Sussex Research Online

Education, language and identity

Article (Accepted Version)

Clots-Figueras, Irma and Masella, Paolo (2013) Education, language and identity. Economic Journal, 123 (570). F332-F357. ISSN 0013-0133

This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/42501/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

http://sro.sussex.ac.uk

Education, Language and Identity*

Irma Clots-Figueras† and Paolo Masella‡

This Version: October 3, 2012

Abstract

The process of individual identity formation is still an enigma, as is the capacity of public bodies to intervene in it. In 1983, the Catalan education system became bilingual, and Catalan, along with Spanish, was taught in schools. Using survey data from Catalonia we show that respondents who have been exposed for a longer time period to teaching in Catalan have stronger Catalan feelings. The effect also appears to be present among individuals whose parents do not have Catalan origins; in addition the reform affects political preferences and attitudes towards the organization of the State.

1 Introduction

> "Of all political questions, that [of education] is perhaps the most important. There cannot be a firmly established political state unless there is a teaching body with definitely recognized principles. If the child is not taught from infancy that he ought to be a republican or a monarchist, a Catholic or a free-thinker, the state will not constitute a nation; it will rest on uncertain and shifting foundations; and it will be constantly exposed to disorder and change". [Napoleon][1]

Education has often been invoked as an important instrument of nation building and State power consolidation. History provides some examples: in his lectures delivered in Berlin (Ad-

*Corresponding author: Irma Clots Figueras, Department of Economics, Universidad Carlos III de Madrid, Calle Madrid 126, 28903 Getafe (Madrid), Spain. Email: iclots@eco.uc3m.es/ irmaclots@gmail.com. The authors would like to thank the editor, Antonio Ciccone, and two anonymous referees for their very valuable comments and suggestions. They would also like thank Oriol Aspachs, who for professional reasons did not continue with the project. They also thank Manuel Arellano, Oriana Bandiera, Robin Burgess, Antonio Cabrales, Francesco Caselli, Maitreesh Ghatak, Luigi Guiso, Eliana La Ferrara, Ramón Marimón and seminar participants at the London School of Economics, Universidad Carlos III de Madrid, Universitat Pompeu Fabra, EUI, CEMFI, Universita' Bocconi, Universita' di Pisa, IMT Lucca, University of Bristol, University of Mannheim, University of Warwick, Universita' di Bologna, University of Essex, Queen Mary, University of London, IAE-CSIC, University of Cambridge and University of Edinburgh. The authors also thank Olympia Bover for the migration data provided and participants at the 2007 EEA meeting, the 2007 SAE, the 2008 annual RES conference, the 2008 ESPE conference and the 2008 NASM of the Econometric Society. This paper combines material presented in two working papers (Aspachs et al. 2007a and Aspachs et al. 2007b). A previous version of this paper circulated under the name "The Effect of Language at School on Identity and Political Outlooks". Irma Clots-Figueras gratefully acknowledges financial support from the MEC grants SEJ2004-07861 and SEJ2007-67436.

†Universidad Carlos III de Madrid

‡University of Sussex.

[1] Quoted in Reisner (1922), p35.

dresses to the German nation, 1807), Fichte explicitly emphasized the importance of state-controlled mass education in teaching Germans to be good Germans and to create the unified national sentiment needed to restore Prussian power. Between 1817 and 1825, new taxes were imposed to revitalize the Prussian school system and, as a result, every Prussian land was required to have primary schools.[2] Similarly, in 1833, the French Minister of Public Instruction, Francois Guizot, introduced a law that obliged every commune or group of neighboring communes to set up and maintain at least one elementary school. As argued by Weber (1976), the school, and in particular the village school, spreading the knowledge and the use of the French language, was a determinant factor in the acculturation process that made the French people French. In the second half of the 19th century, primary education was at the center of political debate in Italy as well. In 1859, the Italian schooling system was completely reorganized and strengthened in order to break the monopoly of the Catholic Church and to build the foundations of a modern liberal State. Alesina and Fuchs-Schundeln (2007) also discussed how indoctrination in public schools might explain the effect of Communism on the preferences of Eastern Germans towards state intervention.

Nation-building policies (such as the implementation of a sole national language in schools and across the entire territory of a State) as well as multicultural policies that explicitly recognize cultural differences (regional autonomy, the devolution of powers, or the use of multiple languages in schools and in other contexts) have been often proposed as sources of conflict management in ethnically divided societies. Several social scientists (Anderson, 1983; Bates, 1983; and Horowitz, 1985) have argued that the boundaries of ethnic groups and the strength of ethnic ties are not exogenous and change over time. They are likely to be affected by social and economic conditions and to depend on a number of economic and policy choices. To date, however, only limited research has been carried out to examine the extent to which ethnic identification can be shifted by cultural policies and regulation. We know surprisingly little about

[2] Ramirez and Boli (1987) study the political origins of mass schooling extensively.

whether government policies in general, and educational systems in particular, can enhance national cohesion and revitalize the national sentiment of a country.

Our paper takes a step in this direction by analyzing the effect on the process of identity formation of a particular educational policy implemented in the Spanish region of Catalonia. Up to 1983, Spanish was the official language of the Catalan education system. That year the education system became bilingual, and both Catalan and Spanish were used as languages of instruction.

Using survey data from Catalonia and exploiting variation in the number of years of compulsory education under Catalan instruction, results show that individuals who experienced greater exposure to teaching in Catalan are more likely to say that they feel more Catalan than Spanish. As is the case for most of the reforms involving changes in the languages of instruction, the introduction of bilingualism in Catalan schools was associated with other adjustments in the educational system, such as changes in textbooks and course contents. Although the existing data do not allow us to disentangle the language effect from the effect of these additional features, which are extensively discussed in the paper, with our identification strategy we deliver a global evaluation of the 1983 reform which is broadly comparable to similar linguistic reforms implemented in other educational systems across the world.

Our identification strategy relies on variation between cohorts in exposure to the Catalan language at school. The larger the number of years of schooling the individual has received in Catalan, the more affected he or she will be, but given that the number of years of schooling is typically an individual choice that could also be related to identity, we consider the number of years of compulsory education under Catalan teaching as the main measure of exposure to the reform. This number of years is not an individual choice and varies across cohorts. We then include polynomial functions of the year of birth in the regressions to account for cohort or age effects. A number of factors lead us to believe that the link between the educational reform and individual identity is causal. Several robustness checks allow us to conclude that the results are

neither driven by cohort specific trends on Catalan identity nor by the impact on identity of an other important event such as the end of the Franco dictatorship. We also do not find evidence that the reform affected migration flows or changed the composition of the Catalan population.

In the second part of the paper we present a set of additional results and discuss extensively how a change in the language of instruction can have important economic implications.

First, we show how the educational reform has had an effect on political preferences. In ethnically divided societies language at school can affect economic outcomes by changing political preferences, therefore influencing the selection of politicians and the policies they implement. Since in ethnically divided societies individual identity is likely to affect the political choices of individuals, the reform could have increased the vote share of ethnic parties. Indeed, we find that individuals who have experienced greater exposure to teaching in Catalan are more likely to declare that they chose a party with a Catalanist (i.e., Catalan regionalist) platform. Results suggest that the reform increased the salience of the ethnic issue in the Catalan society and helped consolidate a political system organized along ethnic lines. Moreover, changes in the language of education could also have an impact on preferences regarding the organization of the State, which could influence the probability of having an open conflict. As a further step, we then study the effect of the 1983 reform on separatist attitudes and we find that respondents affected by the reform are more likely to claim that Spanish regions should have the right to be independent States.

Second, we discuss how the introduction of compulsory bilingual educational system could affect the degree of integration of minorities within a country, which it is likely to be related to their economic outcomes. The 1983 reform can be interpreted as an example of nation-building policy within Catalonia, where individuals living in Catalonia but who were born elsewhere or whose parents do not have Catalan origins are the relevant minority. We find that the effect of the reform on individual identity also appears to be present among respondents educated in Catalonia after the reform but whose parents do not have Catalan origins. Therefore, if both

family and school can affect identity, this result suggests that nation-building policies indeed tend to promote a common national feeling and that education, through language, can partially balance the role of family in the cultural transmission process.

This paper is connected to an emerging literature within economics that explicitly focuses on the identity formation process. In an influential series of papers, Akerlof and Kranton (2000, 2002, 2005) and Akerlof (2007) incorporate identity, a person's sense of self, into an economic model of behavior. They then present several applications of their theory in the fields of the Economics of Education, the Economics of Organizations and Macroeconomics. Several theoretical studies followed the work by Akerlof and Kranton, including research by Benabou and Tirole (2011) and Bisin et al. (2006). Eckel and Grossman (2005), Charness et al. (2006), Mc Leish and Oxoby (2007), Heap and Zizzo (2009) and Chen and Li (2009) provide laboratory experiments that show how group identity affects individual behavior. Ethnic and national identity, by influencing the support for local and national governments, may also have important repercussions on the level of tax evasion and on whether individuals decide to buy public debt issued by governments. Indeed, Li (2011) provides evidence that strong levels of national identity stimulate tax morale.

However, the number of empirical studies which study the determinants of individual identity and in particular how educational policies affect identity remains limited. In Aspachs et al. (2008) we cite results presented in this paper, compare individuals in Catalonia and in the Basque Country to focus on the interaction between family and schooling in shaping identity and suggest that students are sorted into Basque or Spanish classes according to the identity of their parents. That summary piece is the first to discuss the correlation between exposure to language of instruction and identity; however it does not take into account that the number of years of education is typically an individual (or parental) choice, so it does not provide estimates of the causal effect. This paper is the first to provide evidence on the casual impact of Catalan instruction on individual identity by exploiting variation between cohorts in the number of

years of *compulsory* education received in Catalan language. Moreover, it also contributes to the literature by considering and discussing the economic implications of the observed shift in national attachment.

