

The Effects of Linguistic Immersion on Cognitive Competencies

Jorge Calero^a and Álvaro Choi^{a*}

^aDepartment of Economics, University of Barcelona, Barcelona, Spain

*Facultat d'Economia i Empresa. 690, Diagonal Av. 08034, Barcelona, Spain. Email: alvarochoi@ub.edu Tel. +34934021815

Short bios

Jorge Calero is tenured Professor of Applied Economics at the University of Barcelona, department of Public Economics, Political Economics and Spanish Economics. He is currently coordinator of the Interdisciplinary Research Group on Educational Policies GIPE-IGEP. Previous Head of the Council of the Educational System Evaluation (Catalonia). His main research interests are the economics of education, educational policies evaluation, equity issues in education, and economics of the welfare state

Álvaro Choi is Associate Professor at the University of Barcelona and researcher at the Interdisciplinary Group on Educational Policies (GIPE-IGEP). His main research fields are related to the economics of education, public economics and the evaluation of public policies. He has participated in the evaluation of a number of educational policies. He has worked intensively with international large scale assessments and is an OECD's Thomas J. Alexander fellow.

The Effects of Linguistic Immersion on Cognitive Competencies

There is strong controversy over the application and effects of the Linguistic Immersion Policy in Catalonia, a policy that established that the Catalan language should constitute the only vehicular language during the different levels of compulsory schooling. While some of the effects of this policy have been determined, the evidence on other effects is scarce or non-existent. Within this group is the determination of the effect on academic performance of receiving schooling in a language other than the mother tongue. This paper analyses the performance of fifteen-year-old students as a function of their mother tongue. To do this, multivariate techniques are applied to the data provided by the assessments of PISA-2015. The results indicate that Spanish speakers obtain worse results than Catalan speakers all other variables been equal. The reduction in the level of competencies varies, in turn, among different subgroups of this collective.

Keywords: Bilingualism; linguistic immersion; academic performance; educational inequalities; education policy.

1. Introduction

The main aim of immersion programmes is for the students to master a language that is different from the one spoken at home by means of a very high intensity of exposure to that language at school. Immersion programmes are thus fundamentally different from bilingual programmes in the intensity of the exposure of the students to a specific language, their exposure to other languages being of a lesser or zero incidence.

Immersion policies are common in the educational centres of many countries in which parents typically enrol their children so that they can learn a foreign language. There are few cases in which these policies are compulsory or are applied to the population as a whole. As far as we know, Catalonia, Greenland and the Faroe Islands are the only educational systems within the European Union which apply a compulsory

monolingual immersion system (Arraiza, 2015). This is remarkable, bearing in mind that almost all countries in the EU, with the exception of Iceland and Portugal, have more than one native language spoken within their borders (BBC, 2014). This makes the Catalan experience a case of special interest for studying the effects of learning in a language other than the mother tongue. In any case, the analysis of the application of the linguistic immersion policy in Catalonia is especially relevant for at least three further reasons.

Firstly, Catalonia is a bilingual society, where almost all the adult population understands, speaks and writes in Spanish and Catalan, two Romance languages which are closely related to each other and to Occitan, French and Italian. According to the NGO *Plataforma per la Llengua*, there were approximately 10 million Catalan-speakers in 2021. However, schooling for students from ages 6 to 16 in public and private publicly-funded schools is exclusively in Catalan. Secondly, there is the long period of time during which the linguistic immersion has been applied: the so-called Language Immersion Programs were introduced in the 1992-93 academic year. The third point of interest of the linguistic immersion in Catalonia as an object of study is the existence of a hot debate over the advisability of maintaining this model. The immersion system is frequently presented from Catalan nationalist positions as an example of a successful policy that promotes social cohesion and has no negative side-effects (Oller et al., 2021). However, three decades after its implementation, there is no evidence as to whether the policy affects student performance differently based on their mother tongue (Woolard, 2016; Oller et al., 2020).

Therefore, the main aim of this study is to make this important contribution to the existing evidence base. We assess, for the first time, the average and heterogeneous effects by population subgroups of the Catalan Language Immersion Programs on

students' academic performance. This is vital for understanding the effects of this unique social experiment and for landing the Catalan political debate to a rational level. The results are also of interest to countries that are reflecting on the advisability of modifying their school language models.

Assessing the effects of the Catalan Language Immersion Programs on academic performance is however a challenging task because all students are “treated” –and consequently, there is no control group-. Longitudinal data is not available either. However, the 2015 OECD’s Programme for International Student Assessment (PISA) provides an optimal setting for analysing the effects of completing this assessment in a language other than the mother tongue. We apply multivariant techniques to the data, which contains a representative sample for Catalonia. The main findings can be summed up as follows: First, Spanish speakers obtain worse results than Catalan speakers all other variables being equal. Second, we discover that the reduction in the level of competencies varies, in turn, among different subgroups of this collective.

Subsequently, the study is structured as follows. Section 2 presents the institutional background and briefly reviews the literature on the topic. Section 3 describes the methodology and data used. In Section 4 the results are presented and discussed. Section 5 closes the article with the main conclusions.

2. Linguistic immersion in a bilingual society: The case of Catalonia

Catalonia is a region located in the northeast of Spain, with an approximate population of 7.5 million inhabitants. Catalonia is a diverse and bilingual society, in which practically all of the adult population understands, speaks and writes in the Spanish language - 98% approximately in 2018 (IDESCAT, 2019)- and an ever increasing majority understands (94.4%), speaks (81.2%) and writes (65,3%) in Catalan,

a language whose use and knowledge has undergone a rapid recovery after the years of repression of the Franco regime (the analogous figures were, in 1986, 90.3%, 64% y 31.5%). The education system has been one of the factors that allowed this recovery (Vila i Moreno, 2008). According to IDESCAT (2019), in 2018, 36,1% of the Catalan population over 15 years of age 36.1% had Catalan as their usual language; 48.6% Spanish, and 7.4% both. The remaining 7.9% usually speak in a language other than Catalan or Spanish.

The education system in Spain is highly decentralised. The central government establishes a general education act, within which the Spanish regions (*Comunidades Autónomas*) have a wide margin of manoeuvre, both at the regulatory, budgetary and curricular level. For the purposes of this study, one should bear in mind that education in Catalonia is compulsory from ages 6 to 16. This includes six years of primary education and four years of compulsory secondary education (ESO). Education from 0 to 6 is not compulsory but is, from ages 3 to 6, practically universalized.

The Linguistic Standardisation Law of Catalonia of 1983 established that Catalan students should master the basic competencies in Catalan and Spanish by the end of their compulsory. To this end, it was made compulsory to teach contents in both Spanish and Catalan. The number of centres in which the so-called Language Immersion Programs were applied - centres in which all of the classes, except for the subject of Spanish, were taught in Catalan - grew throughout the decade. From the academic year 1992-93, the school in Catalan became the only learning model in Catalonia and Spanish would only be used in the Spanish language classes (typically, three of the 25 teaching hours in a regular school week at the compulsory school levels). This was consolidated in 1998 with the Language Policy Law, which formally established, among other issues, the language immersion policy, that is, the use of

Catalan as the sole language of teaching at all non-university levels. The implementation of this policy, as has happened with virtually every educational reform carried out in Spain since the Transition, has been the subject of intense political debate, often leading to clearly polarised positions. This debate is usually based on ideological reasons.

The defenders of this policy emphasise the need to recover the use of the Catalan language after years of prohibition, in a socio-cultural environment that is expressed mainly in the Spanish language (Plataforma per la Llengua, 2017); that the immersion policy guarantees that students end up being bilingual, and that this brings with it advantages at a cognitive level (Costa & Sebastián-Gallés, 2014; Bialystok et al., 2012) and in the area of work (Capellari & Di Paolo, 2018); which represents a potential to increase social cohesion (Newman et al., 2008); and that the immersion policy does not affect the academic performance of Catalan students (Arnau & Vila, 2013). They further argue that this policy also guarantees Catalan students a command of the Spanish language similar to that of students from other autonomous communities at the end of compulsory schooling (Consell Superior d'Avaluació del Sistema Educatiu, 2017).

Its detractors, on the other hand, maintain that the conditions that arose in the early 1980s, in which the Catalan language found itself in a position of clear inferiority, have changed, and that parents have the right to school their children in the language of their preference (UNESCO, 2016; Coll, 2018); that immersion prevents students from being schooled in their mother tongue, which could affect their academic performance (AEB, 2017); that immersion is being used for political reasons in order to promote the Catalan national identity and to the detriment of the feeling of belonging to Spain (Oller et al., 2021); and that there is no updated and comparable evidence about the level of

Spanish language of Catalan students (Convivencia Cívica Catalana, 2016). Moreover, Oller et al. (2021) show that the immersion policy clashes with the preferences of the Catalan society, where a bilingual (Spanish-Catalan balanced) or even trilingual (incorporating subjects in English) education system is preferred to the current monolingual system. As Clua (2017: 41) describes, “*to teach in Catalan or in Spanish is a sensitive political issue that goes far beyond considerations about linguistic competencies in a bilingual context*”.

