Revista de Economía Aplicada E A *Número 67 (vol. XXIII), 2015, págs. 63 a 91*

THE IMPACT OF LANGUAGE PROFICIENCY ON IMMIGRANTS' EARNINGS*

SANTIAGO BUDRÍA
CEEAplA and IZA

PABLO SWEDBERG
Department of Business, St. Louis University

This article uses micro-data from the Spanish National Immigrant Survey to analyze the impact of Spanish language proficiency on immigrants' earnings. The results, based on Instrumental Variables (IV), point to a substantial return to Spanish proficiency, of approximately 20%. This figure varies largely between high- and low-educated immigrants. The high-educated earn a premium of almost 50%, while the premium for the low-educated fails to be statistically significant under some specifications. This conspicuous complementarity between education and language skills poses a challenge for traditional language training policies that need to address the immigrants' heterogeneous educational background.

Key words: immigration, Spanish language proficiency, earnings, instrumental variables.

JEL Classification: F22, J24, J61.

Greater fluency in the destination language is expected to enhance earnings among immigrants. Host language proficiency enhances productivity in the job by making the worker more efficient in performing particular tasks or by reducing the cost of communication within the firm. Moreover, proficient workers are in a better position to obtain information about job opportunities and earnings, and to transmit information about their skills to employers. There is a significant literature examining how English language proficiency affects earnings positively in the US, Canada, the UK and Australia [Carliner (1981), McManus *et al*. (1983), Grenier (1984), Tainer (1988), Rivera-Batiz (1992), Chiswick (1991), Dustmann & Fabbri (2003), Lui (2007), Chiswick & Miller (1999), (2010), Zhen, (2013)]. Additionally, Dustmann & van Soest (2001, 2002) for Germany and several studies in Israel [Chiswick (1998), Chiswick & Repetto (2001), Berman *et al*. (2003)] likewise show the positive impact that host language proficiency has on immigrant earn-

(*) Financial support from the Spanish Ministry of Education through grant ECO2008-04321 and the Junta de Andalucía through grant P07-SEJ-03261 is gratefully acknowledged. The authors thank Antonio Di Paolo and Hipólito Simón for generous comments and suggestions.

ings. Conversely, Hayfron (2001) finds no evidence of positive returns to host language acquisition in Norway for immigrants from developing countries.

This paper explores the impact of Spanish language proficiency on immigrant earnings in Spain. The motivation of the paper is threefold. Firstly, the great majority of the research has been carried out in countries with a long tradition and history of immigration whereas, in Spain, the great bulk of immigrants arrived after 1998. A particularity of this fact is that it makes the immigrant group more homogenous in terms of time spent in the host country and assimilation. Secondly, due to differences in labor market regulations and characteristics, the bulk of the international evidence is strongly case dependent. Most research conducted in Spain has focused on Catalonia, and its regional language, Catalan [(Rendón (2007), Di Paolo (2011), Di Paolo & Raymond (2012)]. This focus is partly due to the recent availability of the Survey of Living Conditions and Habits of the Catalan Population, carried out by the Statistical Institute of Catalonia. Pioneering, these efforts have not been followed by studies focusing on Castilian Spanish, the language spoken throughout Spain. Furthermore, Isophording (2013) explores the impact of foreign language skills –English, German and French– of immigrants in the Spanish labor market. The present paper is intended to fill this gap. Thirdly, the extent to which language skills affect an individual's labor market performance influences the immigrants' income and poverty levels. This will ultimately affect their social and cultural integration to the host country society. It has been shown that immigrants experience a negative wage gap with respect to natives. Moreover, the degree of earnings assimilation is found to differ across studies [Hu (2000), Friedberg (2002), Lubotsky (2007), Adsera & Chiswick (2007), Beenstock *et al*. (2010)]. For Spain, Amuedo-Dorantes & De la Rica (2007) find that there is a 15% wage gap reduction in immigrants' initial 5 years of residence in Spain. Lacuesta *et al*. (2009) also find that the initial wage differential with respect to natives with the same observable characteristics decreases with time spent in Spain. According to their estimates, assimilation of legal immigrants occurs rapidly, with a 50% reduction of the initial wage gap in the first 5-6 years after their arrival. Conversely, Fernandez & Ortega (2008) do not find any progress in the immigrants' labor market outcomes. Rodríguez-Planas (2012) shows that immigrants with a high school degree do not have an advantage in terms of wage assimilation over those who do not. All in all, the native-immigrant wage gap partly reflects the fact that migrants cannot fully use their human capital attributes. Simón *et al*. (2008) argue that the disparities in the wage distributions for the native-born and immigrants are largely explained by their different observed characteristics, with a particularly important influence of occupational segregation.

While the focus of the present paper is not on earnings assimilation, it addresses the impact on wages of one of the most important human capital components in the country of residence: language ability. Immigrants who possess language proficiency are less likely to be overqualified at their job [Blázquez & Rendón, (2012)], most likely because they are more efficient in seeking out and obtaining employment where their skills are most highly rewarded. As it fulfills a number of functions, language plays a crucial role in the integration process. The case of Spain is particularly interesting, due to the massive migration flows experienced over the last decade. The number of legal foreign workers present in Spain increased by 2,259,330

people during the period 2001-2008, a rise that amounts to 53% of the total labor force increase for that same period. As a result, immigrants currently represent a significant segment of the country's work force.

This paper uses the Spanish National Immigrant Survey (NIS), a large-scale immigration survey released recently by the Spanish National Statistics Institute. The empirical strategy is based on a set of earnings equations extended to include a control for the immigrant's level of Spanish language proficiency. The paper adopts an Instrumental Variable (IV) approach. Several instruments are proposed and results under alternative specifications and combinations of instruments are presented. The estimates are reasonably robust across models and suggest that the labor market value of Spanish proficiency is sizable.

An additional question addressed by the paper is whether the relationship between Spanish language proficiency and earnings differs between immigrants with different levels of educational attainment. There are reasons to believe that language proficiency and schooling are complementary inputs of the earnings-generating process. Firstly, language skills are more likely to represent a valuable asset in occupations that require higher levels of formal education. This intuition is supported by the fact that returns to language knowledge tend to be higher in high-skill occupations [Berman *et al*. (2003)]. Secondly, schooling has a large impact on life opportunities, social mobility and labor market outcomes. Given that poor language skills are likely to hamper such opportunities, stronger effects from language proficiency among the highly educated may be expected. Thirdly, we cannot preclude the possibility that host language proficiency acts as a signal to employers about the quality of the individual's post-compulsory education.