The paper is also related to the literature on the mechanisms of cultural transmission and the effects of culture on economic outcomes. Culture and trust explain differences in financial habits (Guiso et al., 2004), in living arrangements in Western Europe (Giuliano, 2007) and also in GDP per capita across European regions (Tabellini, 2007). If we then consider identity as part of our cultural background, our paper shows that education affects individual cultural attributes.[3] This result seems to be related as well with part of the literature on endogenous preferences. Bowles (1998) extensively discusses how political and economic institutions shape preferences through their effects on social norms, cultural transmission or through other channels. If identity is part of the utility function, our paper isolates a particular institutional arrangement (bilingual education) that is able to influence individual preferences.

Finally, this work is linked to a panoply of papers that study the relationship between ethnic diversity and economic and political outcomes. Easterly and Levine (1997) claim that high levels of ethnic fragmentation were at the root of Africa's growth tragedy. La Porta et al. (1999) show that ethnic heterogeneity is negatively correlated with the quality of government and, more recently, Montalvo and Reynal (2005) and (2006) explore the channels through which ethnic diversity influences economic development. They find that ethnic fractionalization lowers the rate of investment, while ethnic polarization increases the probability of civil wars. In contrast to those papers, our study tries to understand the roots of individual ethnic identification.[4]

The structure of the paper is as follows: Section 2 describes the reform and more generally the institutional setting in Catalonia. Then the data and the empirical strategy employed are discussed. Section 3 reports the main results and Section 4 provides several robustness checks.

[3] Aghion et al. (2010a) and (2010b) investigate how government regulation interferes with culture, social capital and norms of cooperation.

[4] Miguel and Posner (2005) and Masella (2011) study the relationship between ethnic diversity and ethnic and national identity.

Section 5 discusses the possible economic implications of educational changes that are able to shift identities and national attachment. Finally, Section 6 concludes.

2 Background and Empirical Strategy

2.1 *Background*

Catalonia is a region in northeast Spain. Catalan is a language that evolved from vulgar Latin in the 9th century on both sides of the eastern part of the Pyrenees. Its territorial expansion went hand in hand with the expansion of the Catalano-Aragonese Crown, which was established in other Spanish regions such as Valencia and the Balearic Islands, the south of France and the town of Alghero (in Sardinia, Italy). These are the regions where Catalan is nowadays present with more than 9 million speakers, making it one of the most important regional languages in the European Union. Catalonia is the region that hosts most of them, with more than 6 million speakers.

Catalan has not always been the main language in schools in Catalonia. From the middle of the twentieth century, when education became compulsory, it stopped being taught and all education was received in Spanish. During the Franco dictatorship (1940-1975) its use was restricted and the Spanish language was the only one used as an official language and the only language used in education. Once the dictatorship ended, Spain experienced a transition from a centralized to a politically decentralized system. Regions acquired the status of autonomous communities and were allowed to have their own regional parliaments. One of the most important laws approved by the new Catalan government was the "Catalan Law of Language Normalization" (1983), whose main goal was to promote the use of the Catalan language.

In one of its most relevant parts, the new law allowed the educational system (previously all in Spanish as an inheritance from the Franco period) to become bilingual. Although schools had been able to teach the Catalan language as a subject since 1978, until 1983 it was not

recognized as the main language to be used in education. All students, irrespective of their origin, were supposed to use both Catalan and Spanish at the end of their education as the law made it clear that the Certificate of Basic Educational Attainment could not be obtained without proving proficiency in both languages. The law also stated that students should not be allocated in different classes because of language differences and Catalan had to be introduced following the students' learning process. The law referred to all education levels below university education, including secondary, primary and pre-school education. However, Article 14 of the law states that in preschool, at the beginning of their education, students should be able to receive education in Spanish or in Catalan, according to their mother tongue.

Since the Catalan education system had been Spanish-based for such a long time, the transition to a bilingual system was designed to be smooth. Several subsequent decrees and orders explain how the reform was to be introduced in the first years. Students who were in primary or secondary education during the academic year 1983/1984 have been affected by the reform; starting from that academic year at least one subject had to be taught in Catalan as required by the order passed the 5th of August 1982 and, then, the presence of Catalan in schools had to be progressively enhanced in subsequent years as clearly stated by Articles 8 and 9 of Decree 362/1983. For secondary education, the increased presence of the Catalan language in each school had to occur in accordance with the students' prior knowledge of Catalan, in order to minimize its effects on the normal learning process.

Like most of the reforms involving changes in the languages of instruction, the introduction of bilingualism in Catalan schools was associated with other adjustments in the educational system. A direct implication of the linguistic policy is the translation of old textbooks and teaching materials and the adoption of new ones written in the Catalan language. Several orders related to the 1983 law also clarified the contents of the courses that had to be taught in schools: Catalan as well as Spanish culture, history and geography had to be taught in schools at all educational levels. However, as discussed by Siguan (1991), the reform did not

involve a substantial replacement of the teaching corps. Hundreds of teachers benefited from training schemes in the Catalan language and, although tests of knowledge of Catalan were also applied in the recruitment of new teachers, those who failed such a test had to commit to becoming proficient in the Catalan language within a few years. Only since 1989 have tests been compulsory and eliminatory.[5] As already mentioned in the Introduction, although the identification strategy adopted throughout the paper does not allow us to disentangle the pure effect of the introduction of bilingualism from the standard additional features of most of the policies that modify the language of instruction, such as changes in textbooks, course contents and even teachers' attitudes, our empirical exercise delivers a global evaluation of the 1983 reform of the Catalan educational system, which is broadly comparable with similar linguistic reforms implemented in other educational systems (among others, the ones implemented in Quebec (1977), Morocco (1983) and South Africa (1994)).

The Catalan Law of Normalization, however, involved other changes, as the goal of the law was to increase the use of the Catalan language and make it an effective communication vehicle. In fact, it clearly established that citizens had the right to use Catalan and encouraged its use by the media. These changes could represent a threat for our identification strategy since individuals belonging to younger cohorts have been spending a longer part of their life in an environment with a greater use of Catalan language. We try to address this issue by including in the regression analysis a flexible function of the cohort of birth and performing several robustness checks that seem to disregard the possibility that the results presented are driven by the inclusion of younger respondents that have been more exposed to Catalan culture and history.

[5] Therefore our results are unlikely to capture the fact that new teachers with stronger Catalan feelings have been hired because of the reform and have been transmitting their beliefs to students (see Saint-Paul 2010 for a theoretical discussion of such a mechanism).

2.2 *Data and Empirical Strategy*

In our empirical analysis we use data provided by "Centro de Investigaciones Sociologicas" on political and social attitudes of residents in Catalonia. The survey was conducted in 2001. To identify individuals' national attachment to Spain or Catalonia, we rely on the following question: "With which of the following sentences do you identify with more? (1) I feel only Spanish, (2) I feel more Spanish than Catalan, (3) I feel as Spanish as Catalan, (4) I feel more Catalan than Spanish, (5) I feel only Catalan". From this survey question we derive the two main variables we will be using as dependent variables in our empirical analysis: (i) *Identity,* which is defined over the range 1-5 and takes the value 1 if the individual reports feeling only Spanish, and 5 if she/he reports feeling only Catalan (ii) *IdentityD,* obtained by converting the survey question into a dummy variable that is equal to 1 if the respondent feels only Catalan, more Catalan than Spanish or as Spanish as Catalan and 0 otherwise.

Let us now discuss the variables we use to measure the degree of exposure to the reform and therefore to Catalan instruction. We start by defining *years of exposure* to the reform as the number of years of education received in Catalan language. Thus, *years of exposure* depends on whether an individual was in primary or secondary education after the law was implemented in 1983, which depends on his/her year of birth and the number of years of education. Respondents born in or after 1970 were at least partially affected by the 1983 reform, while respondents born between 1966 and 1969 received education in Catalan only if they attended secondary education.[6]

As discussed in the background section, although the reform may have also affected the language of teaching during the preschool time, preschool in Catalonia is not compulsory and parents can choose the language of education of their children; moreover, we do not have information on the number of years of preschool attended in Catalonia. In one of the specifications

[6] Note that the length of primary and secondary education in Spain is 8 and 4 years respectively and children start school at the age of 6. We do not consider university education since the law we study did not affect it.

presented in the next section, however, we introduce a new measure of exposure, *years of exposure from preschool,* that assumes that all the respondents attended 3 years of preschool in Catalonia. This measure, therefore, takes into account the possibility that the identity of respondents who were born after 1977 may have been affected by the fact that they received part or all their preschool education in Catalan language.

The number of years of schooling is, however, typically an individual choice (or a parental one) and may have been affected by the introduction of the educational reform. Respondents with very intense Spanish feelings, for instance migrants from another region of Spain, might have felt less comfortable about receiving education in Catalan and decided to drop out from school as a consequence of the reform. Individuals with very intense Catalan feelings would also be more likely to stay at school longer if the education was in Catalan. Thus, coefficients of a regression of identity on our measure of *exposure* would be biased upwards. In other words, the length of *exposure* to Catalan education could have potentially been determined by the identity of the respondent or by the strength of Catalan feelings of his/her parents (which in turn are likely to be transmitted to their children).