The scientific literature has provided robust evidence regarding some of the effects of linguistic immersion in Catalonia. Aspachs et al. (2008) and Clots-Figueras and Masella (2013) underline the link between this policy and the feeling of national identity of the boys and girls. Capellari and Di Paolo (2018), meanwhile, in a manner consistent with Rendon (2007) and with a methodological approach similar to that of Angrist et al. (2008), find that being bilingual in Spanish and Catalan has positive monetary returns in the Catalan labour market. However, so far there is no robust empirical evidence regarding the possible effects (positive or negative) of the language immersion policy on students' academic performance by mother tongue.

This is due to at least three reasons: first, the successive regional government's lack of interest in analysing a controversial issue; in the second place, the lack of variation in the application of the policy in Catalonia makes it difficult to build a control group to carry out impact assessments using quasi-experimental techniques. Finally, the educational evaluations available in Spain have limitations that make it difficult to analyse the effect of the policy. Neither the State General Diagnostic Assessment, nor the sixth and fourth year ESO (secondary compulsory education) tests organised by the Consell Superior d'Avaluació del Sistema Educatiu, nor the results of the OECD PISA program, have allowed the sign and size of the impact of linguistic immersion on the

overall performance of Catalan students to be determined. There is, therefore, a vacuum of evidence, for the Catalan case, on an issue that may be decisive for the educational trajectory of the students since, as UNESCO (see UNESCO 2016, for example) or authors such as Anghel et al. (2016) point out, receiving the first years of education in their mother tongue has a positive impact on the students' learning process. In contrast, the starting point of the language immersion policy is that it guarantees bilingualism from the initial stages of development, since students learn Spanish, theoretically, outside of school (reading books, watching films, etc.), in parallel to the Catalan language, which is learned mainly in the educational centre. In this regard, Ivanova and Costa (2007), Bialystok (2009) and Sadat et al. (2016) warn of the potential negative effect of bilingualism in the early stages on some dimensions of first language learning. On the other hand, Admiraal, Westhoff and de Bot (2006), analysing the effect of receiving secondary education in English in the Netherlands, obtain that the effect on the performance of these students is nil. Bleakley and Chin (2008) argue, for their part, that bilingual education from an early age favours the development of cognitive skills. It must be remembered here, in any case, that the linguistic immersion model of Catalonia is not bilingual, but monolingual.

Assuming that being bilingual constitutes an asset, on both an individual and a social level, the question that arises is whether, in an environment in which the qualifications obtained may have an impact on the risk of repeating the year (according to Eurostat, in 2018, approximately 1 in 3 Spanish students has repeated a year of studies before the age of 15) or determines the academic trajectory (Spain has the highest rates of early school leaving in the EU; 17.9% in 2018), the application of the language immersion policy is efficient for all students or whether, on the contrary, it has a series of winners and losers. If, as described by UNESCO (2016), being schooled in a

language other than the mother tongue has a negative effect on academic performance, then the Catalan linguistic immersion model may be hindering the equality of education opportunities principle for students from Spanish-speaking households.

Therefore, the review of the literature leads us to propose two initial hypotheses. First, students who use Spanish at home, being educated in a regime of linguistic immersion in Catalan, achieve a lower performance in the skills evaluated in PISA than their peers whose language used in the home is Catalan, all other personal, socio-cultural and economic characteristics being equal. Second, there may be heterogeneity in the relationship between the language used in the home and results in terms of competencies. The immersion policy may have a stronger effect on reading competencies than on numeracy, which uses a specific language. At the same time, Catalonia is a diverse region where the use of language is not homogeneously distributed. For example, there is a higher concentration of Spanish speaking families in Barcelona and its nearby towns than in the rural areas (IDESCAT, 2019). Thus, students in rural areas are more exposed to the Catalan language in their daily lives. Importantly, recent evidence has shown that the acquisition of skills during childhood evolves differently by sex (Borgonovi et al., 2021). At the same time, there are more privately owned centres in urban areas. Authors such as Bonal et al. (2019) have repetitively referred to the Catalan educational system as a segregated system by school ownership, where the combination of residential segregation and catching areas leads to the concentration of students with very different profiles across schools. Finally, the capacity of Spanish-speaking students to overcome the possible negative effects may depend on the economic and socio-cultural level of the household. Therefore, we test for the existence of differences across competencies (reading, numeracy and science) and the previously described population groups (by size of the town; by sex; by

ownership of the centre; and by the economic and socio-cultural level of the household.

The results will allow us to have a deeper understanding of the costs, in terms of equity and academic performance, of language immersion policies.

3. Data and methodology

3.1. Data

The main aim of our analysis consists in identifying the effect on the level of competencies of learning in a language (Catalan) other than that spoken at home. PISA-2015 is an optimal assessment for identifying this effect because a) students participating in PISA-2015 were born in the year 2000, well after the policy was implemented (in the year 1992); b) the region of Catalonia has a representative sample of its population; c) PISA competencies in Catalonia are assessed only in Catalan; d) all students in public and private publicly-funded schools are taught in Catalan; and e) the background questionnaire provides, among a rich set of other variables, information on the language spoken at home. Thus, the conditions under which the Catalan students take the PISA assessment are very similar to those of the other tests they usually take at school.

The cognitive skills that are used as dependent variables are those of science, reading and mathematics. In the full sample of PISA for Catalonia, 52 public and private centres participate, with a total of 1,769 students evaluated. We have withdrawn from the sample independent private schools, since in some of these centres schooling is not carried out under the regime of linguistic immersion in Catalan. With a view to ensuring, in turn, that the students analysed have received identical treatment (language immersion during all years of compulsory schooling), students who declare that they had arrived in Spain with 6 or more years of age are removed from the sample.

After applying the two filters mentioned, the sample finally used in our analysis is of 1,347 students, enrolled in 44 public or private-publicly financed centres.

The explanatory variables of interest are the language spoken at home and its aggregation at the centre level (the percentage of students in the centre who speak Catalan at home). This last variable approximates the linguistic environment in which the student's school life develops outside the classroom. At the same time, it is important to consider all those factors that may be incidentally affecting the results gap, favourable to students whose mother tongue is Catalan, as shown in Figure 1.

[Insert Figure 1 around here]

The rest of the explanatory variables are those which are generally used in the literature on educational performance (more specifically, in the literature derived from PISA), and act both at the level of the student and that of the centre. The individual and home variables are: sex and age of the student; socioeconomic level of the home; place of birth of the student and their parents; language spoken at home; years of preschool; whether the student has repeated an academic year during primary or secondary school; and whether the student has changed school during compulsory education. The detailed description of these variables can be consulted in Table 1. The school level variables included in the models are: city size; school size; the student to teacher ratio; class size; school ownership; the average socioeconomic status of the school; the percentage of immigrant students enrolled in that school; and the percentage of students who have Catalan as their mother tongue.

[Insert Table 1 here]

3.2. Methodology

Our methodological approach consists in a set of multivariate linear regressions. The nested structure of PISA, in which first a number of centres are selected at random and, then, a sample of fifteen-year-old students from these centres, prevents the direct application of ordinary least square models of linear regression, since this would violate the principle of independence of the observations; in other words, two students selected at random from the same centre will probably be more similar to each other than two students selected at random from two different centres. Hence, we have turned to multilevel regression models, in which the hierarchical structure of the data is taken into account. In our model, the performance Y of the students i of the schools j evaluated in PISA depends on a battery of k variables at an individual level (X – level 1) and l school variables (Z – level 2), as shown in equations 1 to 4. Equation 4 is obtained by substituting the β of equation 1 with equations 2 and 3 and it is estimated, separately, for the three cognitive skills evaluated by PISA (reading, mathematics and science).

