

Occupational Selection in Multilingual Labor Markets: The Case of Catalonia

Núria Quella^a, Silvio Rendon^b

^{a,b}*Economics Department, SUNY-Stony Brook University,
Stony Brook, NY 11794, USA*

August 2009

Abstract.- In multilingual labor markets agents with high proficiency in more than one language may be selected into occupations that require high levels of skill in communicating with customers or writing reports in more than one language. In this paper we measure this effect in Catalonia, where two languages, Catalan and Spanish, coexist. Using census data for 1991 and 1996, and controlling for endogeneity of Catalan knowledge, we find that proficiency in speaking, reading, and writing Catalan reinforces selection into being employed, being an entrepreneur, and into white-collar occupations and communication-intensive jobs. In particular, being able to read and speak Catalan increases the probability of selection into white collar occupations by between 9 and 14 percentage points, while writing Catalan increases by 6 to 13 percentage points the probability of engaging in services, and government and educational activities.

JEL Classification: J61, J70, J31, I20.

Keywords: Language; Occupational Selection; Industry; Immigration; Skill Premium.

1 Introduction

There are activities and types of work that are clearly intensive in communication skills, namely providing customer service, writing reports, negotiating with trading partners, closing business agreements, coordinating work activities with co-workers, preparing presentations for prospective customers or colleagues, teaching, or doing paperwork. In multilingual labor markets, such as Quebec in Canada or Catalonia in Spain, language knowledge as a special form of human capital reinforces individuals' comparative advantage to perform certain jobs. Thus, knowledge of French or Catalan may be a potentially powerful determinant of workers' allocation to tasks and, more generally, of occupational selection.

In this article we quantify this effect by measuring the contribution of language skills to several patterns of occupational selection. We use Catalonia's labor market, in which the coexistence of Spanish and Catalan provides a good opportunity to analyze this issue in great detail. We find that knowing Catalan significantly reinforces occupational selection of individuals born outside Catalonia into being employed, being an entrepreneur or an independent worker, and permanent employment. Language proficiency also increases the probability of these individuals selecting into white collar work, and employment in trade, services, and government and education. In particular, being able to read and speak Catalan increases selection into white collar occupations by between 9 and 14 percentage points, while writing Catalan increases by 6 to 13 percentage points the probability of engaging in services, and government and educational activities.

Individuals self-select into activities in which they perform relatively better and, therefore, earn more. This is the basis for Roy's (1951) model: tasks in different occupations require different types of skills. This mechanism underlies selection of individuals into being workers or managers (Lucas 1978), sectorial choice and wage distribution (Heckman and Sedlacek 1985, Gould 2002), or patterns of schooling, employment and occupational choice (Keane and Wolpin 1997).

Because different occupations require different communication skills, language proficiency is an important determinant of occupational selection. The literature on this important connection is, however, scarce. Chiswick and Miller (2007) show that knowledge of English is important in matching immigrants to occupation in the US. Additionally, some authors show that sharing a common language is important both in domestic and international trade: according to Hutchinson (2002, 2005) linguistic distance to English is a crucial determinant of trade between countries; Tadesse and White (2008) find that greater cultural differences between the US and a trading partner reduce exports to that country; Sauter (2009) finds that trade in industries in which direct oral communication with the importer is necessary increases with the probability that people in another Canadian province speak the same language.

In Catalonia, as in other economies characterized by linguistic diversity, an active language policy has increased the economic value of Catalan knowledge. Individuals with more knowledge of Catalan are significantly more likely to be employed (Rendon 2007). Our work goes a step further and analyzes the incidence of language knowledge on the probability that an individual performs certain kinds of jobs over others. That is, the economic value of Catalan differs across activities and occupations, so that some are relatively more attractive to those individuals who have the required level of language knowledge. We use two samples from Census years 1991 and 1996 to estimate the effect of Catalan knowledge on occupational selection.¹ First, we estimate a Probit model for the individuals' level of language proficiency in order to correct for the possible endogeneity of Catalan knowledge, as it may be jointly determined with occupational selection. Then, we estimate a bivariate Probit model for the probability of choosing a given occupation *conditional* on a given Catalan proficiency level. We find that for individuals born outside Catalonia reading and speaking as well as writing Catalan significantly reinforce occupational selection into communication-intensive jobs and positions. Specifically, language proficiency

¹These are the last census with reliable data on language for Catalonia. The census of 2001 had several inconsistencies and therefore has not been used in the analysis of language issues.

increases the probability that individuals select into becoming entrepreneurs or independent workers, as well as into permanent employment, white collar work, jobs in trade and services, and government and educational activities.

The remainder of this paper is organized as follows. The next section presents the background of multilingualism, language policy, and economic structure in Catalonia. Section 3 describes the data set and discusses the main descriptive statistics. Section 4 explains the estimation procedure and results, and Section 5 details the main conclusions of this article.

2 Background

In this Section we explain the language transition in formal communication from Spanish to Catalan, and describe the main features of the economic structure of Catalonia.

2.1 A Multilingual context

Two important language related phenomena occurred on the second half of the twentieth century. On the one hand, the fall of dictatorships in Europe and the process of decolonization in Africa and Asia. This implied that, in many countries, widely spoken languages that were only informally used gained official status and were given priority in education over the former official, usually colonial, languages. On the other hand, subsequent waves of immigration to Northern Europe and North America gave rise to the issue of language assimilation of immigrants to their host countries.

Research by most economists has traditionally focused on this second issue, approaching language as a form of human capital valued by the market, crucial to convergence between wages of immigrants and natives. A recent stream of the economic literature attempts to measure the economic effects of changes in the language of education. For example the switch from French to Arabic in Morocco; from English

to Welsh in Wales; from Russian to Estonian in Estonia; or from English to Spanish in Puerto Rico (see Angrist and Lavy 1977, Grin and Vaillancourt 1998, Sabourin and Bernier 2003, Angrist et al. 2008). However, evidence on the economic effects of this kind of language switching is so far inconclusive. While Angrist and Lavy (1977) find that the language switch in Morocco decreased returns to education, Angrist et al. (2008) show that once education-specific cohort trends were introduced, English instruction had no effect on English-speaking ability among Puerto Rican natives.

As other European countries, Spain is characterized by a vast language diversity. Altogether, around forty percent of the population of Spain live in areas with two official languages. While Castilian (Spanish) is official in the totality of the territory, Catalan, Galician, and Basque share co-officiality with Castilian in their own territories.² A comparison of the importance of these languages in their territories reveals that Galicia, with little immigration, has the highest proportion of speakers of their own language. However, Catalonia exhibits the best evolution of indicators of knowledge, use, and favorable attitudes towards the language. Not only has language knowledge increased through the past twenty years, but also the proportion of individuals who consider Catalan their main language. Language commitment to one's own language though is higher in the Basque Land and Navarra: out of those who speak Basque, the proportion that also write it is higher than its equivalent for Catalan in Catalonia and Galician in Galicia (Siguán 1999)..