In order to avoid biases related to the endogeneity of the length of *exposure* to the reform, in the main specification we rely on a measure of exposure that is not the result of an individual choice. To account for unobservables that could affect both schooling and identity, we define *years of compulsory exposure* to the reform as the number of years of *compulsory* education received in Catalan language. Until 1990 the length of compulsory education established by law was 8 years, that is only primary education was compulsory and students could have left school at the age of 14. In 1990, with the approval of the Law of General Ordering of the Educational System (LOGSE), the first two years of secondary education became compulsory as well, so students could have left school at the age of 16. *Years of compulsory exposure* is a variable that only depends on the year of birth of the respondent; each respondent belonging to the same cohort will be subject to the same amount of exposure to compulsory education.

Respondents from the 1970 cohort received 1 year of *compulsory exposure,* since only their last year of primary (compulsory) education was affected by the reform. In general, the length of *compulsory exposure* to the reform corresponds to the number of years of primary education received after the 1983 reform and therefore it is equal to 8 years for all the cohorts born in 1977 or afterwards.[7] We include in our sample respondents belonging to any cohort of birth (that is cohorts 1908-1983).[8] Therefore the respondents who did not receive any year of compulsory Catalan education are the ones who were born before 1970, the ones who received some years of Catalan education (from 1 to 7) are the ones born between 1970 and 1976, the ones who received all their primary (and compulsory) education in Catalonia are the ones born in or after 1977.

The econometric specification to be tested is then:

$$Identity_{ic} = \alpha + \beta Years\ of\ Comp.Exp._c + \mathbf{x}_{ic}\boldsymbol{\gamma} + f(c) + \varepsilon_{ic} \quad (1)$$

where the dependent variable $Identity_{ic}$ is the identity measure of individual i, from cohort c. $Years\ of\ Comp.Exp._c$ is the number of *years of compulsory exposure* to teaching in Catalan experienced by an individual belonging to cohort c, $f(c)$ are fourth order polynomials of the cohort of birth c to control for cohort trends and age effects, and $\mathbf{x}_{ic}$ is a vector of control variables, that include the gender of the respondent, his/her origin, the origin of his/her parents, his/her province of residence and the size of his/her town of residence. Robust standard errors are clustered at the cohort level to control for the fact that observations in a given cohort group may be correlated.

Our identification strategy relies on the implicit assumption that there is no other variable that affects how certain cohort groups feel about the Catalan identity issue. In the robustness

[7] The only notable exception is represented by the respondents born in 1983. They were subject to 10 years of compulsory education, which were all experienced after the reform. So individuals from the 1983 cohort were subject to 10 years of *compulsory exposure*.

[8] Since the survey includes information on when respondents arrived in Catalonia, we decided to exclude from the sample only migrants who did not receive any education inside Catalonia, as they would not be comparable to the rest (373 observations).

section of the paper, we discuss the fact that the composition of the sample might not be exogenous to the reform and we rule out the possibility that our estimates are simply driven by cohort trends in Catalan feelings or by the impact on identity of the end of the Franco dictatorship.

Table 1 provides the full set of descriptive statistics for the main variables used; details on the construction of each variable can be found in the data appendix. Figures 1 and 2 then provide preliminary evidence on the relationship between exposure to the reform and strength of Catalan identity. We start by realizing that for a set of cohorts (1969-1977) compulsory exposure to Catalan education increased linearly from 0 to 8 years. Older cohorts were not affected at all, while younger cohorts were fully affected by the 1983 reform (they all received all their compulsory education in Catalan language). We then expect the strength of Catalan identity to be increasing (not necessarily linearly) among cohorts 1969-1977, but not among younger/older cohorts. We provide graphical evidence that supports this conjecture. Figure 1 plots the average of the variable IdentityD by cohort for the sample of respondents born between 1960 and 1983; then we separate 3 groups, i.e. cohorts not affected by compulsory exposure, cohorts fully affected and cohorts 1969-1977, and, for each of the 3 groups, we plot a non-linear function, estimated using locally weighted least squares (LOWESS) to fit the curve.

TABLE 1 HERE

As it is possible to notice, while there seems to be a negative trend when considering cohorts fully affected or not affected by the reform (the younger respondents are, the weaker their Catalan feelings), this trend is stopped and partially reversed by the introduction of Catalan education. Catalan identity is a positive and concave function of the number of years of treatment; the effect of instruction in Catalan language becomes smaller once the respondent has already received few years of instruction in Catalan language.

Quite interestingly the reform seems to have already an impact on cohorts 1966-1969. This is because, as pointed earlier in the paper, respondents within those cohorts got their secondary

(and not compulsory) education under Catalan instruction. We then focus on the set of cohorts 1965-1977 (for this set of cohorts potential exposure during compulsory and non-compulsory education to Catalan education increased linearly from 0 to 12 years) and perform a similar exercise. Figure 2 shows very clearly how the reform reversed the negative age trend in Catalan feelings.

FIGURE 1 HERE

FIGURE 2 HERE

3 Baseline Results

We now turn to the empirical evidence. Table 2 shows estimates of equation (1) using *years of compulsory exposure* as the measure of exposure to the reform and controlling for the quartic order polynomial of the cohort of birth. In column 1 we report results obtained using *Identity* as the dependent variable, in column 2 using *IdentityD*. We then introduce the set of control variables discussed in the previous section: province of residence dummies, controls for size of town of residence, gender, family and individual origin. We classified respondents into 4 categories: 1) individuals who were not born in Catalonia 2) individuals who were born in Catalonia but whose parents were not 3) individuals who were born in Catalonia but with only one parent born in Catalonia 4) individuals who were born in Catalonia and whose parents were both born in Catalonia. Columns 3-4 report results for both the dependent variables using a specification that includes both the full set of control variables and a fourth order polynomial of the cohort of birth. Column 4 represents our baseline specification. We choose *IdentityD* as dependent variable in the baseline specification because its binomial nature makes the interpretation of the coefficients of interest easier. Table 2 offers quite a uniform picture. Exposure to Catalan instruction indeed fostered the intensity of Catalan identity. Estimates from column 4 suggest that one year of exposure to the reform increased the probability of feeling

only Catalan, more Catalan than Spanish or as Catalan as Spanish by more than 2 percentage points, that is about 7% of one standard deviation. The intensity of Catalan identity is, as is predictable, correlated with the origin of the interviewed and the origin of his/her parents as well. We find that Catalan identity is strongest among respondents who were born in Catalonia and whose parents were both born in Catalonia, and it is weakest among respondents who were not born in Catalonia.

TABLE 2 HERE

In Table 3 we provide some specification checks. We first run the baseline specification controlling for different functions of the birth cohort, in column 1 we use a third order polynomial and in column 2 a fifth order one. Results are similar to those obtained before; the coefficient when we control for a third order polynomial is, however, slightly smaller. Columns 3-5 present results obtained using our baseline specification but considering three different measures of exposure to Catalan teaching. First, in column 3 we use the variable *years of exposure*, that, as discussed in the previous section, could suffer from possible endogeneity problems; the coefficient obtained is, however, very similar to our baseline coefficient. Second, in column 4 we take into account the possibility that the reform may also have affected individual identity during preschool by using the measure *years of exposure from preschool,* which assumes that all individuals have attended 3 years of preschool. The estimated coefficient is very similar to the baseline coefficient. Finally, in column 5 we consider a binomial measure of treatment, the *compulsory exposure dummy*, which is a variable equal to 1 if the respondent has been subject to at least one year of compulsory education in Catalan language, and zero otherwise. Respondents exposed to the reform are 7.5 percentage points more likely to feel only Catalan, more Catalan than Spanish or as Catalan as Spanish.

The effect of the introduction of bilingual education in Catalan schools on individual identity can be analyzed using Akerlof and Kranton (2000) as a reference framework. In their framework the utility of an individual depends on his/her identity and individuals can choose to belong

to certain groups and therefore adopt the corresponding social identity and follow the related behavioral prescriptions. In this setting there are two alternative mechanisms through which education can affect individual identity. First, the reform may have reduced the level of effort that individuals have to spend in order to behave accordingly to the prescriptions of the Catalan group (i.e. have an adequate command of the Catalan language), therefore making more attractive the identification with such group. Alternatively, the reform may have increased the utility associated to identification with the Catalan group, for instance by changing the perception among students of the social status assigned to such group.[9] Students learn to associate Catalan language with schooling, which is likely to be perceived as the most important institution by children at that age.

TABLE 3 HERE

4 Robustness

In this section, we perform several robustness checks to confirm the validity of the identification strategy used. We first investigate whether our estimates are capturing the existence of cohort trends in Catalan feelings or the effect of the end of the dictatorship, and then we run some robustness checks to account for the fact that the reform could have affected migration flows (and then the composition of our sample).