Estimating equation 4 will allow us to answer to our first hypothesis (i.e. whether students who have Spanish or other languages at home obtain worse results than Catalan speakers all other variables being equal).

$$Y_{ij} = \beta_{0j} + \sum_{k=1}^n \beta_{1j} X_{kij} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim N(0, \sigma^2) \quad \text{Level 1 (individual)} \quad (1)$$

$$\beta_{0j} = \gamma_{00} + \sum_1^j \gamma_{01} Z_{1j} + \mu_{0j} \quad \mu_{0j} \sim N(0, \tau_0) \quad \text{Level 2 (centre)} \quad (2)$$

$$\beta_{1j} = \gamma_{10} \quad \text{Level 2 (centre)} \quad (3)$$

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{kij} + \gamma_{01} Z_{1j} + \mu_{0j} + \varepsilon_{ij} \quad (4)$$

The dependent variable Y gathers the performance of the students in the PISA tests through a system of ten plausible values (estimates of the student's score¹). The

independent variables X and Z introduced in the models, based on the literature on determinants of academic performance², are those described in subsection 2.1.

With a view to analysing the possible heterogeneous effects among population groups of schooling in a language other than the mother tongue (hypothesis 2), we proceed in two ways: in first place, we estimate model (4) but adding interaction terms to our variable of interest (language at home) by sex; ownership of the centre; size of the town; and economic and socio-cultural level of the household; in second place, we reestimate the model (4) but applying it to different subsamples. Table 2 shows the average performance differences, for scientific competence, according to the following variables: sex; ownership of the centre (public versus private publicly-funded); size of the town (Barcelona city compared to the rest of the territory); and economic and socio-cultural level of the household (33% higher in the ESCS index of PISA, compared to 33% lower). Missing values were replaced by the mean of the variable plus binary flags for the imputed values.

[Insert Table 2 here]

All regressions were estimated using the *repest* module for Stata (Avvisati and Keslair, 2020). This module takes into account the nested nature of the database and runs the regressions using the set of replicate weights provided by PISA. Additionally, it also allows for analyses with plausible values.

4. Results

Table 3 shows the results of the multilevel regressions applied to the simple set, results which confirm the first hypothesis for the case of competencies in science and reading and refute it for the case of mathematics. The coefficient of the variable of interest, which must be interpreted as an association, not as a relation of causality, is negative

(-10.36) and significant in the case of the science competence and also negative (-9.84) and significant in the case of the reading competence. Finally, the lack of significance of the variable of interest in the estimation of the mathematics competence can be explained by the use of a specific formalised language in the learning of mathematics.

[Insert Table 3 here]

It should be stressed that, after having controlled for a large battery of individual, family and centre variables, having the Spanish language as their mother tongue affects not only the students' performance in reading skills in Catalan (it is reasonable to assume that a negative association would have been obtained here for Catalan speakers in the case of having taken the PISA test in Spanish) but also their performance in other competencies such as science. These results can be interpreted in at least two ways: first, the fact of taking the PISA test in a language other than the mother tongue, something that also happens in the internal assessment tests conducted by the centres, negatively affects student performance; secondly, the fact of having received education during the previous years in a language other than the mother tongue has hindered the learning processes of Spanish-speaking students, preventing them from matching the performance of their classmates who have Catalan as their mother tongue. Naturally, there is also the possibility that these two effects combine. Our analysis, in any case, is not able to specify the channel through which the identified effect operates; the implications for the educational policy would vary depending on this. Of course, the existence of variables not incorporated into the model that could have different effects on Catalan-speaking and Spanish-speaking speakers could be skewing the results,

although it is difficult to identify any relevant variables in addition to those already incorporated into the models.

It is also interesting to note the lack of significance of the variables related to the “density” of students who use a language in the centre - an approximation of the dominant language of communication between students outside the classroom but during academic hours. A lower presence of Catalan-speaking students does not significantly reduce the results in any of the three competencies.

The rest of the variables that we use as a control in the model presented in Table 3 behave similarly to that established in previous studies of this type, with some differences that we will point out in each case. Regarding the sex of the students, the girls, with the rest of the variables being equal, achieve scores significantly lower than the boys in math and science and they do not present any significant advantages in reading. The ESCS index, of economic and socio-cultural status, is positively related to the scores of the three competencies, both when considered individually (student's ESCS index) and when considered in aggregate form (ESCS index of the families that school their children in the centre). It is these two variables that present the strongest association with student results.

As regards the variables that describe the previous educational background of the students, a positive association is observed in the three competencies with the number of years in which the student has attended early childhood education. There is also a strong negative association with the repetition of previous years (especially if this has occurred in primary education), although these variables suffer from inverse causality problems that prevent the results from being interpreted as causal effects. In this same area, one noteworthy observation is the loss of points linked to student

trajectories that imply changes of centre, with greater intensity when the change of centre occurs in secondary education.

The variables related to immigration are only significant when they refer to the density of immigrants in the centre. Immigrant status, the rest of the variables being equal, is not significantly associated with a loss of score in any of the three competencies. This is a result that differs from those obtained in previous analyses applied to the whole of Spain (see, for example, Calero and Escardíbul, 2013), in which a significant negative association is frequently detected, especially for first-generation immigrants. It is necessary to take into account, in any case, the size of the standard errors associated with the small number of students of immigrant origin included in the sample. Regarding the density of immigrants in the centre, the estimate indicates a negative association between high densities (greater than 30%) and scores, although this result only occurs in the case of the reading and science skills.

Among the variables related to the characteristics of the centre, none of the three related to their physical capacities and ratios (size of the centre, student-teacher ratio and classroom size) is significant, as is usual in this type of estimates. The variable referring to ownership of the centre indicates a significant negative association between private ownership (publicly funded centres, remember that independent private centres do not participate in the sample) and the score obtained in science. In several previous studies, such as Calero and Escardíbul (2007 and 2013), this variable was not significant.

Tables 4 to 8 allow us to verify the second hypothesis which, remember, referred to the existence of heterogeneity. Table 4 shows the coefficients of the interaction terms we added to the base model presented in Table 3. Tables 5 to 8 present the results of

replicating the analysis by population subgroups. We use in these tables the results referred to the regressions in which the science competence is the explained variable³.

Table 4 presents the values of the four interaction effects that were added to the base model (full results presented in Appendix Table A.2). These variables capture the existence of heterogeneous effects of sex, ownership of the centre, city size, and economic and sociocultural status, by the language at home variable, on academic performance. In other words, we test whether these variables affect students depending on their mother tongue.

[Insert Table 4 here]

The sex of the student does not affect differently to the students' performance by language at home. We only find a (weak) statistically significant effect for reading. Thus, Spanish speaking girls have a larger performance premium over boys than Catalan speaking girls. The findings for the ESCS interaction are certainly more interesting. Negative statistically significant coefficients were obtained for the three competencies. This implies that a one-point increase in the ESCS index (which, remember, is positively associated to academic performance) will have a smaller effect on the test scores of Spanish speaking students.

Combining the direct –negative- effect on performance of living outside the city of Barcelona with the positive coefficients obtained for the interactions, we may affirm that the negative effect of being a Spanish speaker is lower for those who do not live in Barcelona. Finally, the positive effect on performance of attending a private publicly funded school is higher for Spanish speaking students than it is for Catalan speaking students.

We now turn to the results of re-estimating the base model presented in Table 3 by subsamples (tables 5 to 8). Table 5 shows the results of the hierarchical regression applied separately to girls and boys. It can be observed that the coefficient of the variable of interest is not significant in the case of girls and it is, with a value of -15.62 , for boys. Formal tests confirmed that the coefficients for boys and girls were statistically different⁴. This greater sensitivity of the results of the boys with respect to the vehicular language in the centre also occurs in the case of the students who use a language different from Catalan and Spanish in their homes. In contrast, it can be seen in this same table how the coefficient associated with the ESCS variable is considerably higher in the case of girls than in the case of boys.

[Insert Table 5 here]

The second hypothesis is also confirmed in Table 6, where the estimations are shown as a function of the ownership of the centre. Our variable of interest (having Spanish as the language spoken at home) shows a significant negative coefficient (-26.1) in the case of public centres but not in the case of private publicly-funded centres. It can be observed, then, that due perhaps to different levels of intensity in the application of the linguistic immersion in Catalan, students from private publicly-funded centres do not present any differences in their performance associated to the language. Neither in the subsample for public centres nor in that of private publicly-funded centres are the variables of Catalan-speaking student density in the centre significant. One notable aspect of the control variable coefficients is the higher value associated to the economic and sociocultural status of the students schooled in public centres.

[Insert Table 6 here]

Table 7 presents the results of the estimation of the model for cities with more than 1,000,000 inhabitants (in this case, we are selecting exclusively in the city of Barcelona, where the sample has 12 centres) and for the other populations (where the subsample has 32 centres). The lack of significance of the variable related to the use of Spanish in the home in the case of Barcelona city stands out in these estimates. The coefficient, on the other hand, is significant (-10.82) for the rest of the sample. Regarding the control variables in this estimate, in the sub-sample of the city of Barcelona the coefficient associated with being in a school with a high density of Catalan-speaking students has a positive effect on performance in science.

