does not predict migration to or out of Catalonia among individuals born between 1961 and 1982, which rules out the possibility of confounding effects due to selective migration.

an alternative to our baseline specification we estimate the model substituting total years of exposure with years of exposure at compulsory school. Since it varies only by year of birth (and not also by years of schooling), year of birth fixed effects are not identified in this model, and we group them into birth cohort fixed effects  $\theta_c$ , yielding the following equation<sup>10</sup>

$$w_i = \alpha + \beta' x_i + \gamma s_i + \lambda \tilde{e}_i + \theta_c + \varepsilon_i \quad (5)$$

Therefore, if changing selection into schooling does not represent the main driver of the effect of language exposure on earnings, the results from the baseline model (2), which exploits the potentially endogenous variation of  $e_i$ , would be similar to those obtained from Equation (5), which exploits exogenous variation of  $\tilde{e}_i$  across birth cohorts. On the contrary, finding lower or no effects of compulsory language exposure would suggest that changes in unobserved individual heterogeneity, rather than language exposure, generate the effects that we instead attribute to the policy change. We also run a falsification exercise for the compulsory exposure variable by imputing  $\tilde{e}_i$  to individuals from other non-bilingual Spanish regions (again using the EU-SILC data), which would capture potential contemporaneous trends across the cohorts.<sup>11</sup>

## 6. Results

### 6.1 Bilingual schooling and labor market outcomes

We begin the presentation of results in Table 2 reporting selected parameter estimates for the effects of the reform on various labor market outcomes.<sup>12</sup> Column (1) reports results of the earnings model. The DiD estimates show a sizeable positive effect of bilingual education on earnings. One additional year of bilingual education increases earnings by 1.4 percent. This incremental return to

---

<sup>10</sup> Birth cohorts are defined as birth-year ranges formed on the basis of reform exposure: 1961–1965 (Spanish-only schooling), 1966–1969 (exposure only after compulsory education), 1970–1976 (partial exposure also during compulsory schooling) and 1977–1982 (full exposure).

<sup>11</sup> We are unable to use the never-treated cohort of Catalans to falsify the effect of exposure during compulsory schooling, since individuals belonging to the 1945–1960 cohort were not subject to the same compulsory schooling rules. Specifically, they were not affected by the “*Ley General de Educación (LGE)*”, approved in 1970 and enforced in 1974, which increased the duration of compulsory education to eight years.

<sup>12</sup> In this and the following tables we report only key parameter estimates; the full set of estimated coefficients of the earnings regression are reported in Table A4 in the Appendix. In the same table, we also report separate estimates by wave, as well as estimates on hourly wages (only available in 2011), with and without the exposure variable.

schooling comes on top of a yearly return of 6.5 percent and represents a proportional increase of about 20 percent of returns to schooling due to exposure to bilingual education.

In Columns (2) to (4) of Table 2 we apply the same specification of Column (1) to labor market outcomes other than earnings. Column (2) focuses on employment. In Column (3) we consider whether the effect observed on monthly earnings could stem from differences in hours of work between treated and non-treated individuals and use weekly hours as the relevant outcome. Such exercise is important because we are unable to construct an hourly wage measure consistently in both waves. Finally, we consider whether the earnings effect could reflect better access to more prestigious occupations and estimate a linear probability model for being in highly skilled white-collar jobs (Column (4)). For none of these additional labor market outcomes do we find a significant effect of the reform. Finding an effect on monthly earnings and not on hours indicates that the uncovered impact of bilingual education is operating through increased earning capacity rather than through increased working time within the month. The detected earnings effect might also transmit to reservation wages, which would explain the null employment effect as the net effect of increased wage offers and increased reservation wages for the treated. Also, the effect is not a reflection of compositional changes in the distribution of broad occupational groups, as shown by the null effect on the probability of holding white collar jobs. In the light of this evidence, we focus on earnings in the remainder of this section.

## *6.2 Sensitivity checks*

Table 3 reports the results of sensitivity checks on the earnings effects. Column (1) reports the baseline estimates. In Column (2) we augment the baseline model to correct for selection into employment, using unemployment rate in the province of birth at the time the individual was 16 plus all the regressors of the earnings equation to model individual employment probabilities. This is relevant since our data come from two pooled cross-sections from different phases of the business cycle. In Column (3) we augment the baseline specification (that already includes a quadratic trend

in experience) with age dummies. Finally, in Column (3), to control for potential concurrent compositional changes of the labor force by ability that may come from the contemporaneous educational expansions, we include the educational CDF (and its square) as additional control in the earnings equation.