Different cohorts might have been raised by parents with different values and preferences (the younger parents being more pro-Catalan), therefore our estimates could be the artifact of cohort specific trends in Catalan feelings. To rule out this possibility we perform several falsification exercises: we consider only the sample of respondents who were not affected by Catalan instruction either in compulsory or in secondary education, that is we consider only individuals who were born before 1966, and we assign a pseudo treatment to the youngest

[9] Similarly, using the Benabou and Tirole (2011) framework, we can interpret the increase in the strength of Catalan identification as a consequence of the rise in the perceived salience of the ethnic cleavage due to the exposure to Catalan language in schools.

cohorts as if the reform had been implemented in a year X rather than in 1983, where the year X is a generic year between 1939 and 1978. In this way we perform a battery of 40 falsification tests; the coefficients of the pseudotreatment of each of the 40 tests are displayed in Table 4. It is worth noting that when positive, the coefficients of the pseudo treatment variable are never significant and always smaller in size than the coefficient of our variable of interest as reported in our baseline specification in Column 4 of Table 2. Only one of the 40 coefficients is negative and significant at the 10% level, the one obtained when assuming the reform was implemented in 1976; the size of that coefficient is not negligible, but this is very likely to be due to the very small number of observations affected by the pseudo treatment (in that case the only cohorts considered as treated are 1963-1965).[10]

TABLE 4 HERE

We then test whether our results are driven by the fact that respondents exposed to Catalan education were also affected by the end of the Franco regime in 1975. The previous exercise may also be considered the first robustness check in this direction as it seems to disregard the possibility that the reform is driven by events that took place between 1975 and 1978. We conduct, however, two further robustness checks. First we add as a control in the baseline specification the total number of years exposed to democracy (we consider 1975 as the first year of democracy in Spain) before age 18. Results are reported in Column 1 of Table 5. Reassuringly, the coefficient of our exposure variable remains positive and significant. Results do not change if instead we consider as relevant years either 1978, when the reform of the constitution took place, or 1981, when a coup d'etat was attempted (see Columns 2 and 3). As a further exercise, in Column 4 we restrict our sample to respondents who started their education after the end

[10] We also perform another test in which we use only cohorts who received the same amount of treatment (8 years) and that therefore were born in 1977 or after. Within this very small sample, younger cohorts may have received a different level of exposure to Catalan culture, history, or language outside of school than older cohorts. In order to check whether this is actually what we are capturing in our results we assign a pseudotreatment to the youngest cohorts as if the reform had been implemented in a year X>1983 rather than in 1983 (in this set of exercises given the extremely limited number of cohorts we use linear cohort trends). We never find any evidence that these pseudo-reforms have a positive effect on the identity of the respondents. Results are available upon request.

of the Franco´s regime, that is we restrict the sample to cohorts born in 1970 or later. Results are robust to this additional check.

The appearance of Catalan media could also have affected our results; since the mid-1980s, a TV channel (TV3) has broadcast in the Catalan language and the diffusion of the daily press in Catalan has increased considerably. Younger cohorts therefore have been, in relative terms, more exposed to Catalan-language media. We cannot observe the lifetime exposure to Catalan-language media of respondents, but we do have information on their current behavior, which will most likely be correlated to past behavior. We know whether at the time of the survey the respondent watches Catalan TV, Spanish TV or both. Column 3 of Table 5 shows results when we control for current exposure to Catalan language TV: the coefficient of interest is still positive and significant (and also very similar in size).[11] It must be acknowledged, however, that controlling for current consumption of Catalan media is very problematic since, clearly, this variable can be affected by the treatment itself as the 1983 reform is likely to have stimulated proficiency in the Catalan language.

TABLE 5 HERE

The reform might have affected the composition of the Catalan population, as it could have changed migration patterns into and out of Catalonia. Schooling in Catalan could have implied an additional cost of migration to Catalonia for Spanish speakers. This additional migration cost could have been higher for potential migrants with very intense Spanish sentiments who, as a result, could have decided not to migrate towards Catalonia. Similarly, people with very intense Spanish sentiments could have decided to leave Catalonia after the reform. If this is the case, our results could be capturing a change in the composition of Catalan society rather than the effect of the reform.

As a first step, we rule out the possibility that the results we find are related to a change

[11] As a second exercise, we check whether the effect of the 1983 reform varies with the current exposure to Catalan language media. We introduce interactions between our measure of compulsory exposure to the language reform and a dummy that is equal to 1 if the respondent either watches Catalan TV or both Catalan and Spanish, and 0 otherwise. We find that the effect of the reform is homogenous and independent of current exposure to Catalan language TV. Results are available from the authors on request.

in the pattern of migration inflows in Catalonia. We restrict the sample to individuals whose parents were already in Catalonia when the reform was implemented. For these individuals, their parents' decision to migrate to Catalonia is less likely to be affected by the educational reform. Column 6 of Table 5 shows that the results are robust to this check.[12] However, it can still be the case that patterns of migration outflows, although very small in size, are affected by the reform and, at the same time, contribute to our results. Unfortunately, the survey we have does not allow for any checks that could help us in this direction. We then use other data sources to exclude the possibility that the size and the composition of migration outflows (inflows) have been affected by the 1983 reform of the educational system.

We then use "Residential Variation Data" from the INE (Instituto Nacional de Estadística). This is the only source of annual aggregate migration flows that goes back earlier than 1987, but allows us to provide some evidence that the size of migration outflows (inflows) were not affected by the introduction of bilingualism. We consider the following specification,

$$Y_{i,t} = \alpha Cat_{i,t} + \beta After_{i,t} + \gamma Cat_{i,t} * After_{i,t} + \varepsilon_{it}$$

where $Y_{i,t}$ are the size of outflows (inflows) in per capita terms of region i in year t (from 1978 to 1987), $Cat_{i,t}$ is a dummy variable equal to 1 if the observation refers to Catalan migration outflows or inflows and $After_{i,t}$ is a dummy variable equal to 1 if the observation refers to years from 1983 onwards (after the reform was implemented). In the first specification, we consider all the Spanish regions; in the second one, only Catalonia and the other 5 richest regions;[13] the coefficient of the interaction term (Table 6) is always very close to zero and not significant.[14] This suggests that the size of per-capita migration flows in Catalonia were not affected by the 1983 law.

[12] Results (available on request) do not change if we restrict the sample even further to control for the possibility that parents somehow anticipated the 1983 reform and then took the migration decision before 1983.

[13] Catalonia, The Basque Country, The Balearic Islands, Madrid, Navarra, and Valencia. These are the richest regions according to GDP per capita measures in 2001 (provided by INE).

[14] Results do not change if we use the logarithm as the dependent variable.

As a further robustness check, we consider individual data from the 2001 Spanish Census. For each individual interviewed we know the current province of residence, the previous province of residence (if any) and the year in which they moved (if they moved) from one province to the other. Results in Table 7 show the similarity of several characteristics of individuals who moved out of (into) Catalonia in 1982 (the first year before the reform) and those who moved out of (into) Catalonia in 1984 (the first year after the reform). We cannot reject the hypotheses of equality in age of migration, years of education, province of birth and other parental characteristics. This suggests that the composition of migration flows (both inflows and outflows) was not affected by the 1983 law.

TABLE 6 HERE

TABLE 7 HERE

5 Additional Results

In this section we discuss how the introduction of bilingual education might have repercussions on economic outcomes by affecting the political preferences and attitudes towards the organization of the State and the level of integration of minorities

Our findings can potentially be of general interest and not only related to the Catalan society. The number of languages spoken in the world is estimated to be between 6000 and 7000 and there are more than 20 States with more than one official language.[15] Several countries, in addition to the official language, recognize other national languages, occasionally also compulsory in education, and in many others, a variety of languages are widely used without having the legal authority of an official language.[16] Multilingual societies have often adopted more than one language of instruction. Canada and India have extensive experience in multilingual education; for four decades in India, there were at least three languages of instruction, the official language

[15] The choice of one language over another, however, has often been often an important and divisive issue. In 1956, in Sri Lanka, Sinhalese was made the official language, provoking a strong reaction by the Tamil minority, and until 1994, in Turkey the use of the Kurdish language in public was prohibited by law.

[16] See UNESCO (2003) and the UN report (2004) for a discussion of the topic.

of each Indian state plus English and Hindi. In Latin America, indigenous groups receive instruction in their own language and in the official language of the country, while in most Sub-Saharan African countries, children are educated in their local languages during the very first grades and in the colonial language (French, English or Portuguese) during the later stages of their education. Bilingual education has been a feature of schooling throughout the history of the United States. In 1968 the Bilingual Education Act, a provision of the Elementary and Secondary Education Act, authorized funds for programs for students who spoke languages other than English. Later on, the 1974 Supreme Court decision maintained that school programs exclusively in English denied equal access to education to nonnative students and determined that districts with such students had a responsibility to help them overcome their language disadvantage.

5.1 *Integration of Minorities*

Bilingual education can affect the integration of ethnic minorities. The large migration inflows experienced by Western countries in the last few decades and several episodes of social unrest involving ethnic minorities, first and second generation immigrants (ethnic riots occurred in US in 1992, UK in 2001 and, more recently in France in 2005), placed the issue of the cultural and economic assimilation of immigrants at the center of the political debate.

Recent economic literature investigated whether the cultural integration of immigrants, and more in general, members of ethnic minorities, is conducive to better performance in education and in the labor market. As formalized by Austen-Smith and Fryer (2005), part of the gap in academic achievements between blacks and whites in the United States may be explained by the adoption of "oppositional" identities by students belonging to the black minority. Learning standard English and performing well at school may be regarded as "acting white" and adopting a "white" identity and therefore punished by the members of the black community. Evidence in this sense is provided by Fryer and Torelli (2010). Algan et al. (2010) found that in France,

Germany and UK the educational gap between natives and immigrants is much smaller in the second as compared to the first generation. This finding may be partially explained by the higher level of integration of second generation immigrants. Mason (2004) documents that Hispanic American with a non-Hispanic white racial identity tend to work and earn more and similar findings are reported by Nekby and Rodin (2010) using Swedish data. In the UK a strong ethnic identity is associated with an employment penalty according to results obtained by Battu and Zenou (2010), while in Germany we have a a less clear picture: Constant et al. (2006) and Zimmermann et al. (2007) found that the probability of being employed is significantly lower for immigrants that are separated and marginalized, while Casey and Dustmann (2010) do not report any pronounced effect of ethnic identity on labor market outcomes. A broader view is proposed by Bisin et al. (2011) who, using a cross country framework, provide suggestive evidence that a strong ethnic identity is associated with a lower probability of being employed among non-EU immigrants in Europe.