[Insert Table 7 here]

Finally, Table 8 shows the results corresponding to two different subsamples according to the level of the index of economic and sociocultural status. In a first subsample, students whose families are in the top tertile of this index are selected; in a second subsample, those whose families are located in the lower tertile. We observe that it is in the first subsample, only, where the coefficient of the variable "language in the home: Spanish" is significant, with a value of -22.9. The lack of significance, in the case of the sub-sample with the highest level of ESCS, of the ESCS variable itself (both at the individual level and at the centre level) is explained by the fact that its variability is logically restricted to the selected range (the higher tertile, in this case). Finally, the coefficients associated to the density of Catalan speaking students seem to have a

positive effect for students with a lower level of ESCS, but are statistically not significant.

[Insert Table 8 here]

5. Discussion

This work was based on two hypotheses. The first one was formulated as follows:

“Students who use Spanish at home, being educated in a regime of linguistic immersion in Catalan, achieve a lower performance in the skills evaluated in PISA than their peers whose language used in the home is Catalan, all other personal, socio-cultural and economic characteristics being equal”. The empirical analysis undertaken, based on the estimation of hierarchical regression models applied to PISA-2015 data of public and private publicly-funded schools in Catalonia, confirms this hypothesis in the case of science and reading competences, and refutes it in the case of mathematics.

The second hypothesis referred to the existence of heterogeneity in the relationship between the language spoken at home and the level of competence. Our analyses confirm the second hypothesis. Specifically, we have focused on four variables, obtaining the following results: with regard to sex, boys are more negatively affected by the language at home-language at school mismatch than girls. As a tentative hypothesis for future research, this may be to differences by sex in maturation. As for the ownership of the centre, the coefficient is significant and negative for public centres, but it is not for the private publicly-funded centres. Indeed, our results have also shown that students who speak Spanish at home who are enrolled in private publicly-funded schools perform better -or, at least, not as bad- as their counterparts in public schools.

There is also heterogeneity depending on the municipality of residence: while in the group of municipalities with the exception of the city of Barcelona the coefficient is significant (and negative), it is not in the city of Barcelona.

Finally, when distinguishing according to the level of economic and socio-cultural status, we found a significant negative coefficient in the upper tertile of the index, but not in the lower tertile. We also found that the effect of improving the ESCS status is lower for Spanish-speaking students than for students who speak Catalan at school. To recap: our results confirm that Spanish-speaking students perform worse than Catalan-speaking students in tests in Catalan and that this negative effect varies across population subgroups.

The results that we have presented point clearly to the existence of side effects which had not been explored in the past, linked to the linguistic immersion policy in Catalonia which generates winners and losers. Obvious as some of our results may sound –students who study and take exams in their mother tongue outperform their classmates- they indicate that the immersion policy has a problematic dimension that has not yet been sufficiently addressed and, even worse, in some cases, even denied -the language immersion policy has been repeatedly described in Catalonia as a “success policy” (Generalitat de Catalunya, 2015) without side-effects (Consell Superior d’Avaluació del Sistema Educatiu (2017)). However, this qualification, which in addition to taking into account the recovery of the use of the Catalan language incorporates components of a political, ideological or identity nature, has not considered the effects described in this study. As Clots-Figueras and Masella, (2013) describe, the close relationship between language and feelings of national identity makes it difficult to establish rational debates around these types of policies. Indeed, some of the currently existing analyses which conclude that the language normalisation policy had

neutral effects on academic achievement were based on very poor evidence. In many cases, this evidence included not even sheer correlations between the application of the policy and educational outcomes, thus constituting a misuse of educational evaluations.

Our analyses suggest that the language immersion policy guarantees greater success for Catalan speakers. This is challenging from the point of view of equity, since Spanish speakers in Catalonia have, in turn, a lower socio-economic level, so the immersion policy forces them to face a double risk. Our results also show that improving their ESCS would not suffice to compensate the performance gap with Catalan speakers. Thus, policies aiming to tackle this gap should consider at least two dimensions: the economic and the linguistic.

This article shows that those studies that focused on the monetary effects of mastering both languages, ignored an important factor in their analysis: access to the labour market, strongly linked to previous educational trajectories, can be weighted down for part of the population, in this case the Spanish-speaking part, as a consequence of the application of the language immersion policy. The existing evidence clearly shows the link between reduced academic performance, a higher risk of early school leaving and a lower probability of completing higher education (European Commission, 2019; Schleicher, 2007).

We have revealed the existence of a series of side effects not described above and that seem to recommend that this design should be reconsidered. The recognition of the existence of winners and losers can contribute to the correction of some of the inequities that this immersion policy entails for users of the public educational service in Catalonia. One possible path for overcoming the problems derived from the policy of linguistic immersion might be to adapt the practices to a model of “linguistic conjunction” oriented to a real school bilingualism, which allows the coexistence of

Spanish and Catalan as vehicular languages. Some recent legal sentences, which require the provision of 25% of school content in Spanish in some centres, point in the same direction. A further change that might be considered is that the evaluable tests should be taken in the language of choice of the students.

These results underline the need for further empirical analyses to be conducted in order to examine more hypotheses related to the ones proposed here. For example, whether the performance gap we identified was due to the fact that the exam was taken in a language other than the mother tongue or to the fact of having received education during the previous years in a language other than the mother tongue. Analysing the relationship between language at home and early school dropout rates -the main problem of the Spanish and Catalan educational systems- would also be among them. The undertaking of quasi-experimental studies would provide further, essential knowledge about the causal relationship between the immersion policy and the results in terms of competencies, but also in the understanding of the channels through which said relationship operates. The possible limitations linked to the sample size of PISA reinforce the need for additional studies on this issue.

The currently applied language immersion policy has contributed to the public use of the Catalan language and, from that perspective, it has been positive. However, a rational and honest debate should acknowledge that the policy has had costs too. All in all, this study has provided for the first time the existence of some of the costs of monolingual immersion systems. This should be useful to policymakers aiming to improve the academic performance and reduce education inequalities of all Catalan students and to the international stakeholders who might be interested in rethinking the linguistic model of their educational system.

Footnotes

1. Broadly speaking, plausible values are the values imputed from the results that a student might obtain in a given test. They are calculated using a probabilistic model, based on a series of relevant characteristics of the student, once the distribution of the results has been obtained (this distribution is obtained a posteriori, on completion of the test) and taking into account the weighting obtained by means of the Item Response Theory. For a detailed explanation of this and other matters related to the design of PISA-2015, consult OECD (2017).
2. See, for example, Hanushek (2003), Calero and Escardíbul (2007) or Guio and Choi (2014).
3. Each wave of PISA focuses its interest mainly on one of the three competencies. In 2015, it was the turn of science.
4. This is relevant given the limited size of our samples. In fact, also of statistical significance is the difference between these coefficients, when we compare subgroups by socioeconomic level of the home (Table 7). In the other analyses, shown in Tables 5 and 6, the betas associated to the language of the home are not statistically different for the subgroups. Hence, the results should be interpreted with caution and further studies should be undertaken in the future with bigger samples. This does not prevent, however, the identification of patterns corresponding to population subgroups.

References

- Admiraal, W., Westhoff, G. & de Bot, K. (2006). Evaluation of Bilingual Secondary Education in The Netherlands: Students' Language Proficiency in English. *Educational Research and Evaluation*, 12, 75–93.
- Angrist, J.D., Chin, A. & Godoy, R. (2008). Is Spanish-Only Schooling Responsible for the Puerto Rican Language Gap? *Journal of Development Economics*, 85, 105-128.
- Anghel, B., Cabrales, A. & Carro, J.M. (2016). Evaluating a bilingual education program in Spain: the impact beyond foreign language learning. *Economic Inquiry*, 54, 1202-1223.
- Arenas, J. & Muset, M. (2007). *La immersió lingüística. Una obra de govern, un projecte compartit*. Centre d'Estudis Jordi Pujol.
- Arnau, J. & Vila, F.X. (2013). Language-in-education Policies in the Catalan Language Area. In J. Arnau (Ed.), *Reviving Catalan at School: Challenges and Instructional Approaches* (pp. 1-28). Multilingual Matters.

Arraiza, J. M. (2015). *Making Home Rules for Mother Tongues. The Legal Implications of Linguistic Diversity in the Design of Autonomy Regimes*. Åbo Akademi University.