During most of its history Catalan has been official in Catalonia. However, from the forties to the seventies, during Franco's regime, Spanish was declared the only official language in Catalonia (as in the rest of Spain), and Catalan was relegated to informal use. This, combined with massive immigration of Spanish speakers to Cat-

²Catalan is official in Catalonia (6,995,206 inhabitants in 2005), in Valencian Land (4,692,449 inhabitants), and in Balearic Islands (983,131). Galician is official in Galicia (2,762,198) and Basque is official in the Basque Land (2,124,846) and in the north of Navarra. These regions represent 39.81% of the total Spanish population (44,108,530). Other languages are Asturian or Bable (with around 600,000 speakers and not official in its territory: Asturias and the north of Castilla-Leon), and Aranés or Occitan, official in Valh d'Aran, within Catalonia (9,100 inhabitants).

alonia in the sixties and seventies, contributes to explain how an important proportion of native Catalans of a certain age do not master Catalan. During the eighties and nineties the autonomous Catalan government carried out the ‘Normalization’ policy with the explicit goal for Catalan to match and replace Castilian (Spanish) as the official language, particularly, in the fields of education, public administration and public media. These public policies significantly contributed to the recovery of the Catalan language in Catalonia, so that in Catalonia there is an economic premium for knowing Catalan language (Rendon 2007). Before showing how the economic incentive for knowing Catalan varies by occupation and activities, in the next subsection we will describe briefly the economic structure of Catalonia.

2.2 Economic Structure of Catalonia

Within the European Union Catalonia has approximately the same size and population as Denmark, a higher per capita GDP than Portugal or Greece, and inflows of foreign direct investment above Austria’s or Estonia’s. A densely populated and heavily urbanized area in the Northeast of Spain, Catalonia has more than 7 million inhabitants, one quarter of which are concentrated in its capital, Barcelona. It represents 16% of the Spanish population in a territory that is only 6% of the country’s total. Catalonia’s GDP is one fifth of Spain’s, and its per capita income level has historically been above the average for Spain. Currently, it is 23% above the average for the whole European Union (EU-27) in PPP terms, although still below the average for the EU-15.

Within Spain’s economy Catalonia has traditionally had a preponderant role. Currently it occupies the first place in terms of overall economic activity, about 20% of the Spanish total, industry, around 21%, and trade, 18%; occupying the third place in tourism, 15%, Spain’s main industry (Servei d’Estudis 2008). Its economic structure is strongly dominated by the manufacturing sector, which represents about 30% of its GDP, above Spain and EU-15 averages. Its industrial composition is balanced and

diversified: no single sector's macromagnitude (production, employment, gross value added, etc.) represents more than 20% of the total.

In recent years traditional sectors such as textiles have lost ground, both in terms of gross value added and employment, to more technology intensive industries such as chemicals, transportation equipment, and optical & precision instruments. Namely, more than 60% of all Spanish pharmaceutical production and over 50% of all of Spain's laboratories are located in Catalonia, which has attracted major biotech companies. Specifically, Catalonia attracts more than half of all foreign direct investment in Spain related to R&D activities (INE 2004). Overall, Catalonia's investment in R&D, at 24.6% of Spain's total, has only very recently fallen second to Madrid's 28.4%, still far ahead of the Basque Country's 9.5%, ranked third. Also food & beverage processing, and paper, publishing & printing have expanded.

Within the services sector, tourism, health, financial services, and firm-oriented services are pivotal. Trade has undergone a two-decade transformation toward larger retailers and distribution chains. Nevertheless, small and medium enterprises constitute 99% of Catalonia's business sector. Specifically, firms with 100 employees or less still absorb about 66% of workers. The overall unemployment rate in Catalonia, although above the EU-15 average, has consistently kept below the Spanish average for the past two decades: in 1992 it was 13.6%, reached its 21.2% peak in 1994, decreased to 18.9% in 1996, and stabilized around 8.8% in year 2000.

[Table 1 here]

Table 1 illustrates the sectorial evolution of the employment composition in Catalonia and Spain. Catalonia is clearly more industrial than Spain: in Catalonia industry represents 30% of employment in 1991 and 26% in 1996 while in Spain these percentages are 21% and 18%, respectively. Nevertheless, the service sector is clearly the largest employer in both Catalonia (from 56% of employment to 62%) and Spain (from 58% to 64%).

In the next sections we show how in this multilingual growing service economy, proficiency in reading, speaking and writing Catalan influences agents' choices for certain sectors of production.

3 Data

We use two samples of 250,000 randomly selected individuals extracted from Census data for 1991 and 1996,³ provided by the Catalan and Spanish National Statistical Institutes (IDESCAT-INE). These datasets contain information on personal attributes such as gender, age, marital status, schooling, place of residence, place of birth, number of years in Catalonia, occupational status, and knowledge of Catalan. We combine the Census data with data on municipalities in order to capture the externality effects on sample individuals of residing in areas with high employment rates or widespread Catalan knowledge. We restrict the sample to parents and children from 16 to 60 years of age, born in Spain but not in Catalonia, and participating in the labor force.⁴ We include individuals living in several types of households: singles, divorced or separated individuals living alone, and individuals in multi-personal households, in which we only consider both parents and their children. The final sample contains 47,053 individuals for year 1991, and 69,043 individuals for 1996. Appendix A.1 details the sample selection.

Descriptive statistics for all variables by Census year and gender are presented in Table 2.⁵ Because we cannot relate individuals across Census years or link two individuals of the same family the data only allow us to analyze the cross-sectional effect of individual attributes on language selection.⁶

³Unlike the census of 1991, applied in all of Spain, the census of 1996 was only applied in Catalonia.

⁴International immigration rates to Catalonia for the two census years are low. E.g. 4% in 1991. Moreover, in several estimations we found origin to be non-significant in determining employment status. However, origin may be more important for explaining wages, for which unfortunately there is no data in the census.

⁵Appendix A.2. explains the definition of these variables in further detail.

⁶In order to study intergenerational language persistence or to restrict the sample to children of

[Table 2 here]

Occupational variables in Table 2 are ordered by employment status, type of work, occupation, and activity. Men exhibit higher employment rates, are more likely to be entrepreneurs and self-employed, and more likely to be permanent workers than women. Women, on the other hand, are more likely to be found in white collar occupations, mostly in activities within services, and government and education. There are no important gender differences among individuals working in trade.

Respondents' level of Catalan knowledge is self-declared which, given the linguistic proximity of Catalan and Castilian, may lead respondents to over-report proficiency.⁷ In order to solve this problem we class individuals who claim to understand and either only read or only speak Catalan as having a basic level of Catalan knowledge; individuals who report to read *and* speak are in the intermediate level; while those who can also write belong in the advanced level.