Therefore we now try to understand whether the reform of the Catalan education system has had an impact on the minorities living in the Catalan region. We define as a minority group all respondents of non-Catalan origin or with parents who are not of Catalan origin (first and second generation immigrants). We then construct 2 dummy variables, creating two groups. The first group includes individuals not born in Catalonia and individuals born in Catalonia with both parents who were not born in Catalonia. The second group includes individuals born in Catalonia for whom at least one of the parents was born in Catalonia. We then interact these two dummy variables with our exposure variable. Results in column 1 of Table 8 show that the reform had a positive impact also on the Catalan identity of respondents belonging to the minority group as defined previously. Nation-building policies (such as the implementation of a unique national language in schools and across the entire territory of a State) seem to promote the growth of a common national feeling.

The 1983 reform of the Catalan education system can, however, also be interpreted as an

example of a multicultural policy within Spain, where individuals living in Catalonia are the relevant minority. Catalan residents are indeed listed as an ethnic minority in many of the most important datasets that list the ethnic groups within each country (such as for instance the datasets provided by Fearon (2003) and Alesina et al.(2003)). According to such interpretation, we can conclude that multicultural policies instead tend to favor the development of regional identities.

5.2 *Political Preferences*

The Catalan educational reform might also have affected economic outcomes by changing the political preferences of Catalan residents and therefore influencing the selection of politicians and the policies implemented by them. Indeed, Banerjee and Pande (2007) found evidence that stronger voter ethnicization, defined as a greater preference by a voter for the party representing his/her ethnic group, is associated with lower quality and more corrupt politicians. A governance outcome very widespread in severely divided societies (for instance Nigeria, Uganda and Burundi) is patronage, the system of granting benefits to members of a particular ethnic group while penalizing other ethnic groups. Alesina et al. (1999) and (2000) provide evidence that politicians in more fragmented US cities tend to lower the provision of public goods and increase targeted benefits such as public employment.

The CIS 2001 survey includes a question in which individuals are asked who they voted for in the 1999 elections for the Catalan Parliament. This feature of the data allows us to compare the voting behavior of individuals affected by the reform with the voting behavior of individuals not affected and then to perform a direct test on whether the reform increased the salience of the ethnic issue. In Catalonia, at the time of the survey there were mainly five political parties: Convergència i Unió (CIU), the party that was in power from 1980 to 2003, Partit dels Socialistes de Catalunya (PSC), Esquerra Republicana de Catalunya (ERC), Partido Popular (PP), and Iniciativa per Catalunya Verds (ICV). We consider parties to be Catalan if their manifestos

include the approval of a law that would give Catalonia the right to self-determination. These parties are CIU, ERC and ICV. Column 2 of Table 8 reports estimates for Equation (1), using as a dependent variable a dummy equal to one if the respondent voted for a Catalan regionalist party in the 1999 regional election. We restrict the sample to respondents who declared having voted during such elections. Results suggest that those who were more exposed to Catalan instruction tend to be significantly more likely to vote for Catalan parties; in fact, an extra year of compulsory education in Catalan increases the likelihood that the respondent declared that he/she voted for a Catalan party by 4 percentage points, which corresponds to 6% of the average.

We then use two other questions of the survey in which respondents are asked about their opinion on how the Spanish state should be organized and how strong the degree of independence of the Catalan region should be. In the first question respondents are asked whether they would like the Spanish State to be more centralized than it is now, as centralized as now, less centralized than now or whether Spanish regions should even have the right to claim independence. We generate a dummy that is equal to 1 if the respondent chooses the more extreme option, the one that allows regions to be independent States, zero otherwise. Similarly in the second question respondents are asked whether (i) the region of Catalonia should be part of a more centralized State, (ii) should have the same powers that it has now (iii) should increase its competencies (iv) should be able to become an independent State. Again, we create a dummy variable that is equal to one if the respondent selects the fourth alternative. Results are in column 3 and 4 of Table 8 and show that those affected by an additional year of instruction in Catalan during compulsory education are 2 percentage points more likely to say that Spanish regions, and Catalonia in particular, should have the right to ask for independence. The rise in separatist attitudes might affect economic outcomes as they increase the probability of an open conflict as extensively discussed by Horowitz (1985). Recent literature discusses the large effects of conflicts on economic outcomes (Besley and Muller (2011) and Eckstein and Tsiddon

(2004)); Abadie and (2003) quantify the economic penalties connected to the Basque conflict in terms of 10% of the Basque region GDP. Moreover, in Catalonia there has always been the widespread perception that, being among the richest regions of Spain, the amount of taxes paid to the central government is considerably higher than the amount of resources and services received from it. In fact, in a recent survey conducted by the Catalan government, 43% of the respondents were in favor of a referendum for independence and, among those who would vote yes to such referendum, 36.4% argued that they would do so to allow Catalonia to manage its own economic resources. The rise in local identity and separatist attitudes caused by the reform of the Catalan education system may have increased the pressure towards the central government to release fiscal powers by increasing the responsibilities of the Catalan government in terms of tax collection and public good provision (indeed reforms that increased the level of fiscal decentralization were approved in Spain in 1997 and 2001).

TABLE 8 HERE

6 Conclusions

Much has been said about the possibility that education can affect individual identity and preferences. However, to date, not many studies have been conducted on this matter. We have considered the 1983 educational reform, through which the Catalan education system became bilingual and Catalan as well as Spanish was taught in schools, and we have found a positive and large effect of this policy on Catalan identity. We then interpret the reform as an example of nation-building policy within Catalonia and try to build a bridge with the current literature on the integration of first and second generation immigrants. We find that the effect of the reform extends to respondents whose parents were not born in Catalonia and conclude that nation-building policies encourage the growth of a shared national sentiment among minorities. The survey used for Catalonia further allows us to show that the 1983 reform increased voting for Catalan parties and separatist attitudes. The shift in identity may be one of the channels

through which this takes place.

Nowadays, most countries in the world could be classified as multilingual. However, not all of them have multilingual educational systems. The UNESCO report "Education in a Multilingual World" explains that the choice of language of education constitutes an important challenge in the development of educational policies. In addition, mother tongue instruction is considered important for the quality of the education provided. Language can be regarded not only as a communication tool but also as an attribute of empowerment and cultural identity. Research of this sort along with the study of the effects of such reforms on political and economic outcomes can be highly relevant in ethnically divided societies where linguistic policies can be seen as mechanisms of conflict reduction. We have considered the 1983 educational reform, through which the Catalan education system became bilingual and Catalan as well as Spanish was taught in schools, and we have found a positive and large effect of this policy on Catalan identity. We need to stress, however, that our work does not fully separate the effects of the introduction of bilingualism from the other contemporaneous changes in the Catalan educational system, such as the revision of course contents and textbooks.

References

Abadie, A. and J. Gardeazabal (2003)'The economic costs of conflict: a case study of the Basque country,' *American Economic Review,* 93(1): 113-132.

Aghion, P. Y. Algan and P. Cahuc (2010), 'Can policy affect culture? Minimum wage and the Quality of Labor Relations', *Journal of the European Economic Association, Forthcoming*

Aghion, P. Y. Algan, P. Cahuc and A. Shleifer (2010), 'Regulation and distrust', *Quarterly Journal of Economics*, *Forthcoming*

Akerlof, G.A. (2007), 'The Missing motivation in Macroeconomics,' *American Economic Review* 97(1), 5-36

Akerlof, G.A. and R.E. Kranton (2000), 'Economics and identity,' *Quarterly Journal of Economics* 115, 715-753

Akerlof, G.A. and R.E. Kranton (2002), 'Identity and schooling: some lessons for the economics of education', *Journal of Economic Literature,* 40(4), 1167-1201

Akerlof, G.A. and R.E. Kranton (2005), 'Identity and economics of organization', *Journal of Economic Perspectives,* 19, 9-32

Alesina, A., R. Baqir and W. Easterly (1999), 'Public goods and ethnic divisions', *Quarterly Journal of Economics* 114, 1243-1284.

Alesina, A., R. Baqir and W. Easterly (2000), 'Redistributive public employment', *Journal of Urban Economics* 48, 219-241.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat and R. Wacziarg (2003), 'Fractionalization', *Journal of Economic Growth* 8(2), 155-194.

Alesina, A. and N. Fuchs-Schündeln (2007), 'Goodbye Lenin (or not?): the effect of Communism on people' *American Economic Review*, 97(4): 1507–1528

Algan Y., C. Dustmann, A. Glitz and A. Manning (2010), 'The economic situation of first and second-generation immigrants in France, Germany and the United Kingdom', ECONOMIC JOURNAL, vol. 120(542), pages F4-F30, 02

Anderson, B (1983). *Imagined Communities: Reflections on the Origins and Spread of Nationalism*, London, Verso.

Aspachs-Bracons, Oriol, Irma Clots-Figueras and Paolo Masella (2007a). 'Identity and language policies'. Universidad Carlos III Working Papers 07-77(46).