Asamblea por una Escuela Bilingüe (2017). Informe de la Asamblea por una Escuela Bilingüe. Retrieved from <https://www.aebcatalunya.org/es/> [06/11/2020]

Aspachs-Bracons, O., Clots-Figueras, I., Costa-Font, J. & Masella, P. (2008). Compulsory Language Educational Policies and Identity Formation. *Journal of the European Economic Association*, 6, 434-444.

Avvisati, F. & Keslair, F. (2020). REPEST: Stata module to run estimations with weighted replicate samples and plausible values.

BBC (2014). Languages across Europe. Retrieved from: https://www.bbc.co.uk/languages/european_languages/countries/index.shtml [20/08/2021]

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, 12, 3-11.

Bialystok, E., Craik, F.I. & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences*, 16, 240-250.

Bleakley, H. & Chin, A. (2008). What Holds Back the Second Generation? The Intergenerational Transmission of Language Human Capital among Immigrants. *Journal of Human Resources*, 43, 267-298.

Bonal, X., Zancajo, A., & Scandurra, R. (2019). Residential segregation and school segregation of foreign students in Barcelona. *Urban Studies*, 56(15), 3251-3273.

Borgonovi, F., Choi, Á., & Paccagnella, M. (2021). The evolution of gender gaps in numeracy and literacy between childhood and young adulthood. *Economics of Education Review*, 82, 102119.

Calero, J. & Escardíbul, J.O. (2013). El rendimiento del alumnado de origen inmigrante en PISA-2012. In INEE (Ed.), *PISA-2012 Programa para la Evaluación Internacional de los Alumnos. Informe Español. Volumen II. Análisis secundario* (pp. 4-31). Instituto Nacional de Evaluación Educativa.

Calero, J. & Escardíbul, J.O. (2007). Evaluación de servicios educativos: el rendimiento en los centros públicos y privados medido en PISA-2003. *Hacienda Pública Española*, 83, 33-66.

Cappellari, L. & Di Paolo, A. (2018). Bilingual schooling and earnings: Evidence from a language-in-education reform. *Economics of Education Review*, 64, 90-101.

- Clua, M. (2017). Language, national identity and school: The role of the Catalan-Language Immersion Program in contemporary Catalan nationalism. In *The State, Schooling and Identity* (pp. 41-59). Singapore: Palgrave Macmillan.
- Clots-Figueras, I. & Masella, P. (2013). Education, Language and Identity. *The Economic Journal*, 123, 332-357.
- Coll, J. (2018). Los socialistas y la escuela catalana. In *El País*, 26th February 2018.
- Consell Superior d'Avaluació del Sistema Educatiu (2017). Conocimiento de las lenguas catalana y castellana del alumnado de Cataluña. *Informes d'Avaluació*, 21.
- Convivencia Cívica Catalana (2016). La lengua española en la educación de Cataluña. Retrieved from <http://convivenciavicacatalana.blogspot.com/> [04/11/2020]
- Costa, A. & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews Neuroscience*, 15, 336-345.
- European Commission (2019). *PISA 2018 and the EU: Striving for social fairness through education*. Brussels: European Commission.
- Generalitat de Catalunya (2015). La llengua a l'escola catalana: un model d'èxit. Barcelona: Generalitat de Catalunya.
- Guio, J. & Choi, Á. (2014). The Evolution of School Failure Risk During the 2000 Decade in Spain: Analysis of PISA Results with a Two-Level Logistic Model. *Estudios sobre Educación*, 26, 33-62.
- Hanushek, E.A. (2003). The failure of input-based schooling policies. *The Economic Journal*, 113, 64-98.
- IDESCAT (2019). *Enquesta d'usos lingüístics*. IDESCAT.
- Ivanova, I. & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, 127, 277-288.
- Newman, M. Trenchs-Parera, M, & Ng, S. (2008). Normalizing bilingualism: The effects of the Catalan linguistic normalization policy one generation after. *Journal of Sociolinguistics*, 12, 306-333.
- OECD (2017). *PISA 2015: Technical Report*. Paris: OECD.
- Oller, J.M; Satorra, A.& Tobeña, A. (2020). Privileged Rebels: A longitudinal analysis of distinctive economic traits of Catalan secessionism. *Genealogy*, 4: 19.
- Oller, J.M; Satorra, A.& Tobeña, A. (2021). Parochial Linguistic Education: Patterns of an Enduring Friction within a Divided Catalonia. *Genealogy*, 5: 77.

- Plataforma per la Llengua (2017). *La immersió lingüística a Catalunya. Un model eficaç i exitós*. Plataforma per la Llengua.
- Rendon, S. (2007). The Catalan premium: language and employment in Catalonia. *Journal of Population Economics*, 20, 669-686.
- Sadat, J., Martin, C.D., Magnuson, J.S., Alario, F.X. & Costa, A. (2016). Breaking down the bilingual cost in speech production. *Cognitive Science*, 40, 1911-1940.
- Schleicher, A. (2007). Can competencies assessed by PISA be considered the fundamental school knowledge 15-years-olds should possess? *Journal of Educational Change*, 8, 349-357.
- UNESCO (2016). Si no entienes, ¿cómo puedes aprender? *Documento de Política*
- Vila i Moreno, F. X. (2008). Language-in-education policies in the Catalan language area. *Aila Review*, 21(1), 31-48.
- Woolard, K. A. (2016). *Singular and Plural: Ideologies of Linguistic Authority in 21st Century Catalonia*. New York: Oxford University Press.

Table 1. Descriptive statistics

	N	Mean	Standard deviation	Min	Max
Dependent variables					
- Performance in mathematical competencies*	1,347	499.84	83.53	219.67	740.82
- Performance in scientific competencies*	1,347	504.82	88.57	230.57	752.12
- Performance in reading competencies*	1,347	498.94	86.93	192.31	743.47
Individual and household level variables					
- Female	1,347	48.7%	0.50	0	1
- Age (years)	1,347	15.85	0.29	15.42	16.42
- Economic and sociocultural status (ESCS index**)	1,347	-0.41	1.11	-3.70	2.09
- Native	1,347	88.6%	0.37	0	1
- First generation immigrant	1,347	7.0%	0.25	0	1
- Second generation immigrant	1,347	4.4%	0.20	0	1
- Language at home: Catalan	1,347	42.1%	0.49	0	1
- Language at home: Spanish	1,347	53.9%	0.50	0	1
- Language at home: Other	1,347	4.0%	0.20	0	1
- Preschool attendance (number of years)	1,347	2.81	1.25	0	5
- Repeated during primary education	1,347	7.73%	0.26	0	1
- Repeated during secondary education	1,347	16.7%	0.37	0	1
- Changed school during primary education	1,347	22.4%	0.52	0	1
- Changed school during secondary education	1,347	15.6%	0.45	0	1
School level variables					
- City size: over 100,000 inhabitants	1,347	49.4%	0.50	0	1
- City size: 10,000 to 100,000 inhabitants	1,347	28.1%	0.45	0	1
- City size: below 10,000 inhabitants	1,347	22.5%	0.42	0	1
- School size (number of students)	1,347	655.41	432.66	25	2,635
- Students to teacher ratio	1,347	11.52	3.35	1.85	16.86
- Class size (number of students)	1,347	25.35	5.57	13	48
- School ownership: public	1,347	71.7%	0.46	0	1
- School ownership: private publicly-funded	1,347	28.3%	0.45	0	1

- Average ESCS of school	1,347	-0.42	0.55	-1.40	0.96
- Immigrant students' density in school >30%	1,347	15.4%	0.36	0	1
- Percentage of students with Catalan as their mother tongue: >60%	1,347	25.1%	0.43	0	1
- Percentage of students with Catalan as their mother tongue: between 30% and 60%	1,347	38.5%	0.49	0	1
- Percentage of students with Catalan as their mother tongue: <30%	1,347	36.3%	0.48	0	1

* Performance variables display average results for a set of 10 plausible values. ** The ESCS index is derived, using Principal Component Analysis, from the following variables: number of books and educational resources at home, parental occupation, parental education and home possessions. Higher values indicate a higher ESCS status. Goodness-of-fit measures available upon request.