We can see a gender gap has opened in 1996. Relative knowledge of Catalan was similar for both genders in 1991, with the bulk of sample individuals declaring a basic level of knowledge. In 1996 the proportion of men in the basic level has decreased by eleven percentage points, whereas for women the decrease is twenty-three percentage points. A similar proportion of men and women, 27%, claim to have intermediate knowledge of Catalan, but the proportion of women with and advanced knowledge of the language, at 23.2%, has more than doubled since 1991, surpassing men's 13.9%.

Individuals in the sample are on average about 43 years of age, mostly married, with a growing level of education and Catalan knowledge, especially women.⁸ The proportion of people affected by the Normalization process is growing over time for

non-Catalan speakers, for whom learning Catalan means more of an investment, we need information on the parents' language proficiency.

⁷Self-assessed data can lead to biased inference (Charette and Meng 1994). Corrections for missclassification, such as the one used by Dustmann and van Soest (2001), require panel data and are not possible in our current framework.

⁸The lowest level of schooling is illiteracy, followed by 'no schooling.' The percentage of illiterates in the sample is negligible, while the percentage with no schooling is very small. As for Catalan knowledge, the improvement persists even after controlling for variation in the samples, as will be seen below.

both genders, but more so for women, who were less proficient in Catalan than men in 1991 but overtake them in 1996. Most people reside in Barcelona, but a growing percentage of individuals resides in Lleida, Girona and Tarragona, the other three Catalan provinces. On average, individuals in the sample arrived to Catalonia at the end of the sixties, 25 years prior to the censuses, when they were no older than 10, and mainly from Andalusia. There is a slight increase in the number of individuals born in other Catalan-speaking areas in Spain outside Catalonia (Valencia, Balearics and La Franja). Most sample individuals live in municipalities where people are (increasingly) proficient in Catalan, were born in Catalonia, and work in services.

4 Language Effects on Occupation Selection

In this Section we use the Census samples to estimate the effect of language knowledge on occupational selection. First we correct for the possible endogeneity of language knowledge, as it may be jointly determined with occupational selection. To this purpose we estimate a Probit model for the probabilities of an individual exhibiting either the intermediate or the advanced level of proficiency in Catalan for 1991 and 1996, for males and females separately. Because individuals in the sample are not born in Catalonia, they are more likely than natives to have faced a language knowledge decision. In a second stage, we estimate the effect of Catalan knowledge on the probability of choosing a given occupation. Language knowledge is shown to decisively influence individuals' self-selection into occupations, activities and types of work that are intensive in communication.

4.1 Estimation and Identification of Language Effects

We consider two types of occupational selection: functional⁹ and sectorial. Individuals are selected into functions or roles in producing income, i.e. into entrepreneurship

⁹We borrow this term from what is usually called the “functional distribution of income” between individuals who own different factors of production.

(and self-employment) or salaried work; into employment or unemployment; and into permanent or temporary employment. Individuals are also selected into sectors of production, that is, into different occupations and activities. Our results show that proficiency in Catalan influences both types of occupational selection.

A Probit estimation of language effects in occupational selection yields unbiased estimates only if language knowledge is exogenous. As this may not be the case, we also estimate a bivariate Probit for both types of occupational selection which does account for language selection. Thus, the correction for selection is done in two stages, as in Willis and Rosen (1979).

In the first stage, we estimate a Probit model for level of language knowledge. We use the following sources of exogenous variation for language proficiency: level of Catalan knowledge in the individual's municipality; whether the individual was affected by the Normalization policy; whether the individual arrived to Catalonia before age 10; number of years since migration; an interaction term between years since migration and whether the individual arrived before age 10; and origin variables such as whether the individual was born in Andalusia or in other Catalan-speaking regions of Spain.¹⁰ These variables represent the externality effect the community of residence has on the individuals' level of Catalan knowledge, and the individual's exposure to schooling in Catalan and to the Catalan economy. Therefore, they should *directly* affect the individual's proficiency in Catalan, but not the probability of occupational selection.

In the second stage, we use the estimated parameters of the language selection equation to estimate the probability of selection into being employed; being an entrepreneur or self-employed; into permanent employment; white collar occupations; and trade, services, and government and educational activities, *conditional* on a given Catalan proficiency level. Additional control variables in this equation are: age, schooling, marital status, province of residence, percentage of individuals employed

¹⁰The estimation results of the first stage are shown in Appendix A.3.

in the municipality of residence and, of these, percentage employed in the service sector.

4.2 Estimation Results

The language effect on functional selection for the simple Probit and the bivariate Probit models can be seen in Table 3. We present estimates for effects on employment status, on being an entrepreneur, an entrepreneur or self-employed, and on having a permanent job. Clearly, all estimated discrete effects, regardless of gender and census year, are significantly different from zero. We also report the correlation coefficient and a chi-square statistic that tests for independence of occupational selection and Catalan language selection.

[Table 3 here]

We see that knowing Catalan reinforces selection into being employed, being an entrepreneur, being an entrepreneur or independent worker, and being a permanent worker. The language effect on being employed is significant for all groups, yielding systematically lower values once language selection is accounted for,¹¹ except for women in 1991. The correlation coefficient is positive and low, except for women in 1991, at about 0.155. The contribution is higher for women, for whom it decreases over time and increases with the level of language proficiency, whereas for men the evolution is precisely the opposite. These results are similar to those obtained by Rendon (2007).

Selection into being an entrepreneur is significantly affected by knowledge of Catalan, more so for men than for women. The language contribution is higher when language choice is accounted for, with lower correlation between equations for women. This contribution decreases the higher the level of Catalan knowledge. The effect of

¹¹Between a simple probit model and the baseline bivariate probit model there is an average difference of up to 0.7 percentage points, that is, 30% in relative terms.

language on being an entrepreneur *or* a self-employed worker is higher than on being an entrepreneur, while it is still higher for men, with the wrong sign for women. It is also decreasing over Catalan skills and higher once corrected for language selection.

Finally, knowledge of Catalan increases significantly the probability that a worker has a permanent job. This contribution is generally lower once Catalan selection is accounted for. It is higher for the intermediate level than for the advance level of Catalan, and higher for men in 1991, but lower in 1996.

In sum, a simple Probit estimation over-estimates language effects on the probability that an individual is employed and permanently employed, but under-estimates an individual's selection into entrepreneurial and independent work. Language effects on functional occupational selection are generally stronger for men than for women, and decreasing over time and language proficiency level.

Table 4 exhibits the effects of Catalan knowledge on sectorial selection by gender, year, and level of language knowledge. The contribution of language to the probability of an individual selecting a white collar occupation is high, around 10 percent on average, for both genders and census years within either level of language proficiency. This contribution is always higher for women, increasing over skills for men and decreasing for women. It is also decreasing over time, except for men that can write Catalan. This contribution is lower for Biprobit than for Probit estimations for all groups, while correlation between language and occupational choice is strong, between 0.15 and 0.24.