Aspachs-Bracons, Oriol, Irma Clots-Figueras and Paolo Masella (2007b).'Education and political behavior. Evidence from the Catalan linguistic reform'. Universidad Carlos III Working Papers 07-78(47).

Aspachs, O., I. Clots-Figueras, J, Costa and P. Masella (2008), 'Compulsory language educational policies and identity formation', *Journal of European Economic Association, Papers and Proceedings*, 6 (2-3), April-May

Austen-Smith D. and Fryer, Jr, R.D. (2005), 'An economic analysis of "acting white"',*Quarterly Journal of Economics*, vol.120, (May), pp.551-83.

Banerjee A. and R. Pande (2007), 'Parochial politics: ethnic preferences and politician corruption', Faculty Research Working Papers Series, John F. Kennedy School of Government

Bates, R (1983), *Modernization, Ethnic Competition and the Rationality of Politics in Contemporary Africa* in State versus Ethnic Claims: African Policy Dilemmas.

Battu, H. and Y. Zenou (2010), 'Oppositional identities and employment for ethnic minorities: evidence for England', ECONOMIC JOURNAL, 120, F52–F71.

Bénabou, R. and J. Tirole (2011), 'Identity, morals, and taboos: beliefs as assets', *Quarterly Journal of Economics*, vol. 126(2), pages 805-855.

Besley T. and H. Mueller (2011), 'Estimating the peace dividend:the impact of violence on house prices in Northern Ireland', *American Economic Review,* forthcoming

Bisin A, E. Patacchini , T. Verdier , Y. Zenou (2006), 'Bend it like Beckam': identity, socialization and assimilation', *CEPR Discussion paper* No 5662

Bisin, A., E. Patacchini, T. Verdier, T. and Y. Zenou (2011), 'Ethnic identity and labour market outcomes of immigrants in Europe', *Economic Policy,* 26: 57–92.

Bisin A and T. Verdier (2000), 'The economics of cultural transmission and the dynamics of preferences', *Journal of Economic Theory* 97, 298-319

Bover, O. and P. Velilla (2001), 'Migrations in Spain: historical background and current trends', *Banco de España Working Papers* No 9909

Bowles, S. (1978), 'Capitalist development and educational structure', *World Development*, 6, 783-796.

Casey, T. and C. Dustmann (2010), 'Immigrants' identity, economic outcomes, and the transmission of identity across generations', ECONOMIC JOURNAL*,* 120, F31–F51.

Charness, G., L. Rigotti and A. Rustichini (2007), 'Individual behavior and group membership', *American Economic Review,* 97, September, 1340-1352

Chen, Y. and S. Xin Li (2009), 'Group identity and social preferences'. *American Economic Review*, 99(1): 431–57.

Constant, A., L. Gataullina and K. F. Zimmermann (2006), 'Gender, ethnic identity and work', *IZA Discussion Paper* No 2420.

Easterly, W and R. Levine (1997), 'Africa's growth tragedy: policies and ethnic divisions', *Quarterly Journal of Economics*, Vol. 112, Issue 4

Eckel, C.E. and P.J. Grossman (2005), 'Managing diversity by creating team identity', *Journal of Economic Behavior and Organization*, 58(3), 371-392.

Eckstein Z. and D. Tsiddon (2004), 'Macroeconomic consequences of terror: theory and the case of Israel', *Journal of Monetary Economics*, Volume 51, Issue 5, July 2004, Pages 971-1002

Fearon, J. (2003) 'Ethnic and cultural diversity by country', *Journal of Economic Growth*, 8(2), 195-222.

Fernandez, R., A. Fogli and C. Olivetti (2004), 'Mothers and sons: preference formation and female labor force dynamics', *Quarterly Journal of Economics* 119 (4) pp. 1249-1299.

Fryer, R.G. Jr. and P. Torelli (2010), 'An empirical analysis of 'acting white'', *Journal of Public Economics* 94, 380-396.

Giuliano, P. (2007). 'Living arrangements in western Europe: does cultural origin matter?', *Journal of the European Economic Association, MIT Press*, vol. 5(5), pages 927-952, 09

Guiso, L., P. Sapienza and L. Zingales (2004), 'The role of social capital in financial development', *American Economic Review,* 94(3), pp. 526-556.

Hargreaves Heap, S., and D. Zizzo (2009), 'The value of groups', *American Economic Review*, 99(1): 295–323.

Horowitz, D. (1985), Ethnic Groups in Conflict. *Berkeley: University of California Press.*

La Porta, R., F. Lopez-de-Silanes, A. Shleifer and R. Vishny (1999), 'The quality of government', *Journal of Law, Economics and Organization,* 222-279

Li, S. (2011), 'Social identities, ethnic diversity, and tax morale', *Public Finance Review,* Vol. 38, No. 2, March 2010: 146-177

Mason, P. L. (2004), 'Annual income, hourly wages, and identity among Mexican-Americans and other Latinos', *Industrial Relations*, 43, 817–34.

Masella, P (2011), 'National identity and ethnic diversity', *Journal of Population Economics,* forthcoming

McLeish, K . and R . Oxoby (2007), 'Identity, cooperation, and punishment', *IZA Discussion Paper* No. 2572

Miguel T and D. Posner (2004), 'Sources of ethnic identification in Africa', *Afrobarometer Working Paper No. 44*

Montalvo J. and M. Reynal (2005), 'Ethnic diversity and economic development', *Journal of Development economics*, 76, 293-323

Montalvo J. and M. Reynal (2006), 'Ethnic polarization, potential conflict and civil wars', *American Economic Review* 95 (3), 796-816

Nekby, L. and M. Rödin (2010), 'Acculturation identity and employment among second and middle generation immigrants', *Journal of Economic Psychology,* 31, 35–50.

Ramirez F. and J. Boli (1987), 'The political construction of mass schooling: European origins and worldwide institutionalization', *Sociology of Education*, Vol. 60, No. 1. (Jan., 1987), pp. 2-17.

Reisner, E (1922), *Nationalism and education since 1789.* New York, Macmillan.

Saint-Paul, G (2010), 'Endogenous indoctrination: occupational choice, the evolution of beliefs, and the political economy of reform', ECONOMIC JOURNAL*,* vol. 120, n. 544, p. 325-353.

Siguan, M. (1991), 'The Catalan language in the educational system of Catalonia', *International Review of Education*, Volume 37, Number 1, March, 1991

Tabellini, G. (2007), 'Culture and institutions: economic development in the regions of Europe', *IGIER, Bocconi University, mimeo*

UN report (2004), 'Building multicultural democracy'

UNESCO (2003), 'Education in a multiligual world'

Weber, E. (1976), *Peasants into Frenchmen: the modernization of rural France, 1870-1914.* Stanford University Press.

Woolard, K (2003): *We don't speak Catalan because we are marginalized'; ethnic and class connotations of language in Barcelona.* In R.K. Blot, ed., Language and Social Identity. Westport, CT: Praeger Publishers, pp. 85-103.

Zimmermann, L., K. F. Zimmermann and A. Constant (2007), 'Ethnic self-identification of first-generation immigrants', *International Migration Review*, 41, 769–81.

7 Appendix

7.1 *Data Appendix-Definition of the variables*

-Identity: ordered variable which assumes the following values: (1) if the respondent answered "I feel only Spanish"; (2) if the respondent answered "I feel more Spanish than Catalan"; (3) if the respondent answered "I feel as Spanish as Catalan"; (4) if the respondent answered "I feel more Catalan than Spanish" and (5) if the respondent answered "I feel only Catalan". Source: CIS

-IdentityD: Dummy variable that is equal to 1 if the respondent answered "I feel as Spanish as Catalan", "I feel more Catalan than Spanish" or "I feel only Catalan", 0 otherwise. Source: CIS

-Catalan origin, Catalan family: dummy equal to 1 if the respondent answered that he was born in Catalonia and both his parents were born in Catalonia. Source: CIS

-Catalan origin, mixed family: dummy equal to 1 if the respondent answered that he was born in Catalonia and only one of his parents were born in Catalonia. Source: CIS

-*Catalan origin, non-Catalan family:* dummy equal to 1 if the respondent answered that he was born in Catalonia and neither of his parents were born in Catalonia. Source: CIS

-*Non-Catalan origin:* dummy equal to 1 if the respondent answered that he was not born in Catalonia. Source: CIS

-*Female:* dummy equal to 1 if the respondent is female. Source: CIS

-*Voting:* dummy equal to 1 if the respondent declares to have voted in the 1999 regional election. Source: CIS

-*Catalan Voting (IC+ERC+CIU):* dummy equal to 1 if the respondent declares to have voted for a Catalanist party in the 1999 regional election. We consider IC, ERC and CIU to be Catalanist parties. Source: CIS

-*Catalan TV: :* dummy equal to 1 if the individual currently watches Catalan TV or both Catalan and Spanish TV, 0 otherwise. Source: CIS

-*Outflows:* annual per capita outflows. Source: Residential Variation Data, INE.

-*Inflows:* annual per capita inflows. Source: Residential Variation Data, INE.