Accounting for selection into employment has an impact on the baseline return to education, but it only marginally affects the estimate of the treatment effect. The overall picture remains the same when we include age dummies in the model. When controlling for the educational CDF in Column (4), the estimate of the underlying return to education decreases and loses precision while the estimated treatment effect increases by approximately a third. The result on returns to education is consistent with a supply side effect, while the increased effect of exposure suggests that bilingual schooling may shelter treated cohorts from the increased competition induced by the entry into the labor market of a more educated workforce.<sup>13</sup> Overall, these sensitivity checks point to a substantial robustness of the estimated earnings effect of exposure to bilingual schooling.<sup>14</sup>

### *6.3 Falsification analysis*

In this section we run several falsification experiments that are aimed at ruling out the existence of spurious relationships between the treatment variable and earnings. The main assumption underlying the consistency of our DiD estimator is that any (education/cohort-specific) earnings trend unrelated to the reform is common between treated and control groups. To assess its validity, in Table 4 we report results of placebo regressions and contrast the resulting estimates with the baseline results shown in Column (1).

---

<sup>13</sup> The estimated coefficients on the educational CDF and its square are very precisely estimated and indicate an increasing and concave relationship between educational expansion and earnings.

<sup>14</sup> In Table A4, we report estimates by wave. By construction, these estimates cannot fully control for life-cycle effects, but anyway point towards the robustness of the results. In 2006, the baseline return to education was 5.9 percent and the treatment increased it by 1.1 percent, while in 2011 the baseline return was 6.5 percent and the effect of the reform was 1.1 percent. In both waves, the effect of the reform is less precisely estimated than in the baseline model due to smaller sample sizes, and we cannot reject the null hypothesis that returns by wave are equal to 1.4 as in the baseline analysis of Table 2. Larger returns to education in 2011 suggest a protective effect of education in the face of the Great Recession.

First, we check for the existence of contemporaneous spurious relationships by using EU-SILC data and imputing placebo exposure to individuals born in the same cohorts as our main sample (1961–1982) but in other non-bilingual Spanish regions, who, therefore, were not exposed to the reform. The results are displayed in Column (2). The coefficient on placebo exposure is virtually zero. This evidence suggests that our exposure variable is not capturing contemporaneous trends in earnings that apply to other Spanish regions.

Second, we consider the presence of a pre-existing trend within the Catalan labor market. We assign placebo exposure ( $\pi^*_i$ ) to older cohorts from the EHCVP data (namely individuals born 1945–1965) that have not been actually treated by bilingual education. For these cohorts we mimic a set of reforms assuming they were implemented in each of the years 1963 to 1970, and run a placebo regression for each of these hypothetical reforms. The results obtained from this exercise are displayed in Columns (3) to (10) of Table 4. The underlying return to education is larger for these older cohorts compared to the baseline estimating sample, reflecting underlying trends in the supply of educated workers reducing returns among younger and more educated cohorts, while all the placebo effects are estimated close to zero and statistically insignificant. Taken together, the evidence from Table 4 rules out any threat to identification coming from pre-dating earnings trends across cohorts and schooling levels or concurrent trends outside Catalonia.<sup>15</sup>

#### *6.4 Endogenous selection into education*

Another fundamental concern with our DiD estimator is selection into schooling, which is an issue if the estimated earnings effects reflect changing unobserved ability of more educated individuals before and after the reform, rather than the true effect of bilingual schooling. We dispel these doubts in Table 5, which reports the effects of reform exposure at compulsory schooling following the approach developed by Clots-Figueras and Masella (2013) for analysing the effects of the 1983

---

<sup>15</sup> Combining placebo and baseline analyses we also derived a triple-difference estimator for the effects of exposure to bilingual education, finding results that are entirely in line with those discussed in the text.

reform on feelings of Catalan identity. They used the *potential* number of years of compulsory education (the variable  $\tilde{\epsilon}_i$  in Section 5) to capture the exogenous amount of exposure to the reform.