Such heterogeneous effects may have pronounced implications for the design of effective integration policies. A common policy priority in OECD countries is labor market integration and the strengthening of educational aspects, including language training [OECD (2012)]. In line with this view, the Spanish Strategic Plan for Citizenship and Integration acknowledged the fact that immigration poses specific challenges that must be tackled, including "the promotion of improvements in immigrants' knowledge of the official languages and social norms in Spain, prerequisites for a cohesive society and for the social integration of immigrants" [Ministry of Labor and Social Affairs (2007)]. Unfortunately, the scope attributed to such policies may be more modest than presumed if workers with low qualifications fail to reap relevant returns from language training. The evidence collected so far is scarce and suggestive of diverging degrees of complementarity between schooling and language skills [Chiswick & Miller (2003), Casale & Posel (2011)]. This paper sheds further light on this issue by assessing the interaction between Spanish proficiency and educational attainment in the Spanish labor market

The paper is organized as follows. Section 2 describes the dataset, the estimating sample and the Spanish language proficiency question. Section 3 describes the IV approach and discusses the choice of excluded instruments. Section 4 examines the determinants of Spanish proficiency obtained from the first stage equation, and presents estimates of the impact of Spanish proficiency on earnings. The results are paired with a variety of validity and relevancy tests to assess the quality of the instruments and the robustness of the IV estimates. Results broken up by education level are also presented. Section 5 contains the concluding remarks.

1. Data and definition of variables

The data is taken from the Spanish National Immigrant Survey (Encuesta Nacional de Inmigrantes), a large-scale immigration survey carried out by the Spanish National Statistics Institute. The data collection was conducted between November 2006 and February 2007 and was based on the Municipal Census (Padrón Municipal). The original survey sample comprises approximately 15,500 individuals. The NIS provides detailed information on the socio-demographic characteristics of immigrants and their previous and current employment status. Immigrants are defined as any individual born abroad (regardless of their nationality) who, at the time of being interviewed, had reached at least 16 years of age and had resided in a home for at least a year or longer, or had the intention to remain in Spain for at least a year[1].

The estimating sample consists of private sector men who are between 18 and 65 years old and work regularly between 15 and 70 hours a week. Self-employed individuals, as well as those whose main activity status is paid apprenticeship or training, and unpaid family workers have been excluded from the sample. Women are disregarded on account of the extra complications derived from potential selectivity bias. Dropping observations, including item non-response, leaves us with a final sample of 3,089 individuals.

1.1. Spanish proficiency

The Spanish proficiency question in the NIS is:

– Thinking of what you need for communicating at work, at the bank or with the public authorities/administration. How well do you speak Spanish?

with answers ranging from 1 ('very well') to 4 ('need to improve'). These responses were used to define SP, a dummy variable that takes value one if the immigrant has Spanish proficiency (1-very well), zero otherwise[2]. According to this criterion, nearly 67% of the sample report being proficient in Spanish.

The question included in the NIS provides a unique opportunity to investigate meaningful relationships contained in the data and fits well the conventions of the literature. The use of subjective evaluations is standard in the field, partly due to the high costs of test-based assessments of language ability. Admittedly, respondents may have different perceptions under identical circumstances of how well they speak a foreign language. Notwithstanding, subjective questions are typically found to be highly correlated with scores from tests designed to accurately measure language skills as well as functional measures of language skills' [Bleakley & Chin (2010)].

(1) Target population of the NIS.

(2) The distribution of responses was: 1. 'very well', 66.8%; 2. 'well', 20.2%; 3. 'sufficient', 7.3%; 'need to improve', 5.7%. Ideally, one would like to differentiate each increment using dummy variables for each possible skill level and, correspondingly, would need an instrument for each dummy variable. We are, however, constrained by the limited set of available instruments. Therefore, the paper follows a stringent criterion by considering only individuals who claim to be able to speak Spanish 'very well'. Results under the alternative classification 1-2 against 3-4 displayed slightly lower returns and are available upon request.

It must be noted that respondents to the NIS are asked (when they report Spanish as a foreign language) to self-assess (yes/no) whether they possess a satisfactory skill level in different areas, including comprehension, speaking, reading and writing[3]. However, this information is provided on a yes/no basis and as many as 99.8% (comprehension), 100% (speaking), 92.0% (reading) and 84.8% (writing) of the sample answer 'yes' to the corresponding question. These figures are far higher than the 66.8% of proficient immigrants that emerge from the central question used in the paper. Therefore, relying on these indicators provides a far less stringent criterion for language proficiency. This is the reason why the present paper does not attempt to differentiate between different types of skills[4].

Table 1 provides summary statistics by proficiency level. Nearly 67% of the sample reports being proficient in Spanish while the remaining 33% has limited language ability. There are some relevant disparities between the two sub-samples. Proficient immigrants earn 22.7% higher hourly wages (6.80 €/hour) than non-proficient immigrants (5.54 €/hour)[5]. They also exhibit higher levels of educational attainment (11.3 against 8.2 years of schooling), a longer period of stay in Spain (13.8 against 8.2 years since migration), slightly shorter professional experiences (20.0 against 21.7 years) and a higher probability of enjoying a permanent contract (55.4% against 38.3%). In terms of marital status and household size, the two groups are quite similar. Proficient immigrants are mainly from North and Latin-America (55.9%) or Central and Western Europe (24.8%) and are more likely to work in the Technological & Scientific sectors (19.5%) and in the Manufacturing & Construction sectors (16.1%). By contrast, non-proficient immigrants are mainly from Northern Africa (33.1%) and Eastern Europe (36.9%) and work mainly in Agriculture & Fishery (36.4%) and other Non-qualified occupations (39.9%).

(3) The precise wording of the questions is: "*Comprehension? 1. Yes. 2. No; Can you speak this language? 1. Yes. 2. No; Can you read in this language? 1. Yes. 2. No; Can you write in this language? 1. Yes. 2. No"*. Immigrants whose mother tongue is Spanish are not required to provide such information, the underlying assumption being that they are fully proficient in the four areas.

(4) Some studies define language proficiency as the self-reported ability to read and write [Casale & Posel (2011)] or the ability to speak and write [Di Paolo, (2011)]. Chiswick (1991) and Dustmann (1994) suggest that reading and writing proficiency are more important determinants of earnings than speaking fluency. On the contrary, Chiswick & Miller (1999) find a speaking premium of about 8 pp, while the reading premium is 6 pp. Interestingly, the authors find that neither speaking but not reading nor reading but not speaking English are beneficial in terms of wages. In Chiswick & Repetto (2001) speaking only Hebrew or Hebrew as a primary language increases earnings by 13 pp, whereas being able to write a letter in Hebrew brings an additional 12 pp increase. In Dustmann & Fabbri (2003) English language fluency is associated with about 22 pp higher wages among immigrants in the UK, whereas conditional on fluency written literacy fails to attract a significant wage premium. Surprisingly, Hayfron (2001) finds that speaking; reading and writing proficiency have no positive impact on earnings among immigrants from developing countries in Norway.

(5) They are also more likely to work. The employment rate is 71.9% among proficient immigrants and 67.1% among the non-proficient.