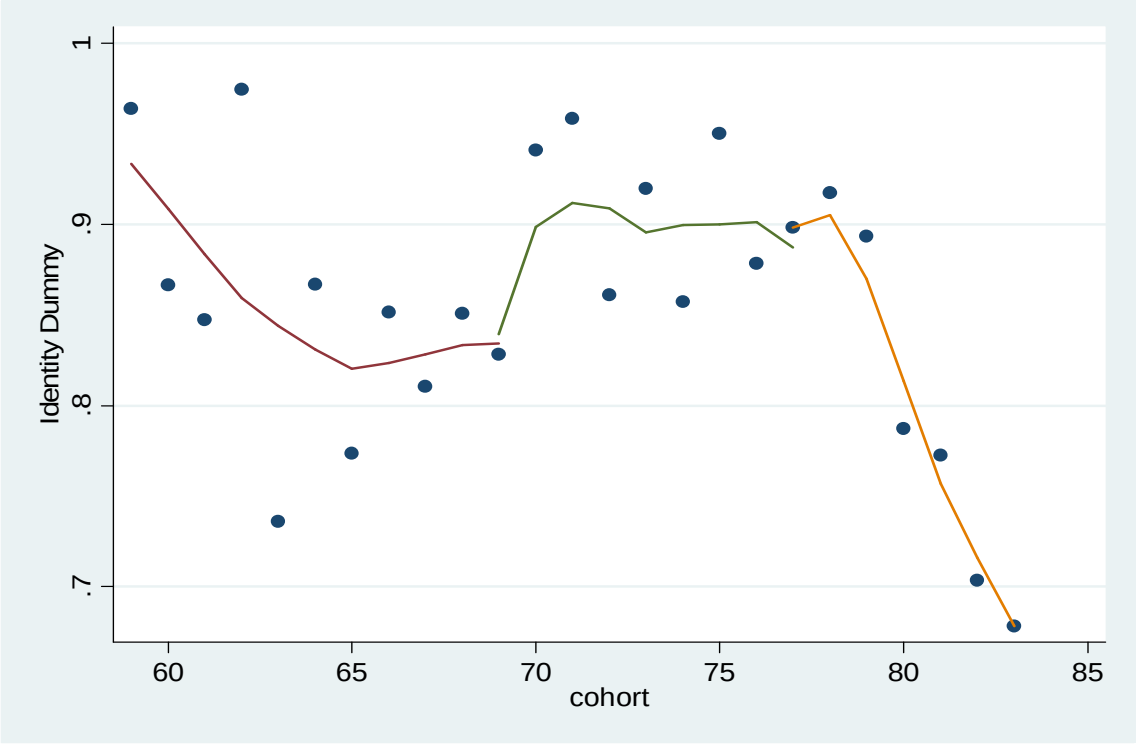

Fig. 1 : *Identity by cohort*

Note: Each dot represents the average of Identity(D) by cohort. The lines are lowess fits for 3 groups of cohorts: 1959-1969, 1969-1977 and 1977-1983.

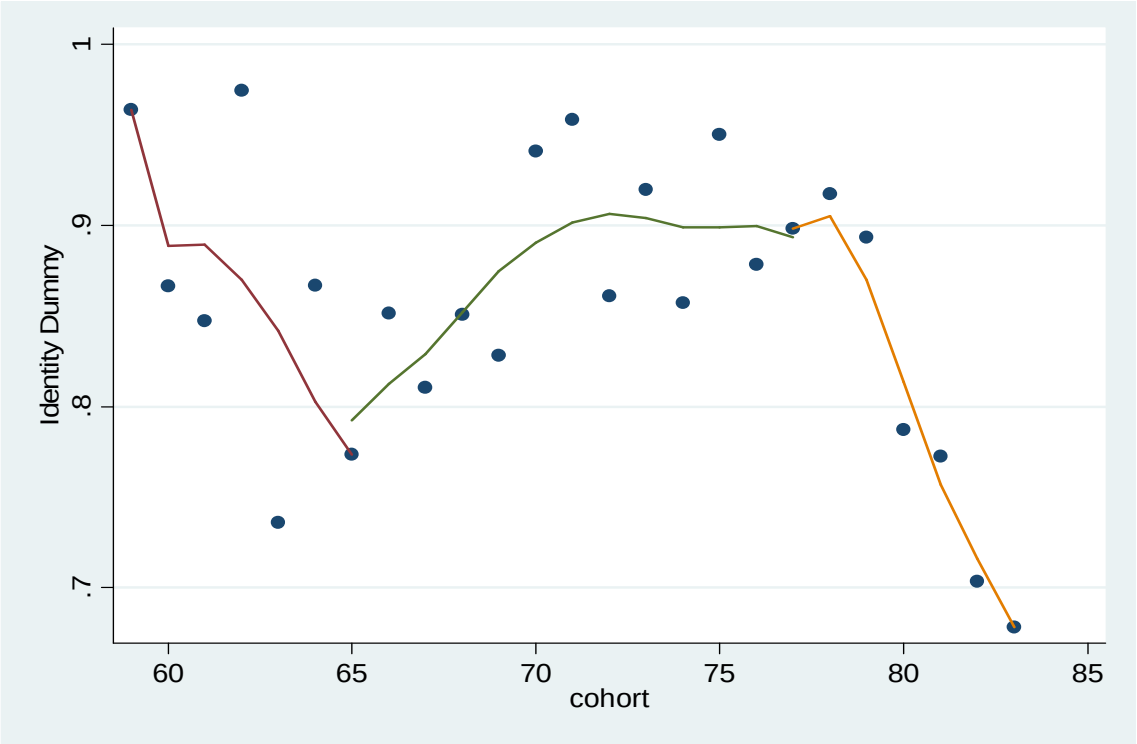

Fig. 2: *Identity by cohort(2)*

Note: Each dot represents the average of Identity(D) by cohort. The lines are lowess fits for 3 groups of cohorts: 1959-1965, 1965-1977 and 1977-1983.

Table 1*: Descriptive Statistics*

Variable	Obs	Mean	Std. Dev.
Identity	2309	3.5725	1.1296
IdentityD	2309	0.8727	0.3334
Catalan Voting	1888	0.6112	0.4876
Independence Preferences (1)	2263	0.2342	0.4236
Independence Preferences (2)	2282	0.2441	0.4296
Years of exposure to the reform	2292	2.5537	4.1753
Years of exposure during compulsory education	2309	1.7540	3.1291
Compulsory Exposure Dummy	2309	0.2811	0.4496
Years of exposure to the reform + preschool	2292	2.8538	4.7848
Years of exposure to democracy before 18	2309	6.4036	7.4054
Years of exposure to constitution (1978) before 18	2309	5.3274	6.9142
Years of exposure after coup d'etat (1981) before 18	2309	4.2369	6.1885
Watches TV in Catalan	2294	0.8923	0.3100
Age	2309	44.9861	18.0871
Female	2309	0.4829	0.4998
Province: Barcelona	2309	0.4153	0.4929
Province: Girona	2309	0.1958	0.3969
Province: Lleida	2309	0.2040	0.4030
Province: Tarragona	2309	0.1849	0.3883
Not born in Catalonia	2309	0.1589	0.3657
Born in Catalonia with parents born outside Catalonia	2309	0.1793	0.3837
Born in Catalonia with one parent born in Catalonia	2309	0.1308	0.3372
Born in Catalonia with both parents born in Catalonia	2309	0.5310	0.4991
Less than 2000 inhabitants	2309	0.1455	0.3527
Between 2001 and 10000 inhabitants	2309	0.2096	0.4071
Between 10001 and 50000 inhabitants	2309	0.2252	0.4178
Between 50001 and 100000 inhabitants	2309	0.1065	0.3086
Between 100001 and 40000 inhabitants	2309	0.1763	0.3811
More than 1000000 inhabitants	2309	0.1369	0.3438

Table 2: *Baseline Results*

	(1) Identity	(2) IdentityD	(3) Identity	(4) IdentityD
Years of exposure during compulsory education	0.0745***	0.0279***	0.0580**	0.0237***
	(0.0225)	(0.00680)	(0.0245)	(0.00706)
Born in Catalonia with parents born outside Catalonia			0.665***	0.301***
			(0.0890)	(0.0337)
Born in Catalonia with one parent born in Catalonia			1.138***	0.350***
			(0.0751)	(0.0309)
Born in Catalonia with both parents born in Catalonia			1.503***	0.403***
			(0.0716)	(0.0283)
Individual Controls			x	x
Polynomial	4th	4th	4th	4th
Observations	2309	2309	2309	2309
R-squared	0.011	0.017	0.318	0.222

Note: In Columns 1 and 3 Identity is the dependent variable. Identity is a variable that takes values between 1 and 5, 1 if the respondent answers he/she feels Only Spanish, 2 if he/she answers he/she feels More Spanish than Catalan, 3 he/she answers he/she feels As Spanish as Catalan, 4 he/she answers he/she feels More Catalan than Spanish and 5 he/she answers he/she feels Only Catalan. In Columns 2 and 4 IdentityD is the dependent variale. IdentityD is a dummy variable equal to one if Identity is equal to 3,4 or 5. Years of exposure during compulsory education is the number of years of compulsory education after 1983. Specifications 3 and 4 include dummy variables to control for gender, the province of residence and the size of town of residence, the origin of the respondent and the origin of the respondent´s parents. Robust standard errors are reported between parentheses and are clustered at the cohort level. * Significant at 10%, ** significant at 5%, *** significant at 1%.