Using the same measure as Clots-Figueras and Masella (2013) delivers an effect of potential compulsory exposure that is larger than the baseline (2.1 vs 1.4 percent), although we cannot reject the null hypothesis that the estimated effect of  $\tilde{\epsilon}_i$  is different from 1.4 percent. This result shows that changing selection into schooling does not affect our results, because potential exposure to bilingualism during compulsory education is not driven by individual choices and related unobservable characteristics. Nevertheless, there is a final important issue that deserves additional attention. By only depending on year of birth, the potential exposure variable rules out selection on unobservables, but, similarly to the full exposure variable, could be biased by differential earnings trends across cohorts. We address the relevance of these perils with a falsification analysis based on the contemporaneous cohorts of individuals from non-bilingual Spanish regions, which we assume to receive pseudo years of compulsory exposure. The results are reported in Column (3) of Table 6. The estimated treatment effect obtained from these cohorts is close to zero and not significant.

## **7. Heterogeneous effects**

As discussed in Section 3, with the migrations of the 1950s and the 1960s many low-skilled workers arrived in Catalonia from the rest of Spain; raising Catalan proficiency in the population became an important target for the new regional government that came into power in the late 1970s, also as a mean of fostering equal opportunities and preserve social cohesion. To the extent that the reform managed to achieve its goals of levelling the playing field, we would expect a reduction of earnings differences among individuals with different origins, with different languages or with a different socio-economic background, but equally exposed to bilingual education. We now consider these aspects by looking at the heterogeneous effects of the reform.

Our data allow us to take heterogeneous effects into account, thanks to variables recording information about parental regional origins, language use and parental education. In particular, we

observe whether respondents' parents were born in Spanish regions other than Catalonia, and we define as of non-Catalan origins those individuals who report both parents being born outside Catalonia. Moreover respondents report their usual language (either Spanish or Catalan). We estimate models including indicators of either regional origins or language use, and their interactions with years of education and years of exposure to bilingual education. Results are reported in Table 6. Column (1) reports the baseline estimates, while columns (2) and (3) report estimated heterogeneous effects. Patterns of heterogeneous effects are similar for both regional origins and language use. Returns to education in the absence of bilingual exposure are lower for individuals of non-Catalan origins and those reporting Spanish as the main language used, by 1.4 and 2.5 percentage points (p.p.), respectively. Exposure to bilingual education reduces the gap between these two groups and those of Catalan origins or those using Catalan, by about half of a p.p.. The estimated heterogeneous effects are marginally significant. In both cases, it seems that the reform achieved at least partially its egalitarian goals, because the pre-existing gap in returns to education is reduced by about a third by exposure to bilingual education.

The observed heterogeneous effects are consistent with a positive effect of bilingual education on the language proficiency of non-Catalan originals.<sup>16</sup> In principle, however, the effects of the reform may not only have been confined to non-Catalans, but also spread to the rest of society. This could have happened if the reform affected the overall quality of the educational system, for example through the recruitment of more skilled teachers capable of teaching bilingual programs, with an effect on the overall ability of the teaching body. To the extent that school quality matters more for children with low parental background as it compensates for the lack of endowment, we would expect disadvantaged children to benefit the most from the reform irrespective of their origins, and the reform to level the playing field between individuals coming from different backgrounds. The estimates reported in Table 6, Column 4, are obtained defining low background as both parents having completed at most compulsory education, and indicate that this was actually

---

<sup>16</sup> Indeed, Caminal and Di Paolo (2017) show that the reform raised language proficiency for native Spanish speakers.



the case. One year of education is less worth for low background individuals compared with more advantaged ones (by 3 p.p.), but exposure to bilingual schooling reduces the gap by a third. We further explore the role of parental origins splitting the sample by parental background and re-estimating the models with heterogeneous effects by regional origins or language use (Columns 5 to 8). As far as regional origins are concerned, the table shows that while low background individuals benefit from the reform irrespective of whether their parents are Catalan or not, among individuals from high background the reform has an effect only among non-Catalans. Patterns are less clear-cut for language use: low background individuals still benefit from the reform irrespective of language, but for high background individuals effects are generally non-significant, suggesting that for high background Spanish speakers proficiency in Catalan may have not been an issue even prior to the reform.

The results presented in this section are broadly consistent with the idea that the reform made the Catalan society more equal, reducing the gap in returns to education between high and low background individuals, and for high background individuals between Catalans and non-Catalans. The reform seems to have at least partly achieved its goal of levelling the playing field in the bilingual labor market.

## **8. Discussion and conclusion**

The language-in-education reform that in 1983 introduced Catalan alongside Spanish as a medium of instruction in the Catalan schooling system provides a unique opportunity for evaluating the effects of bilingual schooling. In this paper we exploit the differential exposure to this reform across birth cohorts and levels of education to provide the first evaluation of the earnings effects of bilingual education, complementing strands of literature that have focused either on the effects of bilingualism on educational outcomes, or the labor market effects of changing one language of instruction with another. We find positive earnings effects: one year of bilingual education raises earnings by on average 1.4 percent, representing one fifth of the baseline return to education. These

effects are robust to alternative specifications, and placebo analyses exclude that they could reflect spurious trends. Conversely, we do not find effects on employment, hours of work or occupation.