[Table 4 here]

We can also see a differentiated gender pattern in the contribution of language to individuals selecting an activity. Regardless of language proficiency level and census year, the contribution of language to men selecting trade is higher than for women. Correction for language selection reinforces this difference for the intermediate level of Catalan, increasing the language effect for men and decreasing it for women. However, correlation between sectorial occupational choice and language is low. Overall,

the contribution of speaking Catalan to trade is positive and significant, whereas writing has a negative or zero contribution; it is strongly negative for women and writing. Thus, reading and speaking Catalan reinforces selection into trade, but writing Catalan reinforces selection away from trade.

Selection into services by language knowledge is positive and increasing in language skills for both genders: for reading and speaking language effects are between 2% and 4%, while for writing Catalan they are between 7% and 13%. They are higher for women than for men, and increasing over time for men, but decreasing for women. Correction for endogeneity of language reveals higher language coefficients than a simple Probit estimation. Even though correlation between sectorial occupational choice and language is low, independence between language and occupational selection it statistically rejected.

Knowing Catalan undisputedly increases the probability of selection into both services and government and education activities, more so for writing, around 8 percentage points, than for reading and speaking, 2 percentage points, and notoriously more for women than for men. The correction for language selection yields higher values of the language contribution for both skill levels. Correlation between sectorial selection and language skill level is relatively higher for government and education, especially for women.

In sum, a simple Probit model over-estimates the effects of intermediate language proficiency for selection into white collar occupations, as well as the effects of advanced knowledge for selection into trade, but under-estimates selection of individuals with advanced knowledge of Catalan into white collar occupations, and services, government and education activities. Language effects in sectorial occupational selection are generally stronger for women than for men, and increasing over time and language proficiency level.

Summarizing, in all of these estimations language proficiency does make a significant difference in occupational selection, increasing the probability that individuals

engage in jobs in which communicative skills are needed, such as entrepreneur, self-employed worker, permanent employment, white collar occupations, and in activities such as trade, services, and government and education.

4.3 Quality of Instruments

The instruments we use in the benchmark estimation are municipality variables capturing the externality effect of the community of residence on Catalan knowledge; origin variables; assimilation variables such as years passed since migration, and age of arrival; and whether the individual was affected by Normalization. These instruments are related to knowledge of Catalan, but not to occupational selection. In order to assess the quality of these instruments, we follow [?and](#) proceed to exclude subsets of them and compare the resulting estimated language effects on occupational selection.¹²

Table 5 presents the language contribution to functional occupational selection for all subsamples by instruments. That is, for the benchmark Bivariate Probit estimation, which includes all instruments; and for the estimations excluding, first, municipal variables (percentage of people born in Catalonia and percentage of people who write Catalan in the municipality), then origin variables (born in Andalusia, Valencia-Balearics, Franja), and finally assimilation variables (arrived before 10, years since migration, arrived before 10×years since migration). We can see the premium is sensitive to exclusion restrictions, which shows that it is the instruments that affect the contribution language skills have on occupational selection, not the shape of the nonlinear functional form.

¹²Unlike linear models, limited dependent variable models do not have estimated sample errors nor established tests of instruments. Identification of a recursive bivariate probit model requires exogeneity of the sources of variation (Maddala 1983). Specifically, there might be unaccounted individual characteristics that determine Catalan knowledge that are also correlated with occupational choice, which may produce biased results. To reinforce identification of this effect in the occupation equation we have to exclude the variables used in the language equation, that is, the instruments.

[Table 5 here]

Clearly, assimilation variables have the largest effect on the estimated coefficient, followed by origin and municipal variables. Overall, omission of these variables brings coefficients closer to those in the simple Probit model. For employment status, excluding municipality variables has a larger effect than omitting origin variables. For selection into entrepreneurial activity, leaving out municipality variables is more important in 1991, while in 1996 excluding origin has a larger effect, for both skill levels. These results also apply for selection into entrepreneurial and independent work. As for selection into permanent employment excluding municipality variables is more important for women than excluding origin variables in both years, whereas the opposite holds for men.

In Table 6 we report the results for sectorial occupational selection. Again, assimilation variables are crucial in identifying the premium effect. For selection into white collar occupations, municipality matters more than origin for all groups, except for women in 1991. For trade, however the effect of origin is stronger than the effect of municipality variables, whereas for services it is municipality variables, and for government and education activities is municipality, except for men in 1991.

[Table 6 here]

Summarizing, identification of unbiased language effects on occupational selection relies crucially on assimilation variables, such as years since migration, whether the individual arrived before being 10 years old, and an interaction term between these variables. Municipality variables generally rank second in importance for most selections of occupations, with the notorious exceptions of permanent employment for men and trade activities, for which origin variables matter most.

5 Conclusions

When making occupational choices, workers consider their relative skill levels, i.e. their comparative advantages in different occupations. In this article, we have shown that language skills play a significant role in selecting individuals into communication-intensive jobs and positions.

For a sample of Spanish individuals who were not born in Catalonia, we show that language knowledge significantly reinforces occupational selection into being employed, working as an entrepreneur or independent (self-employed) worker, and having permanent employment. Language knowledge also contributes to selection into white collar occupations and trade, services, and government and education activities. These effects are relatively large: reading and speaking Catalan increases selection into white collar occupations by around 10 percentage points for all subsamples, while writing Catalan increases by around 8 percentage points the probability of engaging in services, and government and educational activities. Interestingly, reading and speaking Catalan reinforces selection into trade, but writing Catalan reinforces selection *away* from trade.

In entrepreneurial jobs, self-employment, and permanent employment the language contribution is stronger for men than for women, and decreasing over time and over language levels. The opposite holds for white collar occupations and trade, services, and government and educational activities, in which language effects are stronger for women than for men and increasing over time and over language levels.

We also find that it is necessary to correct for language selection in order to avoid biased measurement of language effects in occupational selection; there is significant correlation between occupational selection and language selection. To this purpose, we used several types of instruments, namely assimilation, origin, and municipality variables. Of these, assimilation variables are shown to be most important for identification of language effects on occupational selection.

In sum, our results indicate that language knowledge is an important component

in the set of skills that constitute an individual's comparative advantage. As such, it contributes to selection into relatively more communication-intensive jobs, activities, and occupations. Entrepreneurial and independent work require more communication than salaried work, for instance in relations with customers, providers, and associates. In activities such as trade, services, and government and education, individuals need to relate to others, not only through reading and speaking, but also by producing written documents (reports, instructions, etc.), which require higher levels of proficiency in Catalan. Accordingly, individuals with higher language skills engage in these occupations.

This analysis is indicative of how occupational selection occurs in a multilingual economy. Monolingual agents who migrate to a multilingual region in the same country may have the option not to learn the local language. However, they do learn it, because, among other possible reasons, this knowledge improves their job matching, especially in communication-intensive activities and occupations.

Appendix

A.1. Sample Selection

The following table illustrates the importance of the selection criteria in constructing the sample.