Table 1: Summary statistics by Spanish proficiency

	Proficient	Non-proficient
Share	0.6679	0.3321
Hourly wage	6.7974	5.5397
	(3.3343)	(2.5986)
Years of schooling	11.2672	8.1773
	(3.5800)	(4.9949)
Experience	19.9810	21.6660
	(9.6261)	(9.9128)
Permanent contract	0.5535	0.3831
	(0.4972)	(0.4864)
Single	0.3601	0.3363
	(0.4801)	(0.4726)
Divorced	0.0641	0.0799
	(0.2447)	(0.2713)
Married	0.5759	0.5838
	(0.4943)	(0.4931)
Children	0.6403	0.6335
	(0.4800)	(0.4820)
Previous unemployment experience	0.3306	0.3091
	(0.4705)	(0.4622)
Illegal status	0.3107	0.3092
	(0.4629)	(0.4610)
Region of origin		
Maghreb	0.0868	0.3314
	(0.2815)	(0.4709)
Sub-Saharan Africa	0.0194	0.0799
	(0.1379)	(0.2713)
Eastern Europe	0.0751	0.3694
	(0.2637)	0.4828
America	0.5594	0.0331
	(0.4966)	(0.1790)
Asia	0.0107	0.0556
	(0.1027)	(0.2291)
Oceania	0.0011	0.0001
	(0.0311)	(0.0041)
Central and western Europe	0.2476	0.1306
	(0,4490)	(0.3113)

Table 1: Summary statistics by Spanish proficiency (continuation)

	Proficient	Non-proficient
Occupation sector		
Army	0.0019	0.0001
	(0.0440)	(0.0312)
Management	0.0456	0.0195
	(0.2085)	(0.1383)
Technology and Sciences	0.1953	0.0382
	(0.3965)	(0.1913)
Services	0.1386	0.1014
	(0.3456)	(0.3019)
Administration	0.0456	0.0127
	(0.2086)	(0.1119)
Agriculture and Fishery	0.2123	0.3635
	(0.4090)	(0.4812)
Manufacturing, Construction	0.1609	0.0653
	(0.3675)	(0.2471)
Others	0.1997	0.3993
	(0.2999)	(0.4577)

Note to Table 1: a) Source: Spanish National Immigrant Survey; b) Standard deviations are in parenthesis.

Source: Own elaboration.

2. Estimation strategy

The earnings equation is specified as follows,

$$\ln w_i = X_i\beta + \gamma SP_i + \varepsilon_i \quad [1]$$

where w is hourly earnings, X includes educational attainment, potential labor market experience and its square, years since migration, type of contract (temporary or permanent), marital status (single, divorced or widowed, reference: married), number of children at home, previous unemployment spells of 3 months or longer in Spain (yes/no), legal status (documented or undocumented), occupational dummies (according to the one digit level National Classification of Occupations), the immigrant's source region (Maghreb, Sub-Saharan Africa, Eastern Europe, North and Latin-America, Asia or Oceania, reference: Central and Western Europe) and dummies for region of residence in Spain. The choice of these variables is duly motivated by the immigration adjustment literature[6]. Finally, this analysis will focus on SP, Spanish language proficiency.

(6) Other conventional controls such as tenure and work experience in the home and host country were disregarded due to large item non-response.

The interpretation of the results including and excluding information on occupation deserves some consideration. Language proficiency, or any other human capital skill, can be viewed as having both direct and indirect impacts on earnings. Direct effects stem from a better communication with coworkers and customers and enhanced productivity. Indirect effects operate via occupational attainment. Because different occupations require different communication skills, language proficiency may be a determinant of occupational selection. Better language skills may increase immigrants' earnings indirectly by allowing them to enter high-paying occupations or to have access to a broader labor market. Consistent with this view, the occupation channel has been found to be important in explaining the earnings effects of language skills [Wang & Wang (2011)][7]. In practice, the inclusion of occupational dummies in the earnings equation tends to provide conservative estimates of the language proficiency effect [Aldashev *et al*. (2009)]. Still, language proficiency is more strongly associated with intra-occupational increases in earnings than it is with inter-occupational increases in earnings [Chiswick & Miller, (2010)]. To provide some sensitivity analysis, the present paper reports results including and excluding occupational controls.

2.1. The endogeneity of Spanish proficiency

Language knowledge may depend on unobservable individual characteristics that are potentially related to immeasurable earnings determinants. That would be the case if, for example, more productive and capable individuals were more likely to be proficient in Spanish. The estimated coefficient then would not reflect the benefits derived from language skills, but merely a spurious correlation. Addressing this issue involves specifying a first stage equation for Spanish proficiency,

$$SP_i = X_i\delta + Z_i\theta + \nu_i \qquad [2]$$

where Z contains the set of excluded instruments. The use of IV is also intended to reduce the extent of attenuation bias that may stem from errors in the measurement of the individual's self-assessed Spanish language proficiency SP[8].

Instruments must be valid [i.e., uncorrelated with earnings, $E(\varepsilon_i|Z_i) = 0$] and relevant (i.e., they must account for a significant variation in SP). Earlier studies addressing the endogeneity problem leave us with a variety of potential candidates. For example, Chiswick & Miller (1995) explore family composition (veteran status, whether married overseas, and number of children) and residential minority-language concentration as identifying instruments to analyze the role of language skills on earnings in Australia. Almost identical instruments are used in Chiswick (1998) to

(7) These effects have been also documented for Catalan knowledge in Catalonia, Spain [Di Paolo (2011), Quella & Rendón (2012)].

(8) Self-reported measures of speaking fluency typically suffer from misclassification/measurement errors, with the probabilities of over-reporting being higher than the probabilities of under-reporting [Dustmann & van Soest, 2001)]. Dustmann & Fabbri (2003) find that measurement error and endogenous choice bias the language coefficients in opposite directions. However, it is difficult to correct for misclassification using cross-sectional data, as in the present case. There is a need for longitudinal immigration data that is currently not available for this survey in Spain. The use of the IV technique in the present context should be seen as a working compromise to mitigate the effect of measurement errors inherent in self-assessed measures of literacy [Charette & Meng (1994)].

analyze the impact of Hebrew language usage among immigrants in Israel. However, as in his previous article, the validity of the selected variables has not been reported. Utilizing a German panel dataset, Dustmann & van Soest (2002) circumvent the measurement and endogeneity problems by using the leads and lags of language proficiency and parental education as instruments. Bleakley & Chin (2004) improve upon the existing literature by proposing an innovative instrument for language proficiency based on the fact that younger children acquire language skills easier than adolescents and adults do. Rendón (2007) shows that variables capturing the externality effect of the community of residence, origin variables, years since migration and age at arrival are significantly related to Catalan proficiency. However, he warns that some of these instruments, particularly the region of origin, might be also related to labor market performance. Gao & Smith (2011) rely on child information (the number of children living in the host city and having at least one child enrolled in primary school in the host city) as a proxy for the individual's proficiency in Mandarin. This road is also explored by Di Paolo (2011) and Di Paolo & Raymond (2011), who assume that language use with their children (as well as other variables including arrival in the host country before the age of 10) affects the parents' language proficiency but does not directly affect their earnings.

This paper explores three potential instruments for Spanish language proficiency that were available in the NIS. The first instrument is a dummy variable that captures whether the respondent arrived in Spain before the age of 12. This choice is motivated by the well-documented fact that younger children learn languages more easily than adolescents and adults. Cognitive scientists refer to this as the critical period hypothesis according to which there is a critical age range in which individuals learn languages more easily and after which language acquisition is more difficult. Behavioral evidence has been supportive of this hypothesis: late learners tend to attain a lower level of language proficiency. While there is no general consensus on the critical period age, most studies find supportive evidence of the critical period hypothesis, with the range between 5 and 15 [Chiswick & Miller (2008)]. Our choice of the break point, 12, falls into this range[9]. The underlying hypothesis is that the coefficient for this variable should be positive and highly significant in the first stage equation. The second instrument captures whether the respondent has a child who is proficient in the Spanish language. Arguably, parents' exposure to communication with their children in the destination country's official language and access to the children's superior pronunciation skills acts as a transmission mechanism. The third and last instrument accounts for the respondent's willingness to stay in Spain for the next 5 years. These exclusion restrictions are confirmed by the encompassing tests of validity and relevancy reported in the next section.