Our results are in line with the evidence reported in papers that analyse the wage effects of the policies that suppressed a given language from education, as in the case of French in Morocco (Angrist and Lavy, 1997) and English in the Indian State of West Bengala (Chakraborty and Bakshi, 2016). In their cases, removing an additional (and relevant) language from the schooling system generates a wage penalty for the affected individuals, while in ours, the same amount of schooling is more rewarded in the labour market when more than one language are used as medium of instruction.

Our results are consistent with different models of the labor market. The uncovered earnings effects may reflect the human capital impact of bilingual education, which fostered the development of cognitive skills of treated individuals. Also, faced with the need of offering new programs taught bilingually, schools recruited more skilled teachers to replace retiring ones, with an effect on the overall ability of the teaching body. In both cases, bilingual schooling can be seen as increasing the quality of education, raising the wage return of one year of schooling.

The estimated causal impacts are also consistent with search and matching mechanisms. In a bilingual labor market like Catalonia where most of employers are Catalans, being proficient in Catalan may give an advantage in the job search process, favouring the arrival of more and better offers and resulting in higher wages once employed. The existence of multiple channels reinforcing each other helps explain why our estimated effects are relatively sizeable.

Coming after forty years of dictatorship in which the use of Catalan was prosecuted, one of the main motivations of the reform was to level the playing field in Catalonia, promoting an effective integration of non-Catalans, many of whom were first- or second-generation immigrants from poorer Spanish regions. The reform was arguably successful in favouring the development of a unitary society: we provide evidence indicating that individuals of non-Catalan background (in terms of either origins or language) were amongst the main beneficiaries of bilingual education.

More in general, the reform reduced the wage gap across individuals of different social backgrounds irrespective of their regional origins, a fact which we interpret as reflecting a generalized increase in school quality produced by the reform.

## References

- Adesope O. O., Lavin T., Thompson, T. and Ungerleider C. (2010). "A systematic review and meta-analysis of the cognitive correlates of bilingualism," *Review of Educational Research*, 80 (2): 207–245.
- Anghel, B., Cabrales, A., and Carro, J. M. (2016). "Evaluating a bilingual education program in Spain: the impact beyond foreign language learning," *Economic Inquiry*, 54(2): 1202-1223.
- Angrist, J. D. and Lavy, V., (1997). "The effect of a change in language of instruction on the returns to schooling in morocco," *Journal of Labor Economics*, 15(1): 48-76.
- Angrist, J. D., Chin, A. and Godoy, R., (2008). "Is Spanish-only schooling responsible for the Puerto Rican language gap?" *Journal of Development Economics*, 85(1-2): 105-128.
- Arenas, J. (1990). "Llengua i educació a la Catalunya d'avui". La llar del llibre: Barcelona.
- Arenas, J., Muset, M. (2007). "La immersió lingüística. Una obra de govern, un projecte compartit". Centre d'Estudis Jordi Pujol. Barcelona.
- Arnau, J., and Vila, F. X. (2013). "Language-in-education policies in the Catalan language Area." In "Reviving Catalan at school. Challenges and instructional approaches", by Arnau, J. (Ed.), pages 1-28. *Multilingual Matters*, Bristol, UK.
- Artigal, J.M. (1997). "The Catalan immersion program." In Johnson, R.K. and Swain, M. (Eds.), *Immersion Education: International Perspectives*. Cambridge University Press: Cambridge.
- Aspachs-Bracons, O., Clots-Figueras, I., Costa-Font, J. and Masella, P. (2008). "Compulsory language educational policies and identity formation," *Journal of the European Economic Association*, 6(2-3): 434-444.