	1991	1996
Total sample	250 000	250 000
Only main household members: parents and children	-17 654	-17 903
Only individuals between 16 and 60 years old	-82 297	-81 770
Only Spaniards	-5 740	-4 745
Only if year of arrival in Catalonia available		-3 788
Only individuals in the labor force	-47 421	-44 809
Only if Catalan language variable available	-25	
Only if born outside Catalonia	-49 810	-27 942
Selected sample	47 053	69 043

A.2. Definition of the Variables

The construction of each variable is presented below.

Employment.- The census reports a variable called “relationship with the activity.” An individual is employed if he or she reports to be working; unemployed if he or she reports to be not working and looking for his/her first job or having worked before. Individuals who report other options (military service, retired, student, working at home) are excluded from the sample.

Type of work.- This variable is literally called “professional situation;” it includes entrepreneurs who hire workers, which we call “Entrepreneur;” self-employed professionals or independent workers who do not hire other workers, which we call “Self-Employed;” and Temporary Workers. There are other categories, such as cooperatives and salaried workers.

Occupation.- This variable is literally called “occupation, profession or craft” and is specific to the interviewed individual. It is coded according to the Catalan Code of Occupations (CCO-94). We group occupations into “White collar,” mostly professionals in trade and services; and the rest, mostly “Blue collar,” involving more manual labor.

Activity.- This variable provides information on the industrial sector of the firm where the worker performs his or her work. It is coded according to the Catalan Classification of Economic Activities (CCAIE-93). We group them into ‘Trade,’ ‘Services,’ and ‘Government and Education.’¹³

Schooling.- The census reports the maximum level of studies attained by the individual. To each level, we assign the number of years of schooling.

¹³Our exact grouping for the classification of occupation and activities is available from the authors upon request.

Age.- It is the census year, 1991 or 1996, minus the year of birth.

Normalization.- If the individual was younger than 12 years old in 1984 this dummy variable takes the value of one, it is zero otherwise.

Married.- This variable takes the value of one if the respondent reports to be currently married; it is zero if the respondent reports to be a widow(er), separated, or divorced.

Residence variables.- The census reports the municipality and province of residence for each individual. With this information we construct dummies for Lleida, Girona and Tarragona.

Origin variables.-The census reports the municipality and the province of birth for each individual. With this information we construct dummies for individuals who are not born in Catalonia, Andalusia, Valencia, Balearics or La Franja.

YSM (Years since Migration).- The census reports the year of arrival to Catalonia. YSM is the census year minus this number. We also construct a dummy indicating if somebody arrived when s/he was no older than 9.

Municipal variables.- We use the residence variable to assign each individual the corresponding information of the municipality.

A.3. Catalan Knowledge

Table A1 reports the coefficients of the selection equation into reading and speaking Catalan on the left side and into writing Catalan on the right side. The covariates used in the estimations are the variables described before, including squared and interaction terms. There are some common features in the estimated coefficients for the four subsamples. Individuals with more years of schooling are more likely to know Catalan. Older individuals are more likely to read and speak Catalan, but less likely to write it, possibly a result of being schooled during the years in which the use of Catalan was exclusively informal. The probability of knowing Catalan is thus increasing both in schooling and age, but at a decreasing rate. Older individuals exhibit lower returns to schooling, especially in the probability of writing Catalan.

[Table A1 here]

Individuals who are not married or live outside Barcelona are more likely to know Catalan, especially to read and speak it. In areas with a higher percentage of employment in services, arguably the more urban areas, it is less likely that an individual knows Catalan. On the contrary, in areas with a higher percentage of individuals who know Catalan, it is more likely that an individual knows Catalan, which is indicative of the importance of social interactions at the local level. In areas where the percentages of people being employed or being born in Catalonia are higher an individual is more likely to read and speak Catalan, but less likely to write it. As with age, this difference across Catalan skills may stem from growing up in the times when the use of Catalan was purely informal.

Individuals affected by Normalization are more likely to read and speak Catalan. This effect is weaker in 1996 than in 1991, greater for men than for women, and greater for reading and speaking than for writing skills.

Individuals who arrived to Catalonia when they were younger than 10 are more likely to know Catalan. More exposure to the local culture, captured by the variable *years since migration*, makes language assimilation more likely. However, the effect of local exposure on knowing Catalan is stronger for individuals who arrived at a mature age. Individuals who were born in Andalusia are less likely to know Catalan, especially women. On the contrary, individuals born outside Catalonia in areas where Catalan is used, such as Valencia, Balearics and La Franja, are more likely to know Catalan.

We also report indicators of overall fit, pseudo R^2 statistics. They are around 0.30, a relatively high value for a discrete variable, which reveals the fairly good explanatory power of the covariates in this estimation.

References

- Angrist, J., Chin, A. and Godoy, R. (2008), 'Is Spanish-only schooling responsible for the Puerto Rican language gap?', *Journal of Development Economics* **85**(1-2), 105–128.
- Angrist, J. D. and Lavy, V. (1997), 'The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco', *Journal of Labor Economics* **15**, 48–76.
- Charette, M. and Meng, R. (1994), 'Explaining language proficiency. Objective versus self-assessed measures of literacy', *Economic Letters* **44**, 313–321.
- Chiswick, B. R. and Miller, P. W. (2002), 'Immigrant Earnings: Language Skills, Linguistic Concentrations and the Business Cycle.', *Journal of Population Economics* **15**, 31–57.
- Chiswick, B. R. and Miller, P. W. (2007), Matching Language Proficiency to Occupation: The Effect on Immigrants' Earnings. The University of Western Australia, Department of Economics. Economics Discussion Working Papers. 07-07.
- Dustmann, C. and van Soest, A. (2001), 'Language Fluency and Earnings: Estimation with Misclassified Language Indicators', *Review of Economics and Statistics* **83**, 663–674.
- Gould, E. D. (2002), 'Rising Wage Inequality, Comparative Advantage, and the Growing Importance of General Skills in the United States', *Journal of Labor Economics* **20**(1), 105–147.
- Grin, F. and Vaillancourt, F. (1998), Language Revitalisation Policy: An Analytical Survey. Treasury Working Paper 98/6. Mimeo.
- Heckman, J. J. and Sedlacek, G. (1985), 'Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market', *Journal of Political Economy* **93**, 1077–1125.
- Hutchinson, W. K. (2002), 'Does Ease of Communication Increase Trade? Commonality of Language and Bilateral Trade', *Scottish Journal of Political Economy* **49**, 544–556.
- Hutchinson, W. K. (2005), "'linguistic Distance" as a Determinant of Bilateral Trade', **72**, 1–15.
- INE (2004), Encuesta de Innovación Tecnológica en las Empresas. Instituto Nacional de Estadística.
URL: <http://www.ine.es/inebase/cgi/axi>
- Keane, M. and K. Wolpin (1997), 'The Career Decisions of Young Men', *Journal of Political Economy* **10**(3), 473–522.
- Lucas Jr., R. E. (1978), 'On the Size Distribution of Business Firms', *Bell Journal of Economics* **9**(2), 508–523.
- Maddala, G. S. (1983), *Limited-dependent and qualitative variables in econometrics*, Cambridge University Press, Cambridge.
- Rendon, S. (2007), 'The Catalan Premium: Work and Language in Catalonia', *Journal of Population Economics* **20**(3), 669–686.