3. Results

This section reports the estimates from the first stage equation and discusses the determinants of SP. Then, it moves on to assess the causal relationship between Spanish proficiency and earnings.

(9) The results presented very small variations following alternative choices, including 10, 11 and 13.

3.1. The determinants of Spanish proficiency

Table 2 reports the results from Eq. [2]. These shed some light on the determinants of SP and provide complementary evidence on the quality of the instruments. As expected, arrival before the age of 12 is a strong predictor of language ability. Immigrants that arrived at a young age are between 38.7 and 40.1 pp more likely to be proficient in Spanish, whereas the impact of having a child with host language skills is about one fourth of this figure (9.0 pp). Planning to stay in the country for the next five years shows a lower coefficient (around 3 pp) and tends to be statistically significant at lower levels.

Inspection of the table suggests that language skills are stronger among more educated individuals. This observation is consistent with the evidence reported in related research and probably stems from the close relationship between the individual's learning ability and the individual's level of schooling. A complementary explanation is that more educated individuals have strong economic incentives to learn a language due to higher market wages and employment rates [Chiswick & Miller (2003)]. Language skills depend negatively on total labor market experience, though at a decreasing rate. The direction of this effect is probably due to (ceteris paribus) lower language skills among older immigrants and the close relationship between professional experience and age. The positive and statistically significant coefficient for the length of stay in Spain (the years since migration) indicates that a longer exposure to the local language firmly favors its knowledge. The number of children in the household shows a positive although non-significant coefficient. Chiswick & Miller (1995) suggest that children may have counteracting effects on language: first, they may act as a translator between the parent and the host country language speaking community (thus reducing incentives to learn the foreign language). Second, they may enhance exposure to the majority population by forcing the parent to cope with institutional matters, like school and parents of native friends of children. Our results indicate that these effects tend to offset each other.

Having a permanent contract and previous unemployment experience in Spain are both positively associated with language proficiency. These effects match expectations. Immigrants with poor language skills are less likely to be offered a permanent position whereas most immigrants without an unemployment history in the Spanish labor market are either recent immigrants or individuals entering the labor force after periods of inactivity. Spells of inactivity are expected to speed up human capital depreciation and to reduce the extent of social interactions leading to language learning. A similar reasoning, as well as a reduced bargaining power leading to a less favorable labor market position, applies to illegal immigrants. Moreover, there are systematic differences in the degree of fluency by region of origin. Migrants from North and Latin-America are more likely to be language proficient due to the existence of large Spanish-speaking communities (Ecuador, Colombia, Argentina and Bolivia, mainly). Conversely, fluency is weaker ceteris paribus among immigrants from Maghreb, Sub-Saharan Africa, Eastern Europe and Asia. Finally, there are significant differences between occupation sectors, with workers in the Management, Technology & Sciences, and Manufacturing & Construction occupations being more likely to display proper language skills.

Table 2: THE DETERMINANTS OF SPANISH LANGUAGE ABILITY

	(a)	(b)	(c)	(d)
	Coeff.	Coeff.	Coeff.	Coeff.
Arrived before age 12	0.3892***	0.3874***	0.4006***	0.3990***
	(0.0259)	(0.0257)	(0.026)	(0.0258)
Child proficient in Spanish	0.0899***	0.0896***		
	(0.0159)	(0.0158)		
Plans to stay in Spain	0.0320*		0.0351**	
	(0.0172)		(0.0172)	
Years of schooling	0.0151***	0.0151***	0.0158***	0.158***
	(0.0019)	(0.0019)	(0.0019)	(0.0019)
Experience	-0.0096***	-0.0096***	-0.0087***	-0.0087***
	(0.0026)	(0.0026)	(0.0026)	(0.0026)
Experience $(x1000)^2$	0.1003**	0.1009**	0.0878*	0.0883*
	(0.0511)	(0.0512)	(0.0510)	(0.0510)
Years since migration	0.0097***	0.0097***	0.0105***	0.0105***
	(0.0010)	(0.0010)	(0.0010)	(0.0010)
Permanent contract	0.0315**	0.0313**	0.0339**	0.0337**
	(0.0134)	(0.0134)	(0.0135)	(0.0135)
Single	-0.0233	-0.0226	-0.0324*	-0.0318*
	(0.0178)	(0.0178)	(0.0178)	(0.0177)
Divorced	-0.0538**	-0.0540**	-0.0823***	-0.0825***
	(0.0264)	(0.0264)	(0.0258)	(0.0258)
Children	-0.0089	-0.0095	0.0220	0.0215
	(0.0191)	(0.0191)	(0.0181)	(0.0181)
Previous unemployment experience	0.0268**	0.0267**	0.0321**	0.0320**
	(0.0130)	(0.0130)	(0.0131)	(0.0131)
Illegal status	-0.0454***	-0.0457***	-0.0438***	-0.0441***
	(0.0143)	(0.0130)	(0.0144)	(0.0144)
Region of origin				
Maghreb	-0.2417***	-0.2421***	-0.2486***	-0.2489***
	(0.0295)	(0.0295)	(0.0296)	(0.0296)
Sub-Saharan Africa	-0.2538***	-0.2539***	-0.2683***	-0.2684***
	(0.0417)	(0.0417)	(0.0420)	(0.0420)
Eastern Europe	-0.2728***	-0.2729***	-0.2810***	-0.2810***
	(0.0310)	(0.0310)	(0.0310)	(0.0310)
America	0.3425***	0.3425***	0.3418***	0.3418***
	(0.0231)	(0.0231)	(0.0232)	(0.0232)
Asia	-0.3597***	-0.3601***	-0.3760***	-0.3763***
	(0.0524)	(0.0524)	(0.0521)	(0.0521)
Oceania	0.2166***	0.2177***	0.2021***	0.2031***
	(0.0623)	(0.0623)	(0.0400)	(0.0400)

Table 2: THE DETERMINANTS OF SPANISH LANGUAGE ABILITY (continuation)