- Bialystok, E., Craik, F. I., Green, D. W., and Gollan, T. H. (2009). Bilingual minds. *Psychological science in the public interest*, 10(3), 89-129.
- Bialystok, E. (2007). "Acquisition of literacy in bilingual children: a framework for research," *Language Learning*, 57(s1): 45-77.
- Bleakley, H. and Chin, A. (2008). "What holds back the second generation? The intergenerational transmission of language human capital among immigrants," *Journal of Human Resources*, 43(2): 267-298
- Caminal, R. and Di Palolo, A. (2017). Your language or mine? The non-communicative benefits of language skills, *Economic Inquiry*, forthcoming
- Chakraborty, T., and Bakshi, S. K. (2016). "English language premium: Evidence from a policy experiment in India." *Economics of Education Review*, 50, 1-16.
- Chin, A., Daysal, N. M. and Imberman, S. A. (2013). "Impact of bilingual education programs on limited English proficient students and their peers: Regression discontinuity evidence from Texas," *Journal of Public Economics*, 107: 63-78.
- Clots-Figueras, I. and Masella, P. (2013). "Education, language and identity," *The Economic Journal*, 123(570): 332-357.
- Costa, A. and Sebastián-Gallés, N. (2014). "How does the bilingual experience sculpt the brain?" *Nature Reviews Neuroscience*, 15(5): 336-345.
- Garbin, G., Sanjuan, A., Forn, C., Bustamante, J. C., Rodríguez-Pujadas, A., Belloch, V., Hernandez, M., Costa, A. and Ávila, C. (2010). "Bridging language and attention: Brain basis of the impact of bilingualism on cognitive control". *Neuroimage*, 53(4), 1272-1278.
- Isphording, I. and Otten, S. (2014). "Linguistic barriers in the destination language acquisition of immigrants." *Journal of Economic Behavior & Organization* 105: 30-50.
- Ivlevs, A. and King, R. M. (2014). "2004 Minority Education Reform and pupil performance in Latvia," *Economics of Education Review*, 38, 151-166.

Lleras-Muney, A. and Shertzer, A. (2015). "Did the americanization movement succeed? An evaluation of the effect of English-only and compulsory schools laws on immigrants," *American Economic Journal: Economic Policy*, 7(3): 258-90.

Saiz, A. and Zoido, E. (2005). "Listening to what the world says: bilingualism and earnings in the United States," *The Review of Economics and Statistics*, 87(3): 523-538.

**Table 1: Imputed years of exposure**

year of birth	years of schooling							
	3	6	8	11	12	14	15	17-20
1961-65	0	0	0	0	0	0	0	0
1966	0	0	0	0	1	3	4	6
1967	0	0	0	1	2	4	5	7
1968	0	0	0	2	3	5	6	8
1969	0	0	0	3	4	6	7	9
1970	0	0	1	4	5	7	8	10
1971	0	0	2	5	6	8	9	11
1972	0	1	3	6	7	9	10	12
1973	0	2	4	7	8	10	11	13
1974	0	3	5	8	9	11	12	14
1975	1	4	6	9	10	12	13	15
1976	2	5	7	10	11	13	14	16
1977-82	3	6	8	11	12	14	15	17

**Table 2: Effects of the reform on labor market outcomes**

	(1)	(2)	(3)	(4)
	Earnings	Employment	Hours	White collar
years of schooling	0.065*** (0.007)	0.022*** (0.003)	-0.032 (0.114)	0.085*** (0.014)
years of exposure	0.014*** (0.005)	-0.003 (0.002)	0.126 (0.230)	0.001 (0.004)
Number of observations	3,949	4,765	3,899	3,949

\*\*\*, \*\*, \* denote significance at the 1, 5 and 10 percent level. OLS Regressions with Two-way Cluster Robust Standard Errors by year of birth and years of schooling. Additional controls shown in Table A4: wave, gender, potential experience and its square, year of birth dummies.

**Table 3: Earnings effects of the reform: sensitivity checks (N=3,949)**

	(1)	(2)	(3)	(4)
years of schooling	0.065*** (0.007)	0.055*** (0.004)	0.067*** (0.008)	0.042 (0.036)
years of exposure	0.014*** (0.005)	0.016** (0.006)	0.014 (0.009)	0.022*** (0.005)
employment selection correction	<i>no</i>	<i>yes</i>	<i>no</i>	<i>no</i>
age dummies	<i>no</i>	<i>no</i>	<i>yes</i>	<i>no</i>
educational CDF	<i>no</i>	<i>no</i>	<i>no</i>	<i>yes</i>

\*\*\*, \*\*, \* denote significance at the 1, 5 and 10 percent level. Columns (1), (3) and (4) are OLS Regressions with Two-way Cluster Robust Standard Errors by year of birth and years of schooling. Column (2) is the earnings equation of a model with endogenous selection into employment that uses the provincial unemployment rate when the individual was 16 as the exclusion restriction. Additional controls not shown: wave, gender, potential experience and its square, year of birth dummies.