Table 2. Raw gender gap in the performance in scientific competencies of 15-year- old students in Catalonia by language at home, sex, school ownership, city size and household economic and sociocultural status

	Language at home			<i>Diff Cat- Spa</i>	<i>Diff Cat- other</i>
	Catalan	Spanish	Other		
Sex					
A. Males	537.35	494.75	422.25	<i>42.6</i>	<i>115.1</i>
B. Females	520.90	486.79	440.69	<i>34.11</i>	<i>80.21</i>
<i>Difference (A-B)</i>	<i>16.45</i>	<i>7.96</i>	<i>-18.44</i>		
School ownership					
A. Public	528.04	483.22	427.02	<i>44.82</i>	<i>101.02</i>
B. Private, publicly-funded	534.61	515.33	495.95	<i>19.28</i>	<i>38.66</i>
<i>Difference (A-B)</i>	<i>-6.57</i>	<i>-32.11</i>	<i>-68.93</i>		
City size					
A. City of Barcelona	541.71	499.98	466.97	<i>41.73</i>	<i>74.74</i>
B. Rest of Catalonia	529.57	490.87	433.33	<i>38.69</i>	<i>96.24</i>
<i>Difference (A-B)</i>	<i>12.14</i>	<i>9.11</i>	<i>33.64</i>		
Economic and sociocultural status (ESCS)					
A. Top 33%	550.81	524.27	496.66	<i>26.54</i>	<i>54.15</i>
B. Lower 33%	490.57	474.83	414.81	<i>15.74</i>	<i>75.76</i>
<i>Difference (A-B)</i>	<i>60.24</i>	<i>49.44</i>	<i>81.85</i>		

Source. Self-elaborated from PISA-2015. In italics, performance gap between subgroups.

Table 3. Determinants of performance in the scientific, reading and mathematical competencies of 15-year-old students in Catalonia. Hierarchical linear models. PISA-2015

	Science	Reading	Mathematics
Intercept	431.18***	415.92***	453.03***
	(132.47)	(142.34)	(121.89)
Individual and household level variables			
- Sex (ref: males)	-19.43***	6.63	-24.85***
	(4.48)	(4.65)	(4.23)
- Age (years)	9.36	8.29	6.66
	(8.40)	(9.08)	(7.66)
- Economic and sociocultural status (ESCS index)	11.07***	9.78***	10.95***
	(1.99)	(2.30)	(2.44)
- First generation immigrant (ref: native)	-7.33	0.84	-14.65
	(12.20)	(13.26)	(11.77)
- Second generation immigrant	-9.05	-1.09	-9.06
	(11.18)	(13.05)	(10.23)
- Language at home: Spanish (ref: Catalan)	-10.36**	-9.84***	-6.23
	(4.45)	(4.30)	(6.23)
- Language at home: Other languages	-25.54**	-19.22	-18.97
	(10.68)	(14.81)	(14.57)
- Preschool attendance (number of years)	4.43**	4.58**	4.50***
	(1.96)	(1.97)	(2.21)
- Grade repetition during primary school (ref: no repetition)	-82.31***	-83.19***	-74.24***
	(8.60)	(9.28)	(8.47)
- Grade repetition during secondary school (ref: no repetition)	-42.24***	-42.80***	-44.25***
	(6.15)	(7.06)	(7.43)
- Changed school during primary school	-12.30**	-11.08	-9.15*
	(6.10)	(6.88)	(5.32)
- Changed school during secondary school	-24.36***	-26.30***	-23.36**
	(6.32)	(6.19)	(7.86)
School level variables			

- City size (10,000 to 100,000)	-25.16***	-20.92***	-16.26**
(ref. more than 100,000 inhabitants)	(6.67)	(6.31)	(6.95)
- City size (below 10,000 inhabitants)	-12.04	-11.20	-6.24
	(12.73)	(9.99)	(11.40)
- School size (number of students)	0.00	-0.01	0.00
	(0.01)	(0.01)	(0.01)
- Students to teacher ratio	-1.18	-0.65	-1.00
	(1.08)	(1.14)	(1.01)
- Class size (number of students)	0.13	0.17	0.29
	(0.54)	(0.47)	(0.44)
- School ownership: private publicly-funded (ref: public school)	-22.03***	-12.70	-15.50*
	(8.53)	(9.24)	(9.64)
- Average ESCS of school	28.97***	22.52**	24.10***
	(10.31)	(10.28)	(8.76)
- Immigrant students' density in school >30%	-16.39*	-14.60*	-11.03*
	(8.84)	(8.35)	(8.20)
- Percentage of students with Catalan as their mother tongue: between 30% and 60%	7.58	5.62	4.90
	(7.72)	(7.89)	(7.15)
- Percentage of students with Catalan as their mother tongue: <30% (ref: schools with at least 60% of students with Catalan as their mother tongue)	5.73	1.93	4.83
	(11.55)	(10.88)	(10.53)
N	1,347	1,347	1,347

Source. Self-elaborated from PISA-2015. ***, statistically significant at 99%, ** statistically significant at 95%, * statistically significant at 90%. Standard errors displayed in parenthesis.

Table 4. Determinants of performance in the scientific, reading and mathematical competencies of 15-year-old students in Catalonia; model with interactions: coefficients. Hierarchical linear models. PISA-2015

	Science	Reading	Mathematics
Individual and household level variables	Yes	Yes	Yes
School level variables	Yes	Yes	Yes
- Spanish at home x Sex	10.30 (6.47)	15.50* (8.17)	3.46 (6.69)
- Other language at home x Sex	-5.90 (31.29)	-4.07 (32.03)	0.62 (26.93)
- Spanish at home x ESCS	-10.02** (5.12)	-9.07* (5.16)	-9.46*** (4.85)
- Other language at home x ESCS	-11.07 (10.57)	-14.84 (11.36)	-11.59 (9.39)
- Spanish at home x Town	27.64** (11.35)	17.31 (11.20)	23.89** (10.64)
- Other language at home x Town	0.20 (25.56)	-16.28 (36.78)	-3.69 (35.37)
- Spanish at home x Village	23.18* (12.55)	12.43 (13.30)	23.05* (12.29)
- Other language at home x Village	-25.18 (25.78)	-24.86 (30.18)	-22.13 (24.40)
- Spanish at home x School ownership	33.14*** (12.67)	29.60** (11.69)	27.30*** (10.00)
- Other language at home x School ownership	39.93 (41.55)	60.15 (45.13)	20.40 (45.23)
N	1,347	1,347	1,347

Source. Self-elaborated from PISA-2015. ***, statistically significant at 99%, ** statistically significant at 95%, * statistically significant at 90%. Standard errors displayed in parenthesis. The models include the same individual, household and school level variables as the base model in Table 3. Full results available in Appendix Table A.2.

Table 5. Determinants of performance in the scientific competencies of 15-year-old students in Catalonia. Hierarchical linear models, by sex, PISA-2015.

	Females	Males
Intercept	348.31*	508.77***
	(183.10)	(179.10)
Individual and household level variables		
- Age (years)	10.88	6.83
	(11.15)	(11.50)
- Economic and sociocultural status (ESCS index)	16.30***	6.56**
	(3.14)	(2.84)
- First generation immigrant (ref: native)	-9.95	-6.77
	(14.96)	(16.45)
- Second generation immigrant	5.97	-25.68
	(14.61)	(17.49)
- Language at home: Spanish (ref: Catalan)	-1.34	-15.62**
	(5.04)	(6.60)
- Language at home: Other languages	-18.84	-32.18*
	(17.21)	(19.65)
- Preschool attendance (number of years)	7.00***	2.02
	(2.27)	(2.46)
- Grade repetition during primary school (ref: no repetition)	-100.04***	-65.90***
	(14.58)	(13.24)
- Grade repetition during secondary school (ref: no repetition)	-33.82***	-48.50***
	(10.61)	(10.23)
- Changed school during primary school	-5.21	-22.81***
	(9.32)	(8.29)
- Changed school during secondary school	-25.74**	-25.21 **
	(10.52)	(8.20)
School level variables		
- City size (10,000 to 100,000)	-22.85**	-26.57***
(ref. more than 100,000 inhabitants)	(9.01)	(9.14)
- City size (below 10,000 inhabitants)	-2.35	-22.58

	(14.27)	(16.19)
- School size (number of students)	-0.01	0.00
	(0.01)	(0.01)
- Students to teacher ratio	0.30	-2.89*
	(1.51)	(1.50)
- Class size (number of students)	0.58	-0.13
	(0.75)	(0.61)
- School ownership: private publicly-funded (ref: public school)	-24.89**	-20.75**
	(10.84)	(10.45)
- Average ESCS of school	19.15**	38.35***
	(12.02)	(12.99)
- Immigrant students' density in school >30%	-13.21	-21.05*
	(11.29)	(11.00)
- Percentage of students with Catalan as their mother tongue: between 30% and 60%	-5.08	15.99*
	(10.60)	(8.91)
- Percentage of students with Catalan as their mother tongue: <30% (ref: schools with at least 60% of students with Catalan as their mother tongue)	-3.68	10.79
	(14.54)	(13.84)
N	666	681

Source. Self-elaborated from PISA-2015. ***, statistically significant at 99%, ** statistically significant at 95%, * statistically significant at 90%. Standard errors displayed in parenthesis.