Table 3: *Specification Checks*

Dependent variable: IdentityD	(1)	(2)	(3)	(4)	(5)
Years of exposure during compulsory education	0.0156*	0.0270***			
	(0.00818)	(0.00809)			
Years of exposure to the reform			0.0226***		
			(0.00545)		
Years of exposure to the reform+preschool				0.0254***	
				(0.00650)	
Compulsory Exposure Dummy					0.0768**
					(0.0332)
Individual Controls	x	x	x	x	x
Polynomial	3th	5th	4th	4th	4th
Observations	2309	2309	2292	2292	2309
R-squared	0.218	0.223	0.229	0.230	0.221

Note: Years of exposure during compulsory education is the number of years of compulsory education after 1983. Years of exposure to the reform is the number of years of education (compulsory and post compulsory) after 1983. Years of exposure to the reform+preschool takes into account that individuals in preschool after 1983 could have been affected by 3 additional years of education in Catalan. Years of exposure to the Reform Dummy is a dummy equal to one if the individual has received at least one year of education in Catalan during compulsory education. All specifications include dummy variables to control for gender, the province of residence and the size of town of residence, the origin of the respondent and the origin of the respondent´s parents. Robust standard errors are reported between parentheses and are clustered at the cohort level. * Significant at 10%, ** significant at 5%, *** significant at 1%.

Table 4: *Falsification Exercises: Reform in Different Years*

Coefficients: Years of exposure During Compulsory Education Dependent variable: Identity Dummy					
Year	coeff	se	Year	coeff	se
Placebo reform			Placebo reform		
1939	0.00299	(0.00899)	1959	0.00997	(0.0106)
1940	-0.00295	(0.00923)	1960	0.00973	(0.00828)
1941	-0.00540	(0.00924)	1961	0.00857	(0.00782)
1942	-0.00463	(0.00935)	1962	0.00343	(0.00803)
1943	-0.00569	(0.00854)	1963	-0.000438	(0.00881)
1944	-0.00745	(0.00869)	1964	-0.00135	(0.00988)
1945	-0.0106	(0.00749)	1965	-0.00108	(0.0113)
1946	-0.0110	(0.00746)	1966	0.00141	(0.0101)
1947	-0.0102	(0.00804)	1967	0.00473	(0.00989)
1948	-0.00824	(0.00872)	1968	0.00746	(0.0122)
1949	-0.00756	(0.00872)	1969	0.00203	(0.0173)
1950	-0.00952	(0.00785)	1970	0.0113	(0.0194)
1951	-0.00838	(0.00705)	1971	0.0191	(0.0238)
1952	-0.00765	(0.00667)	1972	0.0114	(0.0245)
1953	-0.00506	(0.00704)	1973	-0.00905	(0.0225)
1954	-0.000723	(0.00795)	1974	-0.0153	(0.0183)
1955	0.00375	(0.00917)	1975	-0.0259	(0.0180)
1956	0.00731	(0.0111)	1976	-0.0416*	(0.0228)
1957	0.0116	(0.0129)	1977	-0.0314	(0.0413)
1958	0.0122	(0.0120)	1978	-0.0625	(0.0435)

Note: IdentityD is a dummy variable equal to one if Identity is equal to 3,4 or 5. Years of exposure during compulsory education is the number of years of compulsory education after the reform year. Specifications include dummy variables to control for gender, the province of residence and the size of town of residence, the origin of the respondent and the origin of the respondent´s parents. We consider only the sample of respondents who were not affected by Catalan instruction either in compulsory or in secondary education, that is we consider only individuals who were born before 1966, and we assign a pseudo treatment to the youngest cohorts as if the reform had been implemented in a year X rather than in 1983, where the year X is a generic year between 1939 and 1978. For each possible placebo reform we perform our standard regression analysis and report the coefficient of the variable "years of exposure during compulsory education" in the table. Robust standard errors are reported between parentheses and are clustered at the cohort level. * Significant at 10%, ** significant at 5%, *** significant at 1%.

Table 5: *Robustness*

Dependent variable: IdentityD	1	2	3	4	5	6
Years exposure during compulsory edu	0.0262***	0.0273***	0.0235***	0.0403***	0.0235***	0.0206**
	(0.00766)	(0.0087)	(0.0087)	(0.0096)	(0.00700)	(0.00816)
Years of exposure to democracy (1975) before 18	-0.00774					
	(0.00591)					
Years of exposure to constitution (1978) before 18		-0.0058				
		(0.0061)				
Years of exposure after coup d'etat (1981) before 18			0.0004			
			(0.0091)			
Watches TV in Catalan					0.137***	
					(0.0342)	
Individual Controls	x	x	x	x	x	x
Polynomial	4th	4th	4th	4th	4th	4th
Sample				born in 1970 or later		restricted arrival
Observations	2309	2309	2309	649	2294	2276
R-squared	0.223	0.2226	0.2223	0.2375	0.239	0.200

Note: In the specifications in columns 1, 2 and 3 we control for the number of years of exposure to democracy (that is years of exposure after 1975) before the age of 18, the number of years of exposure after 1978 and before 18, and the number of years of exposure after 1981 and before 18, respectively. In column 4 we restrict the sample to those who started education after democracy in 1975. In column 5 we include a dummy variable that is equal to one if the individual watches TV in Catalan. In column 6 we restrict the sample to individuals who were already in Catalonia when the reform was implemented. All specifications include dummy variables to control for gender, the province of residence and the size of town of residence, the origin of the respondent and the origin of the respondent´s parents. Robust standard errors are reported between parentheses and are clustered at the cohort level. * Significant at 10%, ** significant at 5%, *** significant at 1%.

Table 6: *Other robustness checks: Migration.*

Dependent variable:	*Inflows*		*Outflows*	
	(1)	(2)	(3)	(4)
Catalonia	-0.135	-0.179*	-0.029	-0.040
	(0.089)	(0.096)	(0.070)	(0.067)
After1983	0.040	-0.002	0.034	0.051
	(0.031)	(0.055)	(0.024)	(0.039)
Catalonia x After1983	-0.106	-0.041	-0.032	-0.049
	(0.126)	(0.024)	(0.099)	(0.095)
SAMPLE	all Spain	5 richest	all Spain	5 richest
No. obs.	170	60	170	60
Rsq	0.061	0.15	0.017	0.073

Note: In Columns 1 and 2, *inflows* is the dependent variable. In Column 1, the sample consists of all the Spanish regions; in Column 2, it consists only of the 5 richest regions. In Columns 3 and 4, *outflows* is the dependent variable. In Column 3, the sample consists of all the Spanish regions; in Column 4, it consists only of the 5 richest regions. Robust standard errors are reported between parentheses. * Significant at 10%, ** significant at 5%, *** significant at 1%.

Table 7: *Other robustness checks: Migration (II).*

	Census 2001-outflows					
	1982		*1984*		*difference*	
	mean	*sd*	*mean*	*sd*	*diff*	*sd of diff*
Born in Catalonia	0.429	0.018	0.415	0.022	0.015	0.029
Age migrated	23.801	0.573	23.723	0.705	0.074	0.909
Years of Education	8.149	0.142	8.144	0.173	0.005	0.225
Years of Education-mother	6.164	0.262	6.556	0.323	-0.392	0.419
Years of Education-father	7.259	0.305	7.252	0.385	0.007	0.491
Male	0.497	0.018	0.478	0.022	0.019	0.029
Mother born in Catalonia	0.068	0.017	0.096	0.025	-0.028	0.029
Father born in Catalonia	0.069	0.018	0.042	0.018	0.027	0.027
	Census 2001-inflows					
	1982		*1984*		*difference*	
	mean	*sd*	*mean*	*sd*	*diff*	*sd of diff*
Born in Catalonia	0.168	0.017	0.205	0.018	-0.038	0.025
Age migrated	24.573	0.563	23.923	0.6171	0.6501	0.835
Years of Education	7.904	0.175	8.193	0.176	-0.289	0.248
Years of Education-mother	6.603	0.533	6.246	0.388	0.357	0.661
Years of Education-father	7.189	0.593	7.929	0.456	-0.739	0.742
Male	0.455	0.022	0.427	0.022	0.028	0.032
Mother born in Catalonia	0.254	0.055	0.279	0.041	-0.025	0.069
Father born in Catalonia	0.283	0.063	0.226	0.046	0.057	0.077

Table 8: *Additional Results*

	(1) IdentityD	(2) Catalan Voting	(3) Independence preferences (1)	(4) Independence preferences(2)
Years of exposure comp edu * At least one parent born Catalonia	0.0255***			
	(0.00744)			
Years of exposure comp edu * Individual or parents not born Catalonia	0.0213**			
	(0.00833)			
Years of exposure during compulsory education		0.0406***	0.0191**	0.0130
		(0.0139)	(0.00858)	(0.00908)
Controls	x	x	x	x
Polynomial	4th	4th	4th	4th
Observations	2309	1888	2263	2282
R-squared	0.223	0.179	0.070	0.079

Note: In the specification in Column 1 IdentityD is the dependent variable. We generate 2 dummy variables, creating two groups. The first group includes individuals born in Catalonia for which at least one of the parents was born in Catalonia. The second group includes individuals not born in Catalonia and individuals born in Catalonia with both parents who were not born in Catalonia. We then interact these two dummy variables with years of exposure during compulsory education. In the specification in Column 2 the dependent variable is a dummy variable which is equal to 1 if the respondent voted for a Catalanist party during the regional elections. The sample is restricted to individuals who voted during those elections. In the specification in Column 3 the dependent variable is a Dummy variable equal to 1 if the respondent claims Spanish regions should be allowed to become independent States. In the specification in Column 4 the dependent variable is a Dummy variable equal to 1 if the respondent claims Catalonia should be allowed to become an independent State. All specifications include dummy variables to control for gender, the province of residence and the size of town of residence, the origin of the respondent and the origin of the respondent´s parents. Robust standard errors are reported between parentheses and are clustered at the cohort level. * Significant at 10%, ** significant at 5%, *** significant at 1%.