**Table 4: Falsification exercises for the assumption of parallel trends in monthly earnings**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Baseline	Other Spanish regions	Never treated Catalan, placebo reform implemented in:							
			1963	1964	1965	1966	1967	1968	1969	1970
years of schooling	0.065*** (0.007)	0.059*** (0.003)	0.085*** (0.013)	0.084*** (0.012)	0.084*** (0.010)	0.084*** (0.010)	0.085*** (0.009)	0.086*** (0.009)	0.086*** (0.009)	0.087*** (0.009)
years of exposure	0.014*** (0.005)	0.003 (0.007)	0.001 (0.021)	0.003 (0.018)	0.004 (0.016)	0.004 (0.015)	0.000 (0.013)	-0.002 (0.011)	-0.005 (0.010)	-0.008 (0.010)
Number of observations	3,949	8,086	2,156	2,156	2,156	2,156	2,156	2,156	2,156	2,156

\*\*\*, \*\*, \* denote significance at the 1, 5 and 10 percent level. OLS Regressions with Two-way Cluster Robust Standard Errors by year of birth and years of schooling. Additional controls not shown: wave, gender, potential experience and its square, year of birth dummies. The model of Column (2) uses EU-SILC excluding observations from Catalonia and other bilingual regions. The model of Columns (3)-(10) uses ECVHP data with observations of never-treated individuals born between 1945 and 1965.

**Table 5: Earnings effects of exposure at compulsory schooling**

	(1)	(2)	(3)
	Baseline	Compulsory Exposure	Placebo Compulsory Exposure
years of schooling	0.065 <sup>***</sup> (0.007)	0.074 <sup>***</sup> (0.006)	0.053 <sup>***</sup> (0.004)
years of exposure	0.014 <sup>***</sup> (0.005)	0.021 <sup>***</sup> (0.006)	-0.002 (0.005)
Number of observations	3,949	3,949	8,086

\*\*\*, \*\*, \* denote significance at the 1, 5 and 10 percent level. Column (1) is an OLS Regressions with Two-way Cluster Robust Standard Errors by year of birth and years of schooling. Columns (2) and (3) are OLS Regressions with Standard Errors Clustered by year of birth. Additional controls not shown: wave, gender, potential experience and its square, year of birth dummies (column (1)) and birth-cohort dummies (columns (2) and (3)). Column (3) uses EU-SILC data excluding observations from Catalonia and other bilingual regions.

**Table 6: Heterogeneous effects of the reform on labor market outcomes**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
years of schooling	0.065 <sup>***</sup> (0.007)	0.070 <sup>***</sup> (0.008)	0.071 <sup>***</sup> (0.007)	0.086 <sup>***</sup> (0.012)	0.057 <sup>***</sup> (0.008)	0.093 <sup>***</sup> (0.018)	0.061 <sup>***</sup> (0.007)	0.092 <sup>***</sup> (0.019)
years of exposure	0.014 <sup>***</sup> (0.005)	0.013 <sup>***</sup> (0.005)	0.015 <sup>***</sup> (0.005)	0.008 (0.005)	0.016 <sup>***</sup> (0.006)	-0.000 (0.007)	0.016 <sup>***</sup> (0.005)	0.002 (0.006)
years of schooling × non-Catalan origins		-0.014 <sup>***</sup> (0.005)			-0.007 (0.006)	-0.023 <sup>***</sup> (0.009)		
years of exposure × non-Catalan origins		0.005 <sup>*</sup> (0.003)			-0.003 (0.005)	0.021 <sup>*</sup> (0.011)		
years of schooling × Spanish usual			-0.025 <sup>***</sup> (0.009)				-0.028 <sup>***</sup> (0.006)	-0.011 (0.018)
years of exposure × Spanish usual			0.007 <sup>**</sup> (0.003)				0.005 (0.003)	0.008 (0.007)
years of schooling × low parental background				-0.031 <sup>***</sup> (0.011)				
years of exposure × low parental background				0.010 <sup>***</sup> (0.002)				
Number of observations	3949	3949	3908	3831	2752	1079	2752	1079

<sup>\*\*\*</sup>, <sup>\*\*</sup>, <sup>\*</sup> denote significance at the 1, 5 and 10 percent level. OLS Regressions with Two-way Cluster Robust Standard Errors by year of birth and years of schooling. Additional controls not shown: wave, gender, potential experience and its square, year of birth dummies, dummy for non-Catalan origins (columns (2), (5) and (6)), dummy for Spanish as usual language (columns (3), (7) and (8)), dummy for low parental educational background (column (4)). Estimates in columns (5) and (6) are obtained from the subsample of individuals with high parental education. Estimates in columns (7) and (8) are obtained from the subsample of individuals with low parental education.