- Roy, A. D. (1951), 'Some Thoughts on the Distribution of Earnings', *Oxford Economic Papers* **3**, 135–146.
- Sabourin, C. and Bernier, J. (2001), Government responses to language issues. international examples. Office of the Languages Commissioner of Nunavut. Mimeo.
- Sauter, N. (2009), Talking Trade: Language Barriers in Intra-Canadian Commerce. FIW Working Paper series, 023.
- Servei d'Estudis (2008), *Anuario Econmico de Espaa 2008*, Caja de Ahorros y Pensiones de Barcelona, "La Caixa", Barcelona.
- Siguán, M. (1999), Conocimiento y uso de las lenguas en España. (investigación sobre el conocimiento y uso de las lenguas cooficiales en las Comunidades Autónomas Bilingües. Centro de Investigaciones Sociológicas (CIS), 22, Madrid.
- Tadesse, B. and White, R. (2008), Cultural Distance as a Determinant of Bilateral Trade Flows: Do Immigrants Counter the Effect of Cultural Distance? Forthcoming in *Applied Economic Letter*.
- Willis, R. and Rosen, S. (1979), 'Education and Self-Selection', *Journal of Political Economy* **87**(5), S7–S36.

Table 1: Sectorial Employment

Region Censur Year	Catalonia		Spain	
	1991	1996	1991	1996
Total Employed (in %)				
Agriculture, fishing and forestry	3.48	3.54	9.86	8.15
Mining, quarrying and extracting	0.82	0.75	1.06	1.00
Manufacturing	30.43	25.75	21.03	18.02
Construction	9.45	8.22	9.73	8.97
Services	55.81	61.75	58.32	63.86
TOTAL (in thousands)	2,247	2,463	13,203	13,931

Source: INE, Regional Accounts

Table 2: Summary Statistics

Census Year Gender	1991		1996	
	Men	Women	Men	Women
Occupational Variables				
Employed	91.0	77.0	86.3	74.2
Type of work: % Entrepreneurs	6.0	3.6	7.3	4.8
Type of work: % Self-Employed	11.7	11.4	13.0	10.4
Type of work: % Temporary Workers	22.9	34.7	24.8	32.1
Occupation: % in White Collar	30.1	49.8	38.0	60.7
Activity: % in Trade	27.0	27.0	28.0	28.2
Activity: % in Services	15.0	43.6	17.0	44.4
Activity: % in Government and Education	5.7	11.2	8.3	14.6
Catalan Knowledge				
Basic: Reading or Speaking	69.8	72.7	59.3	49.8
Intermediate: Reading and Speaking	19.9	17.5	26.8	27.0
Advanced: Writing	10.3	9.8	13.9	23.2
Socio-Demographic Variables				
Age	42.4	42.5	44.3	41.5
Years of Education	6.4	5.7	7.3	7.7
% Married	82.3	85.6	81.2	71.8
% Normalized	0.8	0.6	2.2	3.6
% Residence in Lleida	3.1	3.2	3.3	3.2
% Residence in Girona	6.5	5.9	6.9	7.2
% Residence in Tarragona	7.8	7.3	8.7	8.4
Years Since Migration	25.0	24.8	28.4	26.8
% arrived age < 10	28.6	35.4	32.0	38.0
% Born in Andalusia	64.0	62.4	63.0	57.9
% Born in Valencia-Balearics	3.4	3.8	3.5	4.3
% Born in Franja	1.2	1.4	1.2	1.5
Municipality Variables:				
% Advanced Catalan Knowledge	37.4	37.5	43.9	44.5
% Born in Catalonia	63.9	64.0	65.4	65.8
% Work in Services	51.8	52.1	58.1	59.1

Table 3: Language Effects on Functional Occupational Selection
Standard errors in small fonts

Language Skills Census Year Gender	Intermediate: Reading and Speaking				Advanced: Writing			
	1991		1996		1991		1996	
	Men	Wom.	Men	Wom.	Men	Wom.	Men	Wom.
Employed								
Probit	3.06	8.17	5.08	6.80	2.03	7.94	2.45	6.05
	1.58	1.34	2.20	1.39	1.12	1.27	1.15	1.33
Biprobit	1.88	9.67	2.15	1.70	1.15	12.03	1.79	3.54
	0.91	1.79	0.82	0.37	0.57	2.74	0.79	1.13
ρ	0.068	0.155	0.058	0.031	0.034	0.154	0.040	0.057
	0.018	0.013	0.017	0.019	0.024	0.017	0.022	0.023
χ^2	15.06	150.12	12.20	2.64	2.11	83.43	3.36	6.29
Entrepreneur								
Probit	1.63	0.60	1.50	0.48	0.76	0.47	0.95	-0.19
	1.03	0.36	0.84	0.33	0.48	0.28	0.52	0.13
Biprobit	2.48	0.47	2.18	0.74	0.72	0.38	1.25	-0.35
	0.90	0.21	0.70	0.26	0.28	0.16	0.44	0.14
ρ	0.121	0.077	0.102	0.054	0.029	0.050	0.047	-0.023
	0.019	0.031	0.020	0.034	0.026	0.040	0.026	0.039
χ^2	40.65	5.95	24.78	2.55	1.30	1.43	3.25	0.32
Entrepreneur and Self-Employed								
Probit	2.93	2.09	4.23	0.87	-0.02	1.16	1.27	-0.73
	1.23	0.93	1.55	0.43	0.01	0.51	0.45	0.35
Biprobit	5.13	2.05	3.36	-0.34	1.45	1.25	2.46	-2.80
	1.33	0.59	0.79	0.11	0.44	0.38	0.79	1.20
ρ	0.114	0.098	0.075	-0.010	0.027	0.048	0.044	-0.073
	0.014	0.019	0.015	0.024	0.019	0.026	0.020	0.028
χ^2	64.96	24.96	23.60	0.17	1.92	3.38	4.81	6.83
Permanent Worker								
Probit	5.92	4.59	6.55	5.49	4.11	2.83	3.24	4.70
	2.16	1.05	1.63	1.07	1.65	0.70	0.90	0.97
Biprobit	1.98	1.78	3.13	0.17	0.48	-1.04	0.35	1.20
	0.72	0.46	0.89	0.04	0.19	0.33	0.12	0.41
ρ	0.042	0.030	0.062	0.003	0.008	-0.015	0.006	0.018
	0.016	0.020	0.018	0.023	0.022	0.024	0.023	0.027
χ^2	2.11	11.51	9.89	0.23	1.79	16.31	2.71	0.04

Table 4: Language Effects on Sectorial Occupational Selection
Standard errors in small fonts