	(a)	(b)	(c)	(d)
	Coeff.	Coeff.	Coeff.	Coeff.
Occupation sector				
Army	-0.0379	-0.0401	-0.0357	-0.0376
	(0.1048)	(0.1049)	(0.1059)	(0.1060)
Management	0.0846**	0.0839**	0.0931**	0.0925**
	(0.0384)	(0.0383)	(0.0385)	(0.0384)
Technology and Sciences	0.0816***	0.0814***	0.0875***	0.0873***
	(0.0245)	(0.0245)	(0.0245)	(0.0245)
Services	0.0021	0.0021	0.0049	0.0049
	(0.0218)	(0.0218)	(0.0218)	(0.0218)
Administration	0.0382	0.0384	0.0386	0.0387
	(0.0317)	(0.0317)	(0.0315)	(0.0315)
Agriculture and Fishery	0.0036	0.0031	0.0027	0.0022
	(0.0174)	(0.0174)	(0.0175)	(0.0175)
Manufacturing, Construction	0.1099***	0.1093***	0.1230***	0.1225***
	(0.0221)	(0.0221)	(0.0221)	(0.0221)
Constant	0.4812***	0.4724***	0.4612***	0.4539***
	(0.0535)	(0.0520)	(0.0531)	0.0517)
R^2	0.5471	0.5471	0.5424	0.5423
No. of observations	3,089	3,089	3,089	3,089

Note to Table 2: i) Source: Spanish National Immigrant Survey; ii) Heteroskedastic-robust standard deviations are in parenthesis; iii) *** denotes significant at the 1% level, ** denotes significant at the 5% level; * denotes significant at the 10% level; iv) additional controls: 19 dummies for Spanish Autonomous Communities; v) Reference individual: arrived in Spain after age 12, does not have a language proficient child, does not plan to stay in Spain for the next 5 years, is married, has not been unemployed for more than three months in the past, has a non-permanent contract, resides legally in Madrid, comes from Central-Western Europe and has average schooling, experience and years since migration.

Source: Own elaboration.

3.2. The immigrants' earnings equation

The OLS estimates from the earnings equation are reported in the first column of Table 3. According to the results, being proficient in Spanish increases wages by 4.9 pp. However, before discussing how reliable and robust this coefficient is to changes in the estimating strategy, it is noteworthy to unveil the role of the remaining covariates included in the equation.

The results are as follows. An additional year of schooling raises earnings by about 1.2 pp, a figure that is below conventional estimates reported for the total Spanish population [Budría (2005), for a survey]. This result may well reflect labor market discrimination against migrant workers, according to which immigrants end up

Table 3: EARNINGS EQUATIONS: OLS AND IV ESTIMATES

	OLS	(a) Benchmark	(b)	(c)	(d)
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
Spanish language proficiency	0.0491***	0.2044***	0.2056***	0.1671**	0.1684**
	(0.0165)	(0.0659)	(0.0659)	(0.0712)	(0.0712)
Years of schooling	0.0118***	0.0086***	0.0086**	0.0094***	0.0094***
	(0.0017)	(0.0022)	(0.0022)	(0.0023)	(0.0023)
Experience	0.0086***	0.0092***	0.0092***	0.0090***	0.0090***
	(0.0022)	(0.0025)	(0.0025)	(0.0025)	(0.0025)
Experience^2 (x1000)	-0.1173***	-0.1258**	-0.1259**	-0.1238**	-0.1239**
	(0.0449)	(0.0513)	(0.0513)	(0.0508)	(0.0508)
Years since migration	-0.0010	-0.0015	-0.0015	-0.0013	-0.0013
	(0.0008)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Permanent contract	0.0552***	0.0469***	0.0469***	0.0489***	0.0489***
	(0.0121)	(0.0137)	(0.0137)	(0.0138)	(0.0138)
Single	-0.0165	-0.0148	-0.0147	-0.0152	-0.0152
	(0.0153)	(0.0159)	(0.0159)	(0.0158)	(0.0158)
Divorced	-0.0547**	-0.0385	-0.0383	-0.0424	-0.0422
	(0.0239)	(0.0259)	(0.0259)	(0.0258)	(0.0258)
Children	0.0146	0.0191	0.0118	0.0125	0.0125
	(0.0152)	(0.0163)	(0.0163)	(0.0162)	(0.0162)
Previous unemployment experience	-0.0539***	-0.0618***	-0.0619***	-0.0599***	-0.0600***
	(0.0123)	(0.0132)	(0.0132)	(0.0131)	(0.0131)
Illegal status	0.0083	0.0096	0.0096	0.0093	0.0093
	(0.0136)	(0.0137)	(0.0137)	(0.0136)	(0.0136)
Region of origin					
Maghreb	-0.1473***	-0.0943***	-0.0938***	-0.1070***	-0.1065***
	(0.0229)	(0.0325)	(0.0325)	(0.0377)	(0.0337)
Sub-Saharan Africa	-0.1444***	-0.0864**	-0.0859**	-0.1003**	-0.0998**
	(0.0327)	(0.0408)	(0.0408)	(0.0418)	(0.0418)
Eastern Europe	-0.1392***	-0.0724**	-0.0719**	-0.0884**	-0.0879**
	(0.0225)	(0.0361)	(0.0361)	(0.0376)	(0.0376)
America	-0.1790***	-0.2108***	-0.2111***	-0.2032***	-0.2035***
	(0.0180)	(0.0238)	(0.0239)	(0.0243)	(0.0243)
Asia	-0.0922**	-0.0179	-0.0173	-0.0357	-0.0351
	(0.0388)	(0.0647)	(0.0695)	(0.0643)	(0.0643)
Oceania	0.2866	0.2330***	0.2325***	0.2458***	0.2454***
	(0.2188)	(0.0742)	(0.0743)	(0.0738)	(0.0739)

Table 3: EARNINGS EQUATIONS: OLS AND IV ESTIMATES (continuation)

	OLS	(a) Benchmark	(b)	(c)	(d)
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
Occupation sector					
Army	0.2469*	0.2476	0.2476	0.2475	0.2475
	(0.1379)	(0.1661)	(0.1661)	(0.1651)	(0.1651)
Management	0.3921***	0.3755***	0.3753***	0.3795***	0.3793***
	(0.0324)	(0.0542)	(0.0542)	(0.0541)	(0.0541)
Technology and Sciences	0.4905***	0.4806***	0.4805***	0.4829***	0.4829***
	(0.0222)	(0.0264)	(0.0264)	(0.0262)	(0.0262)
Services	-0.0142	-0.0164	-0.0164	-0.0159	-0.0159
	(0.0199)	(0.0196)	(0.0196)	(0.0195)	(0.0195)
Administration	0.1325***	0.1234***	0.1233***	0.1256***	0.1255***
	(0.0331)	(0.0308)	(0.0308)	(0.0307)	(0.0307)
Agriculture and Fishery	0.1541***	0.1546***	0.1546***	0.1545***	0.1545***
	(0.0159)	(0.0151)	(0.0151)	(0.0150)	(0.0150)
Manufacturing, Construction	0.1594***	0.1359***	0.1357***	0.1416***	0.1414***
	(0.0208)	(0.0213)	(0.0213)	(0.0214)	(0.0214)
Constant	10.4700***	10.3925***	10.3919***	10.4111***	10.4105***
	(0.0458)	(0.0590)	(0.0590)	(0.0603)	(0.0604)
R^2	0.3947	0.3773	0.377	0.3846	0.3844
No. of observations	3,089	3,089	3,089	3,089	3,089

Note to Table 3: i) Source: Spanish National Immigrant Survey; ii) Heteroskedastic-robust standard deviations are in parenthesis; iii) *** denotes significant at the 1% level, ** denotes significant at the 5% level; * denotes significant at the 10% level; iv) additional controls: 19 dummies for Spanish Autonomous Communities; v) Reference individual: married, has not been unemployed for more than three months in the past, has a non-permanent contract, resides legally in Madrid, comes from Central-Western Europe and has average schooling, professional experience and years since migration.