Table 6. Determinants of performance in the scientific competencies of 15-year-old students in Catalonia. Hierarchical linear models, by school ownership, PISA-2015.

	Public	Private publicly-funded
Intercept	424.30** (166.87)	281.39 (227.55)
Individual and household level variables		
- Sex (ref: males)	-19.14*** (5.69)	-23.33*** (6.94)
- Age (years)	10.64 (10.48)	8.67 (14.28)
- Economic and sociocultural status (ESCS index)	12.30*** (2.40)	5.93 (3.85)
- First generation immigrant (ref: native)	-9.67 (12.92)	19.58 (30.72)
- Second generation immigrant	-7.11 (11.86)	-13.66 (34.61)
- Language at home: Spanish (ref: Catalan)	-11.48** (5.10)	-11.72 (8.73)
- Language at home: Other languages	-26.08** (11.33)	-18.22 (26.86)
- Preschool attendance (number of years)	5.76** (2.29)	4.59 (3.30)
- Grade repetition during primary school (ref: no repetition)	-81.18*** (9.94)	-86.92*** (17.46)
- Grade repetition during secondary school (ref: no repetition)	-42.72*** (5.92)	-16.01 (21.63)
- Changed school during primary school	-12.37* (6.88)	-12.06 (12.02)
- Changed school during secondary school	-27.50*** (8.73)	-18.40** (9.17)
School level variables		

- City size (10,000 to 100,000)	-16.17**	-50.04***
(ref. more than 100,000 inhabitants)	(6.97)	(8.17)
- City size (below 10,000 inhabitants)	-3.87	93.02***
	(13.56)	(35.16)
- School size (number of students)	0.00	-0.01
	(0.01)	(0.01)
- Students to teacher ratio	-2.04*	8.17***
	(1.06)	(1.98)
- Class size (number of students)	-0.00	0.69
	(0.88)	(0.72)
- Average ESCS of school	43.01***	18.85**
	(15.60)	(9.36)
- Immigrant students' density in school >30%	-5.01	-68.03***
	(8.52)	(14.60)
- Percentage of students with Catalan as their mother tongue: between 30% and 60%	7.48	-10.19
	(9.65)	(10.74)
- Percentage of students with Catalan as their mother tongue: <30%	5.88	0.30
(ref: schools with at least 60% of students with Catalan as their mother tongue)	(16.06)	(12.49)
N	929	389

Source. Self-elaborated from PISA-2015. ***, statistically significant at 99%, ** statistically significant at 95%, * statistically significant at 90%. Standard errors displayed in parenthesis.

Table 7. Determinants of performance in the scientific competencies of 15-year-old students in Catalonia. Hierarchical linear models, by city size, PISA-2015.

	City of Barcelona	Rest of Catalonia
Intercept	647.43*** (213.94)	295.91* (165.93)
Individual and household level variables		
- Sex (ref: males)	-27.51*** (9.69)	-14.44*** (5.43)
- Age (years)	-5.19 (11.47)	15.99 (10.52)
- Economic and sociocultural status (ESCS index)	11.14* (4.34)	10.73*** (2.45)
- First generation immigrant (ref: native)	-18.62 (11.63)	0.73 (17.00)
- Second generation immigrant	-10.68 (18.97)	-16.73 (12.49)
- Language at home: Spanish (ref: Catalan)	-4.94 (8.14)	-10.82* (5.57)
- Language at home: Other languages	-3.06 (27.05)	-31.29** (14.10)
- Preschool attendance (number of years)	-0.66 (3.21)	7.08*** (2.31)
- Grade repetition during primary school (ref: no repetition)	-88.97*** (13.42)	-77.55*** (11.81)
- Grade repetition during secondary school (ref: no repetition)	-33.46** (13.42)	-45.67*** (6.87)
- Changed school during primary school	-2.82 (10.69)	-16.88** (7.00)
- Changed school during secondary school	-19.62* (10.99)	-27.60*** (7.94)
School level variables		
- School size (number of students)	-0.04***	0.01

	(0.01)	(0.01)
- Students to teacher ratio	2.24	-2.10*
	(9.40)	(1.15)
- Class size (number of students)	-1.98***	0.05
	(0.75)	(1.00)
- School ownership: private publicly-funded (ref: public school)	28.76	-10.17
	(51.34)	(10.34)
- Average ESCS of school	48.39	16.37
	(40.81)	(13.45)
- Immigrant students' density in school >30%	17.57	-15.01
	(23.55)	(11.09)
- Percentage of students with Catalan as their mother tongue: between 30% and 60%	50.07***	12.71
	(16.34)	(9.71)
- Percentage of students with Catalan as their mother tongue: <30%	50.99	4.64
	(45.61)	(11.98)
(ref: schools with at least 60% of students with Catalan as their mother tongue)		
N	386	961

Source. Self-elaborated from PISA-2015. ***, statistically significant at 99%, ** statistically significant at 95%, * statistically significant at 90%. Standard errors displayed in parenthesis.

Table 8. Determinants of performance in the scientific competencies of 15-year-old students in Catalonia. Hierarchical linear models, by economic and sociocultural status, PISA-2015.

	Top 33%	Lower 33%
Intercept	533.79** (217.40)	507.09** (205.62)
Individual and household level variables		
- Sex (ref: males)	-19.47** (7.41)	-20.01** (7.64)
- Age (years)	3.28 (13.32)	3.56 (13.04)
- Economic and sociocultural status (ESCS index)	11.30 (11.28)	12.93** (6.59)
- First generation immigrant (ref: native)	20.33 (23.01)	-15.10 (12.09)
- Second generation immigrant	18.65 (37.68)	-7.20 (17.90)
- Language at home: Spanish (ref: Catalan)	-22.16** (9.53)	2.29 (8.10)
- Language at home: Other languages	-36.62* (19.30)	-8.76 (16.91)
- Preschool attendance (number of years)	6.10** (2.92)	5.30 (3.86)
- Grade repetition during primary school (ref: no repetition)	-104.19*** (21.28)	-74.58*** (13.86)
- Grade repetition during secondary school (ref: no repetition)	-45.21** (18.76)	-34.97*** (9.59)
- Changed school during primary school	-12.79 (11.08)	-22.65** (10.29)
- Changed school during secondary school	-12.98 (9.26)	-14.95 (9.94)
School level variables		

- City size (10,000 to 100,000)	-44.66***	-5.64
(ref. more than 100,000 inhabitants)	(9.73)	(10.99)
- City size (below 10,000 inhabitants)	-26.79**	1.99
	(12.20)	(17.87)
- School size (number of students)	0.01	0.00
	(0.01)	(0.02)
- Students to teacher ratio	-1.05	-0.92
	(1.52)	(1.83)
- Class size (number of students)	0.64	-0.62
	(0.47)	(1.14)
- School ownership: private publicly-funded (ref: public school)	-46.28***	-6.70
	(14.54)	(11.80)
- Average ESCS of school	17.02	46.79***
	(15.19)	(15.43)
- Immigrant students density in school >30%	-29.56**	-4.39
	(11.83)	(11.53)
- Percentage of students with Catalan as their mother tongue: between 30% and 60%	-8.19	21.74
	(10.40)	(15.27)
- Percentage of students with Catalan as their mother tongue: <30%	-2.94	26.64
	(17.84)	(18.04)
(ref: schools with at least 60% of students with Catalan as their mother tongue)		
N	449	449

Source. Self-elaborated from PISA-2015. ***, statistically significant at 99%, ** statistically significant at 95%, * statistically significant at 90%. Standard errors displayed in parenthesis.