Language Skills Census Year Gender	Intermediate: Reading and Speaking				Advanced: Writing			
	1991		1996		1991		1996	
	Men	Wom.	Men	Wom.	Men	Wom.	Men	Wom.
Occupation: White Collar								
Probit	10.92	14.74	10.63	14.72	8.25	10.34	7.54	8.95
	3.76	5.43	3.53	6.07	2.80	4.20	2.55	4.13
Biprobit	10.08	14.03	8.93	13.35	9.08	18.26	11.25	12.66
	2.96	3.76	2.50	4.59	2.81	5.80	3.65	5.63
ρ	0.200	0.228	0.169	0.237	0.145	0.236	0.169	0.195
	0.014	0.020	0.015	0.022	0.019	0.025	0.020	0.027
χ^2	203.59	89.14	122.73	114.23	59.41	56.51	72.15	51.04
Activity: Trade								
Probit	2.80	4.42	2.35	4.64	1.88	-1.01	0.10	-1.58
	0.44	0.70	0.33	0.78	0.29	0.15	0.01	0.24
Biprobit	3.75	1.53	3.88	0.23	0.10	-2.65	-0.35	-6.39
	0.70	0.32	0.61	0.04	0.02	0.68	0.08	1.59
ρ	0.065	0.025	0.069	0.004	0.001	-0.036	-0.005	-0.104
	0.014	0.021	0.014	0.020	0.018	0.025	0.019	0.024
χ^2	22.37	34.14	22.64	0.07	0.22	25.00	0.06	18.60
Activity: Services								
Probit	0.50	-2.34	2.69	-3.12	2.77	4.58	5.47	5.09
	0.18	0.28	0.97	0.36	0.97	0.55	1.89	0.59
Biprobit	2.12	3.98	2.80	1.93	7.11	13.44	9.02	11.55
	0.71	0.80	0.95	0.33	1.94	3.29	2.24	3.30
ρ	0.058	0.057	0.074	0.031	0.147	0.155	0.180	0.158
	0.016	0.019	0.017	0.020	0.020	0.023	0.020	0.023
χ^2	12.79	27.15	18.74	2.32	51.71	16.50	77.47	47.98
Activity: Government and Education								
Probit	1.00	2.97	1.88	3.55	2.63	5.43	3.54	7.58
	0.72	2.81	1.25	2.23	1.76	4.75	2.20	4.66
Biprobit	1.73	2.84	1.83	4.54	5.84	8.87	7.31	12.04
	1.17	2.64	1.16	3.09	2.88	5.05	3.00	5.43
ρ	0.102	0.221	0.086	0.167	0.235	0.403	0.243	0.333
	0.020	0.022	0.020	0.026	0.023	0.022	0.023	0.025
χ^2	25.06	92.48	17.68	40.42	96.07	268.77	108.07	159.99

Table 5: Language Effects on Functional Occupational Selection
 .by Exclusion of Instruments Standard errors in small fonts

Language Skills Census Year Gender	Interm: Reading and Speaking				Advanced: Writing			
	1991		1996		1991		1996	
	Men	Wom.	Men	Wom.	Men	Wom.	Men	Wom.
Employed								
Benchmark	1.88	9.67	2.15	1.70	1.15	12.03	1.79	3.54
	0.91	1.79	0.82	0.37	0.57	2.74	0.79	1.13
No muni	1.64	9.76	2.18	1.82	0.84	11.97	1.97	3.71
	0.79	1.60	0.82	0.38	0.41	2.39	0.84	1.11
No origin	1.91	9.31	2.30	1.72	1.22	11.81	1.95	3.65
	0.92	1.60	0.87	0.37	0.60	2.47	0.84	1.12
No assimilation	1.52	9.99	1.47	2.12	1.00	12.66	1.37	4.16
	0.73	1.51	0.55	0.43	0.49	2.50	0.58	1.23
Entrepreneur								
Benchmark	2.48	0.47	2.18	0.74	0.72	0.38	1.25	-0.35
	0.90	0.21	0.70	0.26	0.28	0.16	0.44	0.14
No muni	2.54	0.48	2.10	0.65	0.81	0.46	1.14	-0.38
	0.89	0.21	0.66	0.22	0.30	0.19	0.37	0.15
No origin	2.50	0.49	2.28	0.83	0.74	0.38	1.35	-0.32
	0.88	0.21	0.71	0.28	0.28	0.15	0.45	0.12
No assimilation	2.79	0.54	2.45	0.91	1.08	0.44	1.60	-0.18
	0.95	0.23	0.74	0.30	0.40	0.18	0.51	0.07
Entrepreneur and Self-Employed								
Benchmark	5.13	2.05	3.36	-0.34	1.45	1.25	2.46	-2.80
	1.33	0.59	0.79	0.11	0.44	0.38	0.79	1.20
No muni	5.38	2.13	3.44	-0.17	1.82	1.44	2.52	-2.41
	1.32	0.57	0.77	0.06	0.52	0.41	0.74	1.00
No origin	5.24	2.01	3.55	-0.05	1.62	1.25	2.65	-2.60
	1.30	0.55	0.80	0.02	0.47	0.36	0.80	1.08
No assimilation	5.97	2.29	4.20	0.11	2.37	1.60	3.21	-2.45
	1.40	0.59	0.91	0.03	0.68	0.45	0.94	1.01
Permanent Worker								
Benchmark	1.98	1.78	3.13	0.17	0.48	-1.04	0.35	1.20
	0.72	0.46	0.89	0.04	0.19	0.33	0.12	0.41
No muni	1.97	2.95	3.77	1.17	0.52	-0.25	0.77	2.08
	0.70	0.74	1.04	0.27	0.20	0.07	0.25	0.64
No origin	2.06	1.80	3.25	0.29	0.60	-1.08	0.43	1.30
	0.75	0.47	0.91	0.07	0.23	0.32	0.15	0.41
No assimilation	2.34	3.26	2.60	1.65	0.75	0.36	-0.06	2.61
	0.83	0.79	0.71	0.38	0.28	0.10	0.02	0.79

Table 6: Language Effects on Sectorial Occupational Selection
by Exclusion of Instruments. Standard errors in small fonts