Source: Own elaboration.

in low-pay jobs that are not commensurate with their educational background. As expected, professional experience is associated with higher earnings, though at a decreasing rate. The coefficient for years since migration has a negative sign and is statistically insignificant[10]. This observation may be seen as odd because years since migration are expected to reflect assimilation. However, it should not be so if we take into account that individuals who arrived in Spain many years ago are likely to be less skilled than more recent immigrants. Having a permanent contract is associated with 5.5 pp higher wages whereas previous unemployment experience decreases wages by about 5.4 pp. There are conspicuous earnings differentials among immi-

(10) Adding a quadratic also resulted in non-significant coefficients.

grants from different regions of origin. Compared to the reference individual (an immigrant from Central-Western Europe), workers from Maghreb, Sub-Saharan Africa, Eastern Europe, America and Asia reap significantly lower earnings. The predicted wage penalty is between 9.2 pp for Asians and 17.9 pp for Latin-American immigrants. Finally, the results suggest roughly 40-50 pp higher earnings in the Management and Technology & Sciences sector, compared to the reference category 'Unqualified occupations'. Administration, Agriculture & Fishery and Manufacturing & Construction carry a lower though significant premium.

The impact of these covariates does not present much variation when we switch to the IV estimates reported in the remaining columns of the table. Thus, we move on to analyze the relationship between Spanish proficiency and wages.

3.3. The impact of Spanish proficiency on wages

Table 4 summarizes the point estimates of γ, the coefficient associated with the variable SP in Eq. [1]. It also reports a variety of tests for the quality of the instruments. The left-hand side of the table presents results controlling for occupation while results on the right-hand side were obtained after dropping the occupation variables from the first and second stage equations. For reasons of space, the following analysis mainly focuses on the former results.

The IV estimates suggest that assuming exogenous SP yields a downward-biased prediction. Column (a) is based on the full set of instruments and should be considered the benchmark combination of instruments, for it delivers the most favorable quality tests (see below). In this case, Spanish proficiency is associated with a wage increase of 20.4 pp. This figure is quite robust to alternative specifications. In column (b) the instrument with the lowest, though statistically significant, explanatory power upon SP (whether the individual plans to stay in Spain for the next 5 years) has been dropped. This results in an almost negligible variation in the estimated return to language proficiency (20.6 pp). Column (c) proceeds likewise and drops the instrument with the second lowest explanatory power (children's Spanish proficiency) in addition to the previously dropped instrument. This results in a modest decrease of the coefficient (16.7 pp) and an increase of its standard deviation. The loss of precision is, however, modest and the resulting effect is still significant at conventional statistical levels. The just identified model that appears in column (d) delivers an almost identical prediction (16.8 pp).

These robustness checks suggest that the point estimates are quite stable and that the labor market value of Spanish proficiency is sizable. However, as noted earlier, it is convenient to check the sensitivity of the results to the exclusion of the occupation information from the set of controls. This is performed on the right-hand side of Table 4. Here the underlying assumption is that, if Spanish proficiency is a prerequisite to securing a high-paying occupation, this effect should be regarded as a return to language knowledge rather than as return to occupation. As expected, the R^2 of the earnings equation decreases after dropping the occupation dummies by about 13% and the coefficients of individual-level variables, including schooling, type of contract and legal status (omitted from the table) rise by between 8 pp and 25 pp. Perhaps more interesting is the fact that the OLS returns to Spanish proficiency are smaller when occupation controls are omitted, and the observed difference (4.9

Table 4: The impact of Spanish language ability on hourly earnings

		IV Combination of instruments				IV Combination of instruments				
	OLS	(a)	(b)	(c)	(d)	OLS	(a)	(b)	(c)	(d)
Panel 1. Full sample										
Spanish language proficiency	0.0491***	0.2044***	0.2056***	0.1671**	0.1684**	0.0673***	0.1958**	0.1975**	0.0957	0.0967
	(0.0165)	(0.0659)	(0.0659)	(0.0712)	(0.0712)	(0.0180)	(0.0799)	(0.0800)	(0.0986)	(0.0987)
Occupation dummies	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Instruments										
Arrived before age 10	—	Yes	Yes	Yes	Yes	—	Yes	Yes	Yes	Yes
Child proficient in Spanish	—	Yes	Yes	No	No	—	Yes	Yes	No	No
Plans to stay in Spain	—	Yes	No	Yes	No	—	Yes	No	Yes	No
R^2	0.3947	0.3773	0.3771	0.3846	0.3844	0.2641	0.2517	0.2514	0.2634	0.2633
Sargan-Hansen χ^2 test	—	2.89	2.59	0.31	—	—	3.01	2.06	1.11	—
F-test	—	86.48***	129.78***	119.32***	238.91***	—	74.27***	111.51***	87.95***	175.99***
R^2 in first stage equation	—	0.5471	0.5471	0.5424	0.5423	—	0.5266	0.5266	0.5173	0.5173
Partial R^2		0.0849	0.0847	0.0754	0.0753		0.0663	0.0663	0.0479	0.0479
Durbin-Wu-Hausman test	—	6.18***	6.28***	3.04*	3.11*	—	2.75*	2.82*	0.086	0.093
No. of observations	3,089	3,089	3,089	3,089	3,089	3,089	3,089	3,089	3,089	3,089

Note to Table 4: i) Source: Spanish National Immigrant Survey; ii) Heteroskedastic-robust standard deviations are in parenthesis; iii) *** denotes significant at the 1% level, ** denotes significant at the 5% level; * denotes significant at the 10% level; iv) additional controls: years of schooling, potential labor market experience, years since migration, type of contract, marital status, children at home, previous unemployment experience, legal status, immigrant's region of origin and 19 dummies for Spanish Autonomous Communities; v) Reference individual: married, has not been unemployed for more than three months in the past, has a non-permanent contract, resides legally in Madrid, comes from Central-Western Europe and has average schooling, experience and years since migration.

Source: Own elaboration.