Table A.1. Raw gender gap in the performance in mathematical and reading competencies of 15-year-old students in Catalonia by language at home, sex, school ownership, city size and household economic and sociocultural status

	Language at home				
	Catalan	Spanish	Other	<i>Diff Cat-Spa</i>	<i>Diff Cat-other</i>
Sex					
A. Males	531.79	495.60	420.37	<i>36.19</i>	<i>111.42</i>
B. Females	512.83	479.06	435.45	<i>33.77</i>	<i>77.38</i>
<i>Difference (A-B)</i>	<i>18.96</i>	<i>16.54</i>	<i>-15.08</i>		
School ownership					
A. Public	519.59	480.48	424.41	<i>39.11</i>	<i>95.18</i>
B. Private, publicly-funded	530.28	510.24	477.12	<i>20.04</i>	<i>53.16</i>
<i>Difference (A-B)</i>	<i>-10.69</i>	<i>-29.76</i>	<i>-52.71</i>		
City size					
A. Barcelona	531.96	493.12	451.20	<i>38.84</i>	<i>80.76</i>
B. Rest of Catalonia	522.82	487.54	429.43	<i>35.28</i>	<i>93.39</i>
<i>Difference (A-B)</i>	<i>9.14</i>	<i>5.58</i>	<i>21.77</i>		
Economic and sociocultural status (ESCS)					
A. Top 33%	543.65	519.76	483.73	<i>23.89</i>	<i>59.92</i>
B. Lower 33%	484.65	472.26	416.25	<i>12.39</i>	<i>68.40</i>
<i>Difference (A-B)</i>	<i>59.00</i>	<i>47.50</i>	<i>67.48</i>		
	Language at home				
	Catalan	Spanish	Other	<i>Diff Cat-Spa</i>	<i>Diff Cat-other</i>
Sex					
A. Males	516.36	474.42	412.50	<i>41.94</i>	<i>103.86</i>
B. Females	526.01	495.68	456.84	<i>30.33</i>	<i>69.17</i>
<i>Difference (A-B)</i>	<i>-9.65</i>	<i>-21.26</i>	<i>-44.34</i>		
School ownership					
A. Public	516.97	475.33	429.00	<i>41.64</i>	<i>87.97</i>
B. Private, publicly-funded	529.51	514.01	526.05	<i>15.50</i>	<i>3.46</i>
<i>Difference (A-B)</i>	<i>-12.54</i>	<i>-38.68</i>	<i>-97.05</i>		
City size					
A. City of Barcelona	532.44	492.98	475.89	<i>39.46</i>	<i>56.55</i>
B. Rest of Catalonia	520.93	484.79	439.15	<i>36.14</i>	<i>81.78</i>
<i>Difference (A-B)</i>	<i>11.51</i>	<i>8.19</i>	<i>36.74</i>		
Economic and sociocultural status (ESCS)					
A. Top 33%	540.88	515.80	511.28	<i>25.08</i>	<i>29.60</i>
B. Lower 33%	480.71	469.82	421.93	<i>10.89</i>	<i>58.78</i>
<i>Difference (A-B)</i>	<i>60.17</i>	<i>45.98</i>	<i>89.35</i>		

Source. Self-elaborated from PISA-2015. In italics, performance gap between subgroups.

Table A.2. Determinants of performance in the scientific, reading and mathematical competencies of 15-year-old students in Catalonia; model with interactions.

Hierarchical linear models. PISA-2015

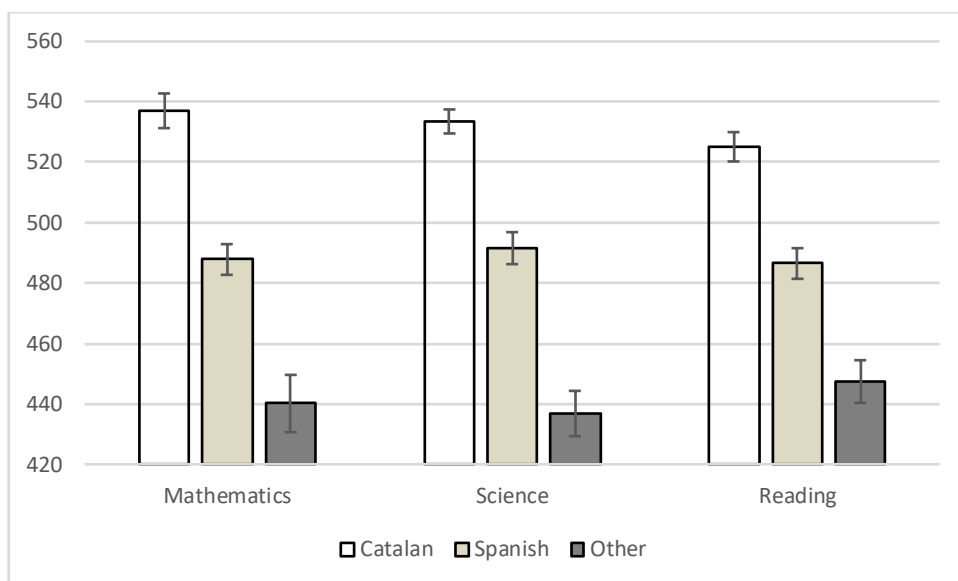
	Science	Reading	Mathematics
Intercept	439.20*** (132.83)	423.03*** (145.31)	460.47*** (120.62)
Individual and household level variables			
- Sex (ref: males)	-26.06*** (4.68)	-2.70 (6.78)	-27.85*** (5.53)
- Age (years)	9.99 (8.38)	8.89 (9.23)	7.05 (7.58)
- Economic and sociocultural status (ESCS index)	16.90*** (3.88)	15.22*** (4.42)	16.64*** (3.88)
- First generation immigrant (ref: native)	-3.52 (11.98)	4.13 (13.36)	-11.30 (11.42)
- Second generation immigrant	-6.95 (11.94)	0.75 (13.48)	-7.57 (10.54)
- Language at home: Spanish (ref: Catalan)	-40.83*** (10.17)	-35.95*** (10.33)	-30.35** (9.75)
- Language at home: Other languages	-29.97** (30.38)	-28.07*** (32.81)	-25.61*** (27.34)
- Preschool attendance (number of years)	4.73** (1.86)	4.83** (1.88)	4.76** (2.11)
- Grade repetition during primary school (ref: no repetition)	-80.97*** (8.72)	-82.41*** (9.22)	-72.88*** (8.69)
- Grade repetition during secondary school (ref: no repetition)	-43.06*** (6.34)	-43.44*** (7.17)	-45.11*** (7.54)
- Changed school during primary school	-12.40** (5.85)	-11.39* (6.75)	-9.18* (5.28)
- Changed school during secondary school	-25.74*** (6.23)	-27.91*** (6.05)	-24.15*** (7.80)
School level variables			

- City size (10,000 to 100,000)	-41.37***	-30.97***	-30.07**
(ref. more than 100,000 inhabitants)	(8.96)	(8.86)	(8.61)
- City size (below 10,000 inhabitants)	-22.72*	-16.58	-16.81
	(11.81)	(10.86)	(11.01)
- School size (number of students)	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)
- Students to teacher ratio	-1.23	-0.83	-1.02
	(1.02)	(1.14)	(1.02)
- Class size (number of students)	0.13	0.18	0.29
	(0.52)	(0.47)	(0.43)
- School ownership: private publicly-funded (ref: public school)	-41.51***	-30.13***	-31.29***
	(9.54)	(10.28)	(9.72)
- Average ESCS of school	30.90***	24.56**	25.39***
	(9.64)	(9.54)	(8.47)
- Immigrant students' density in school >30%	-16.78**	-15.18**	-11.70
	(7.84)	(7.73)	(7.42)
- Percentage of students with Catalan as their mother tongue: between 30% and 60%	4.04	3.26	2.04
	(7.60)	(7.71)	(7.22)
- Percentage of students with Catalan as their mother tongue: <30% (ref: schools with at least 60% of students with Catalan as their mother tongue)	4.96	1.58	4.41
	(10.58)	(10.12)	(10.13)
Interaction effects			
- Spanish at home x Sex	10.30	15.50*	3.46
	(6.47)	(8.17)	(6.69)
- Other language at home x Sex	-5.90	-4.07	0.62
	(31.29)	(32.03)	(26.93)
- Spanish at home x ESCS	-10.02**	-9.07*	-9.46***
	(5.12)	(5.16)	(4.85)
- Other language at home x ESCS	-11.07	-14.84	-11.59
	(10.57)	(11.36)	(9.39)
- Spanish at home x Town	27.64**	17.31	23.89**
	(11.35)	(11.20)	(10.64)
- Other language at home x Town	0.20	-16.28	-3.69

	(25.56)	(36.78)	(35.37)
- Spanish at home x Village	23.18*	12.43	23.05*
	(12.55)	(13.30)	(12.29)
- Other language at home x Village	-25.18	-24.86	-22.13
	(25.78)	(30.18)	(24.40)
- Spanish at home x School ownership	33.14***	29.60**	27.30***
	(12.67)	(11.69)	(10.00)
- Other language at home x School ownership	39.93	60.15	20.40
	(41.55)	(45.13)	(45.23)
N	1,347	1,347	1,347

Source. Self-elaborated from PISA-2015. ***, statistically significant at 99%, ** statistically significant at 95%, * statistically significant at 90%. Standard errors displayed in parenthesis.

Figure 1. Average performance in Catalonia at age 15 in the three competencies assessed by PISA-2015, by mother tongue of the students



Source. Self-elaborated from PISA-2015.