## Appendix Tables

**Table A1: Net monthly earnings (w) by wave, Euros 2006 (N=3,949)**

	wave 2006	wave 2011	waves 2006- 2011
w < 300	1.02	1.08	1.04
300 < w < 451	2.07	2.87	2.36
450 < w < 601	5.09	2.66	4.23
600 < w < 751	8.37	5.81	7.47
750 < w < 901	14.16	16.3	14.92
900 < w < 1051	14.01	9.55	12.43
1050 < w < 1201	18.35	15.72	17.42
1200 < w < 1501	16.47	18.02	17.02
1500 < w < 1801	10.09	16.73	12.43
1800 < w < 2401	6.46	7.9	6.96
2400 < w < 3001	2.43	2.15	2.33
3000 < w < 3601	0.55	0.79	0.63
w > 3600	0.94	0.43	0.76
Total	100	100	100
Estimated average net monthly earnings	1,243	1,185	1,205
Observed average net monthly earnings	1,253	1,198	1,218

Note: average monthly earnings are estimated from reported net monthly earnings in brackets by interval regression on a constant. Observed monthly earnings correspond to the original (continuous) net monthly earnings variable for the 2011 wave and to the mid-point of the net earnings intervals for the 2006 wave.

**Table A2: Descriptive statistics by subsamples**

	employed 2006		employed 2011		employed 2006-2011		not employed 2006		not employed 2011	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
wave 2011	0	--	1	--	0.35	0.48	0	--	1	0
net monthly earnings (in 2006 prices)	1198.4	573.2	1253.3	572.3	1217.8	573.4	--	--	--	--
weekly hours of work	40.15	10.30	39.76	9.89	40.01	10.16	--	--	--	--
white-collar high-skilled occupation	0.38	0.49	0.39	0.49	0.39	0.49	0.22	0.41	0.20	0.40
age	0.54	0.50	0.51	0.50	0.53	0.50	0.20	0.40	0.37	0.48
years of schooling	34.63	5.96	39.62	5.96	36.39	6.42	35.91	5.89	39.96	6.19
years of exposure	12.29	3.29	12.53	3.62	12.37	3.42	10.87	3.37	10.29	3.59
years of exposure at compulsory schooling (pot.)	6.69	5.50	6.97	5.75	6.79	5.59	4.69	4.98	5.28	5.28
potential experience (age - years of schooling - 6)	3.40	3.36	3.41	3.34	3.40	3.35	2.65	3.21	3.28	3.38
local unemployment rate at age 16	16.34	7.02	21.09	7.40	18.02	7.51	19.04	7.00	23.67	7.78
non-Catalan origins	0.38	0.48	0.41	0.49	0.39	0.49	0.48	0.50	0.49	0.50
Spanish usual	0.30	0.46	0.27	0.44	0.29	0.45	0.46	0.50	0.44	0.50
low parental background	0.71	0.45	0.73	0.45	0.72	0.45	0.82	0.38	0.80	0.40
LNA Reform exposure by birth cohort										
Spanish-only at school (1961-65)	0.21	0.41	0.21	0.41	0.21	0.41	0.26	0.44	0.26	0.44
exposure only from secondary education (1966-69)	0.18	0.38	0.18	0.38	0.18	0.38	0.24	0.43	0.15	0.36
partial exposure at all levels (1970-76)	0.37	0.48	0.37	0.48	0.37	0.48	0.32	0.47	0.34	0.47
fully exposed to LNA (1977-1982)	0.24	0.43	0.24	0.43	0.24	0.43	0.18	0.39	0.25	0.43
Number of observations	2,556		1,393		3,949		359		449	

**Table A3: Selected descriptive statistics of the placebo samples (never-treated cohorts from ECVHP and contemporaneous cohorts from EU-SILC from Spanish non-bilingual regions)**