Table 4: The impact of Spanish language ability on hourly earnings (continuation)

		IV Combination of instruments				IV Combination of instruments				
	OLS	(a)	(b)	(c)	(d)	OLS	(a)	(b)	(c)	(d)
Panel 2. Restricted sample										
Spanish language proficiency	0.0681***	0.1659**	0.1649**	0.0893	0.0879	0.1102***	0.1691**	0.1687**	-0.0445	-0.0454
	(0.0179)	(0.0762)	(0.0765)	(0.0921)	(0.0924)	(0.0189)	(0.0857)	(0.0858)	(0.1359)	(0.1364)
Occupation dummies	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Instruments										
Arrived before age 10	—	Yes	Yes	Yes	Yes	—	Yes	Yes	Yes	Yes
Child proficient in Spanish	—	Yes	Yes	No	No	—	Yes	Yes	No	No
Plans to stay in Spain	—	Yes	No	Yes	No	—	Yes	No	Yes	No
R^2	0.4757	0.4665	0.4667	0.4645	0.4651	0.3776	0.3741	0.3741	0.3533	0.4645
Sargan-Hansen χ^2 test	—	2.52	2.42	.11	—	—	3.41	3.39	0.02	—
F-test	—	40.46***	60.42***	48.71***	96.84***	—	33.36***	49.94***	27.94***	55.63***
R^2 in first stage equation		0.4217	0.4216	0.4116	0.4115		0.3791	0.3791	0.3586	0.3585
Partial R^2	—	0.0671	0.0648	0.0507	0.0501	—	0.0541	0.0519	0.0228	0.0207
Durbin-Wu-Hausman test	—	3.90**	3.81**	0.05	0.05	—	4.77**	4.72**	1.27	1.28
No. of observations	1,739	1,739	1,739	1,739	1,739	1,739	1,739	1,739	1,739	1,739

Note to Table 4: i) Source: Spanish National Immigrant Survey; ii) Heteroskedastic-robust standard deviations are in parenthesis; iii) *** denotes significant at the 1% level, ** denotes significant at the 5% level; * denotes significant at the 10% level; iv) additional controls: years of schooling, potential labor market experience, years since migration, type of contract, marital status, children at home, previous unemployment experience, legal status, immigrant's region of origin and 19 dummies for Spanish Autonomous Communities; v) Reference individual: married, has not been unemployed for more than three months in the past, has a non-permanent contract, resides legally in Madrid, comes from Central-Western Europe and has average schooling, experience and years since migration.

Source: Own elaboration.

against 6.7 pp) is statistically significant at the 5% level[11]. Indeed, the estimated returns under the combination of instruments (c) and (d) are almost halved and become non-significant once the occupational controls are omitted. This notwithstanding, the IV estimates with and without occupational controls are strikingly similar under the benchmark combination of instruments (column (a)). In this case, the estimate for γ falls only by 0.8 pp, from 20.4 to 19.6. Moreover, the difference between these two figures fails to be statistically significant. We interpret this observation as evidence showing that omitting occupation controls from the estimating equations does not dramatically affect the causal relationship between Spanish proficiency and earnings.

While these estimates seem high, they are towards the lower range of estimates reported by studies using a similar IV approach for language fluency in other countries. Chiswick & Miller (1995) find that returns to English proficiency are between 40 and 57 pp among immigrants in Canada and the United States and Chiswick & Miller (2003) report estimates within 26 and 42 pp in Canada. For Israel, Chiswick (1998) reports a figure above 35 pp for Hebrew fluency among migrants. Gao & Smyth (2011) analyze the return to standard Mandarin among internal migrants in China, finding that the estimated impact amounts to 40 pp. The results reported in Table 4 suggest that the earnings gains from host language proficiency in Spain are large but relatively lower than in other countries. Nevertheless they appear to be a little higher than the returns to regional language proficiency in Spain because, in Di Paolo & Raymond (2012), the return to Catalan knowledge is about 16 pp among immigrants in Catalonia.

An important concern with IV is the validity of the instruments. Validity is not assured if the excluded instruments have direct effects on earnings beyond those flowing indirectly through Spanish proficiency. This problem may yield biased estimates and will be exacerbated by a weak correlation between the endogenous variable and the instruments [Bound *et al*. (1995)]. To assess these issues, Table 4 reports various relevant tests. In all cases, the Sargan-Hansen test of orhtogonality does not lead to the rejection of the null hypothesis. This suggests that the selected instruments are uncorrelated with the earnings variable[12].

As for the weak IV problem, two diagnosis tests are reported: the F-test for the joint significance of the selected instruments and their relative contribution to R^2 in the SP equation. We recall Bound's *et al*. (1995) Equation 7 (pp. 44), according to which the bias of the IV estimator is given by:

$$BIAS_{IV} = \frac{\rho_{z,\varepsilon} / \rho_{x,\varepsilon}}{\rho_{z,x}} \qquad [3]$$

The bias depends positively on the degree of endogeneity, i.e., the correlation ($\rho_{z,}\varepsilon$) between the instruments (z) and the error term of the second stage equation (ε) and negatively on the correlation ($\rho_{z,x}$) between the potentially endogenous

(11) χ^2 statistic = 4.97, p-value = 0.0258. The statistical significance of cross-models differences was tested using the STATA postestimation command *suest*.

(12) Validity cannot be tested in the just identified model of column (d).

variable (x) and the instrument (z). In other words, having relevant instruments is very important to attenuate the potential bias arising from instrument endogeneity. According to the influential work of Stock *et al*. (2002), the value of $\rho_{z,x}$ is acceptable if the resulting F-test associated with the exluded instruments is above 10. In our estimations, the F-statistics are, in all cases, above 70 (and a little higher under some specifications) and, thus, readily in the safe zone suggested by Stock *et al*. (2002). Similarly, the contribution of the excluded instruments to the partial R^2 of the first stage equation is above 0.075 in models with occupation controls and above 0.045 in models without occupation controls. These figures are above the lower range of values considered admissible in the literature[13].

Finally, the necessity of resorting to IV should be statistically assessed. Since the IV estimator always has larger asymptotic variance than the OLS estimator, there is a loss of efficiency in the latter approach. The Durbin-Wu-Hausman statistic reported tests for the exogeneity of the instrumented variable. A rejection of the null hypothesis implies that the OLS estimator does not yield consistent estimates due to the endogeneity of SP. If this were the case, the IV estimator should be used. In all cases reported on the left-hand side of the table (models with occupation controls) the tests reject the null hypothesis, although rejection occurs only at the 10% level for the combination of instruments (c) and (d). Taken together, the results justify the choice of the IV approach. However, the tests are less conclusive in models that abstract from the occupation information. In this case, rejection occurs only at the 10% level for the set of instruments (a) and (b), whereas (c) and (d) seem to suggest that the OLS estimate is consistent.

As a robustness check, Panel 2 reports complementary results where immigrants whose mother tongue is Spanish have been dropped ('Restricted sample'). Despite the inclusion of explicit controls for the immigrants' region of origin in the earnings equation, unobserved heterogeneity related to SP may persist due to immigrants' diverging language, cultural and social backgrounds. Restricting the sample to individuals whose mother tongue is not Spanish results in a little lower, but still significant, estimates. In models with occupation information, we find that combinations of instruments with a higher explanatory power (partial $R^2 > 0.06$), such as the ones displayed in (a) and (b), yield significant estimates. Specifically, the benchmark model predicts a 16.6 pp return to Spanish proficiency. However, the estimates are now less stable due to smaller sample size. In columns (c) and (d), the coefficient fails to be statistically significant, and the Durbin-Wu-Hausman test fails to reject the endogeneity of the language proficiency variable. This seems to suggest that the 6.8 pp return implied by the OLS estimation is an unbiased estimate. However, these two specifications should be less preferred, for they exhibit a partial R^2 statistic (< 0.06) that is suggestive of inferior instrument quality. These conclusions also apply to the set of

(13) In this scenario, the results of just-identified models, like the one reported in column (d), deserve credit. As Angrist and Krueger (2009, p. 209) put it: "It turns out, however, that just-identified 2SLS (say, the simple Wald estimator) is approximately unbiased. This is hard to show formally because just-identified 2SLS has no moments (i.e., the sampling distribution has fat tails). Nevertheless, even with weak instruments, just-identified 2SLS is approximately centered where it should be (we therefore say that just-identified 2SLS is median-unbiased)".

models that abstract from the occupation information. In this alternative scenario, and under the benchmark combination of instruments, the Durbin-Wu-Hausman statistic clearly rejects the exogeneity of SP. This observation is pertinent because the verdict of this test was less apparent in the results obtained from the full sample.