Language Skills	Interm: Reading and Speaking				Advanced: Writing			
	1991		1996		1991		1996	
	Men	Wom.	Men	Wom.	Men	Wom.	Men	Wom.
Occupation: White Collar								
Benchmark	10.08	14.03	8.93	13.35	9.08	18.26	11.25	12.66
	2.96	3.76	2.50	4.59	2.81	5.80	3.65	5.63
No muni	9.75	13.06	8.75	12.64	8.58	17.25	10.81	12.22
	2.80	3.32	2.41	4.30	2.53	5.15	3.26	5.21
No origin	10.18	13.94	9.22	13.51	9.14	18.30	11.48	12.93
	2.90	3.57	2.53	4.61	2.70	5.56	3.50	5.58
No assimilation	9.60	14.37	7.96	13.60	8.99	18.57	10.25	13.02
	2.67	3.50	2.15	4.58	2.60	5.51	3.01	5.52
Activity: Trade								
Benchmark	3.75	1.53	3.88	0.23	0.10	-2.65	-0.35	-6.39
	0.70	0.32	0.61	0.04	0.02	0.68	0.08	1.59
No muni	3.33	1.10	3.18	-0.53	-0.24	-2.40	-0.94	-6.73
	0.59	0.22	0.47	0.10	0.05	0.58	0.20	1.57
No origin	4.02	1.76	3.99	0.31	0.34	-2.42	-0.28	-6.31
	0.71	0.35	0.59	0.06	0.08	0.59	0.06	1.48
No assimilation	4.08	1.72	3.63	0.27	0.65	-2.00	-0.15	-5.97
	0.66	0.31	0.49	0.05	0.14	0.48	0.03	1.36
Activity: Services								
Benchmark	2.12	3.98	2.80	1.93	7.11	13.44	9.02	11.55
	0.71	0.80	0.95	0.33	1.94	3.29	2.24	3.30
No muni	2.08	3.37	2.82	1.61	6.89	12.43	8.76	10.97
	0.70	0.62	0.96	0.27	1.79	2.76	1.88	2.85
No origin	2.05	3.73	2.82	1.94	7.02	13.26	9.02	11.53
	0.68	0.69	0.94	0.32	1.79	3.03	1.93	3.06
No assimilation	1.25	2.48	1.65	0.33	6.18	11.92	7.31	9.66
	0.42	0.41	0.55	0.05	1.55	2.63	1.50	2.46
Activity: Government and Education								
Benchmark	1.73	2.84	1.83	4.54	5.84	8.87	7.31	12.04
	1.17	2.64	1.16	3.09	2.88	5.05	3.00	5.43
No muni	1.83	2.77	1.83	4.50	5.86	8.48	7.14	11.82
	1.24	2.60	1.16	3.07	2.82	4.71	2.75	4.96
No origin	1.70	2.79	1.82	4.55	5.79	8.76	7.24	12.03
	1.14	2.59	1.15	3.06	2.75	4.73	2.76	5.00
No assimilation	1.07	2.38	0.86	3.69	4.98	7.97	5.54	10.82
	0.74	2.27	0.55	2.48	2.43	4.42	2.21	4.50

Table A1: Language Equation (First Stage Estimation)
Standard errors in small fonts

Language Skills Census Year Gender	Intermediate: Reading and Speaking				Advanced: Writing			
	1991		1996		1991		1996	
	Men	Wom	Men	Wom	Men	Wom	Men	Wom
Constant	-3.10	-4.19	-3.54	-3.82	-1.20	-2.68	-0.94	-2.20
	0.32	0.34	0.40	0.53	0.39	0.42	0.48	0.58
Schooling	27.01	35.07	32.94	37.29	32.85	36.98	34.64	38.05
$\times 10^{-2}$	1.75	2.16	2.50	3.48	2.15	2.55	3.05	3.96
Schooling ²	-67.92	-86.33	-74.23	-91.27	-63.67	-72.15	-50.44	-73.00
$\times 10^{-4}$	5.22	6.37	8.07	11.23	6.45	7.62	10.88	14.18
Age	-0.17	0.28	-0.29	-0.21	-0.88	-0.35	-1.15	-1.10
$\times 10^{-1}$	0.10	0.10	0.13	0.17	0.12	0.13	0.15	0.19
Age ²	0.00	-0.04	0.01	0.00	0.11	0.04	0.12	0.10
$\times 10^{-3}$	0.01	0.01	0.01	0.02	0.01	0.01	0.02	0.02
Age \times Schooling	-6.39	-14.28	-12.95	-10.78	-18.70	-18.62	-22.06	-11.73
$\times 10^{-4}$	3.12	3.75	3.88	5.37	3.67	4.33	4.32	5.62
Married	-0.54	-1.75	0.24	-0.68	-2.31	-2.39	-1.58	-1.99
$\times 10^{-1}$	0.28	0.28	0.29	0.33	0.35	0.34	0.35	0.36
Lleida resident	2.33	1.64	0.48	0.62	2.01	1.08	0.91	0.06
$\times 10^{-1}$	0.57	0.57	0.62	0.88	0.72	0.72	0.74	0.92
Girona resident	2.30	1.71	1.88	1.38	1.53	1.26	2.43	0.54
$\times 10^{-1}$	0.40	0.41	0.45	0.62	0.50	0.53	0.54	0.69
Tarragona resident	1.32	0.66	2.20	1.49	0.91	0.16	1.78	1.24
$\times 10^{-1}$	0.36	0.38	0.37	0.53	0.46	0.49	0.44	0.58
% Mun ^a Employed	-0.21	-1.15	-0.85	0.15	-1.60	-1.24	-2.12	0.68
	0.59	0.62	0.54	0.72	0.76	0.79	0.68	0.82
% Mun ^a Services	-0.82	-0.65	-0.21	-0.25	-1.10	-1.10	-0.77	-0.62
	0.09	0.09	0.10	0.14	0.12	0.12	0.13	0.16
% Mun ^a Born in Catalonia	2.40	2.08	2.21	1.71	3.03	2.71	3.07	1.91
	0.20	0.21	0.26	0.36	0.25	0.27	0.32	0.40
% Mun ^a Write Catalan	0.97	1.55	1.50	1.68	-0.32	0.04	-0.28	0.80
	0.21	0.22	0.25	0.35	0.27	0.29	0.31	0.39
Normalized ^b	0.56	0.49	0.35	0.27	0.52	0.59	0.23	0.19
	0.11	0.12	0.09	0.11	0.12	0.12	0.09	0.11
Arrived younger than 10	-4.27	-4.79	-4.09	-3.92	-1.78	-1.83	-0.92	-1.26
	0.22	0.23	0.28	0.38	0.18	0.20	0.20	0.27
YSM ^c	0.84	0.17	0.97	0.52	0.84	0.17	0.97	0.52
	0.71	0.13	0.71	0.98	0.71	0.13	0.71	0.98
Arrived younger than 10 \times YSM ^c	0.21	0.29	0.78	0.63	0.21	0.29	0.78	0.63
	0.73	0.81	0.77	0.36	0.73	0.81	0.77	0.36
Born in Andalusia	0.56	0.63	0.64	0.60	0.56	0.63	0.64	0.60
	0.57	0.36	0.02	0.89	0.57	0.36	0.02	0.89
Born in Valencia -Balearics	0.26	0.90	0.38	0.57	0.26	0.90	0.38	0.57
	0.90	0.42	0.95	0.40	0.90	0.42	0.95	0.40
Born in Franja	0.95	0.49	0.01	0.43	0.95	0.49	0.01	0.43
	0.20	0.17	0.27	0.30	0.20	0.17	0.27	0.30
Pseudo R ²	0.18	0.22	0.19	0.22	0.21	0.25	0.24	0.28

^a Percentage in *municipi*; ^b Affected by Normalization policy; ^c YSM=Years Since Migration.