Birth-cohort	ECVHP (2006-2011)				EU-SILC (2006-2011)	
	Baseline sample		Placebo sample: never treated		Placebo sample: other Spanish regions	
	1961-1982		1945-1966		1961-1982	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
wave 2011	0.35	0.48	0.35	0.48	0.49	0.5
net monthly earnings (in 2006 prices)	1217.8	573.4	1420.6	755.1	1334.8	809.4
male	0.53	0.50	0.56	0.50	0.53	0.50
age	36.39	6.42	50.03	5.76	37.39	6.78
years of schooling	12.37	3.42	11.50	3.78	12.34	4.23
years of exposure (real)	6.79	5.59	0.00	0.00	6.77	5.87
placebo years of exposure (fake reform in 1970)	--	--	6.18	5.15	--	--
placebo years of exposure (fake reform in 1969)	--	--	6.82	5.23	--	--
placebo years of exposure (fake reform in 1968)	--	--	7.42	5.25	--	--
placebo years of exposure (fake reform in 1967)	--	--	7.97	5.22	--	--
placebo years of exposure (fake reform in 1966)	--	--	8.47	5.14	--	--
placebo years of exposure (fake reform in 1965)	--	--	8.90	5.05	--	--
placebo years of exposure (fake reform in 1964)	--	--	9.30	4.94	--	--
placebo years of exposure (fake reform in 1963)	--	--	9.66	4.81	--	--
years of exposure at compulsory schooling (pot.)	3.40	3.35	0.00	0.00	3.22	3.41
potential experience (age - years of schooling - 6)	18.02	7.51	32.51	7.20	19.05	8.24
birth cohorts (by real/placebo LNA Reform exposure)						
never-treated cohort (1945-60)	--	--	0.62	0.49	--	--
Spanish-only at school (1961-65)	0.21	0.41	0.38	0.49	0.25	0.43
exposure only from secondary education (1966-69)	0.18	0.38	--	--	0.19	0.39
partial exposure at all levels (1970-76)	0.37	0.48	--	--	0.31	0.46
fully exposed to LNA (1977-1982)	0.24	0.43	--	--	0.25	0.43
Number of observations	3,949		2,156		8,086	

Note: summary statistics reported in the first two columns refer to the data from the Survey on Living Conditions and Habits of the Catalan Population (ECVHP), as in table 1. Data in the second column are taken from the same survey, using only observations of individuals born between 1945 and 1966 (never-treated cohorts) who meet the other selection criteria. Summary statistics reported in the last column refer to the EU-SILC, waves 2006 and 2011, excluding individuals from Catalonia and other bilingual regions or born outside Spain.

**Table A4: Coefficient estimates from regressions of monthly and hourly net earnings**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
waves:	2006-2011	2006	2011	2006	2011	2006-2011	2006-2011	2006-2011
constant	5.622*** (0.050)	5.629*** (0.064)	5.724*** (0.097)	5.388*** (0.031)	0.539*** (0.031)	5.343*** (0.095)	5.325*** (0.169)	5.175*** (0.099)
wave 2011	-0.016 (0.015)					-0.063** (0.030)	-0.065 (0.045)	-0.086** (0.034)
male	0.327*** (0.013)	0.349*** (0.017)	0.289*** (0.022)	0.349*** (0.054)	0.289*** (0.060)	0.327*** (0.053)	0.327*** (0.052)	0.328*** (0.055)
years of schooling	0.061*** (0.002)	0.059*** (0.003)	0.065*** (0.004)	0.059*** (0.015)	0.065*** (0.006)	0.061*** (0.009)	0.060*** (0.010)	0.065*** (0.007)
years of exposure				0.011* (0.006)	0.011 (0.007)	0.012*** (0.003)	0.015*** (0.003)	0.014*** (0.005)
potential experience	0.040*** (0.004)	0.040*** (0.005)	0.027*** (0.007)	0.053*** (0.011)	0.042** (0.021)	0.054*** (0.004)	0.060*** (0.008)	0.066*** (0.012)
potential experience <sup>2</sup>	-0.001*** (0.000)	-0.001*** (0.000)	-0.000** (0.000)	-0.001*** (0.000)	-0.001* (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Birth cohort dummies	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>yes</i>	<i>no</i>
Year of birth dummies	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>yes</i>
R <sup>2</sup>	0.247	0.245	0.253	0.246	0.255	0.249	0.250	0.253
Number of observations	3,949	2,556	1,393	2,556	1,393	3,949	3,949	3,949

Note: \*\*\* significant at 0.01, \*\* significant at 0.05, \* significant at 0.1; standard errors (within parenthesis) are robust to heteroscedasticity in columns (1)-(6) and are adjusted for year of birth-years of schooling-wave clusters in columns (7)-(9). Earnings in pooled 2006-2011 regressions are expressed in 2006 prices.