3.4. The complementarity between language proficiency and education

This section explores whether the relationship between Spanish language proficiency and earnings differs between education groups. The suspicion that a lack of academic qualifications may greatly hamper the returns to language skills motivates the analysis.

Table 5 reports the returns to Spanish language proficiency by educational attainment. To avoid problems derived from small cell-size, only two categories are considered. Panel 1 shows the result for immigrant workers that completed at least upper secondary education while Panel 2 is devoted to immigrants with less than upper secondary education. Splitting the sample instead of including a language proficiency-schooling interaction term is intended to allow for different endogeneity and earnings-determination processes within the two groups. Several conclusions emerge from the results. First, the returns to Spanish language proficiency for migrants with an upper secondary or higher education degree are remarkably high. In accordance with to the benchmark specification, highly educated individuals that possess strong Spanish language skills are expected to earn, ceteris paribus, 47.9 pp more than highly educated individuals with a limited knowledge of the host language. This impact is more than double the 20.4 pp obtained from the total sample. Second, the smaller sample size of the highly educated group does not come at the cost of unstable estimates for γ under alternative sets of instruments. A glance at the different columns indicates that the variation in γ is very modest and centred around 47 pp. Third, among the highly educated migrants, the returns to Spanish proficiency rockets to about 65 pp when the occupation information is dropped from the equations. This observation suggests that among this group the occupation channel (i.e., individuals with better language skills accessing high-pay occupations) is relatively more important in explaining the earnings effects of language skills than among less educated immigrants. Fourth, the tests for the excluded instruments suggest inferior instrument quality as we move from columns (a) to (d) both in models with and without the occupation information. Still, the results are broadly satisfactory across specifications.

Panel 2 shows the results for immigrant workers that acquired less than upper secondary education. According to the benchmark combination of instruments, the impact of language proficiency on earnings is 16 pp and 25 pp, depending on whether occupation dummies are included. These figures are more than 30 pp lower than the language premium obtained by highly educated individuals, and the difference between the two subsamples is strongly significant[14]. Furthermore, in some cases (columns (c) and (d)) the language effect is statistically significant only at the 10% level. As a potential concern, we detect a deterioration of instrument quality as

(14) χ^2 statistic = 23.04, p-value = 0.0000 in the model with occupations; χ^2 statistic = 26.78, p-value = 0.0000 in the model without occupations.

we move from the high to the low-educated sample. In the former case, and under some combinations of instruments, the Sargan-Hansen test rejects orthogonality. This outcome suggests that, among the low educated, the excluded instruments have direct effects on earnings beyond those flowing indirectly through Spanish proficiency. This problem may yield biased estimates. However, variation in the returns to Spanish proficiency across specifications that pass and specifications that do not pass the validity tests are found to be relatively modest as, in all cases, these returns are well below the figures obtained from the highly-educated sample.

Table 6 reports the results when immigrants whose mother tongue is Spanish are dropped from the sample ('Restricted sample'). Differences between the two education groups are even larger. Specifically, we find that the returns to Spanish proficiency among the highly-educated rockets to 58.31% (benchmark specification) in the restricted sample, while the low-educated fail to reap a significant premium in models without occupation information. All in all, the estimates are supportive of the substantially lower returns earned by workers with less than upper secondary education.

Earlier studies have examined complementarity effects between education and language proficiency. Chiswick & Miller (2003) find higher returns to education among English proficient immigrants in Canada. In Di Paolo & Raymond (2012), highly educated workers earn a return from Catalan proficiency of about 25 pp while workers with fewer than 9 years of education (presumably, less than secondary education) reap no return. Despite the fact that differences between studies must be interpreted cautiously, the results in this paper suggest that acquiring Spanish language proficiency is a profitable human capital investment for less educated immigrants as well. Finally, Casale & Posel (2011) examine the relationship between English language proficiency and earnings in South Africa. They find a very large wage premium for employed African men who are both proficient in English and have attained either a diploma or academic degree. Specifically, English proficiency together with the completion of some form of post-secondary qualification offers a return of 93 pp, against the 33 pp figure reaped by less educated, language-proficient workers. Despite the fact that sociological, economic and labor market differences between the two countries hamper any thorough comparisons, the results reported in the present study are suggestive of weaker, though substantial, complementarity effects between language and education in the Spanish labor market.

4. Conclusions

This paper examined the impact of Spanish language proficiency upon earnings among immigrants in Spain. In doing so it adds to the less prevalent evidence collected in a non-English speaking country. The more recent and intense nature of immigration in Spain makes it a unique context for analysis; most immigrants arrived in the period 2000-2008, which makes them a more homogenous group and reduces the problems that arise from cohort effects.

The analysis was based on recent data obtained from the Spanish National Immigrant Survey, and stems from an IV approach supported by good quality instruments. The benchmark estimates show that being proficient in Spanish raises immigrant earnings by about 20 pp. This impact is reasonably robust to the combination

Table 5: The impact of Spanish language ability on hourly earnings, by education groups

		IV Combination of instruments					IV Combination of instruments			
	OLS	(a)	(b)	(c)	(d)	OLS	(a)	(b)	(c)	(d)
Panel 1. Workers with upper sec. educ. or more										
Spanish language proficiency	0.0636***	0.4786***	0.4817***	0.4566***	0.4598***	.0916***	0.6475***	0.6502**	0.6603**	0.6624**
	(0.0231)	(0.1447)	(0.1448)	(0.1633)	(0.1635)	(.0248)	(0.1785)	(0.1786)	(0.2291)	(0.2291)
Occupation dummies	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Instruments										
Arrived before age 10	—	Yes	Yes	Yes	Yes	—	Yes	Yes	Yes	Yes
Child proficient in Spanish	—	Yes	Yes	No	No	—	Yes	Yes	No	No
plans to stay in Spain	—	Yes	No	Yes	No	—	Yes	No	Yes	No
R^2	0.4597	0.3582	0.3567	0.3687	0.3672	.3637	0.1795	0.1777	0.1711	0.1695
Sargan-Hansen χ^2 test	—	0.56	0.13	0.43	—	—	0.69	0.01	0.67	—
F-test	—	29.83***	44.81***	40.66***	81.45***	—	25.07***	37.65***	87.95***	62.19***
R^2 in first stage equation	—	0.4509	0.4509	0.4467	0.4466	—	0.4378	0.4378	0.4316	0.4316
Partial R^2		0.0426	0.0419	0.0352	0.0338		0.0315	0.0301	0.0207	0.0201
Durbin-Wu-Hausman test	—	9.62***	9.77***	6.64**	6.75**	—	11.41***	11.51***	6.82***	6.87***
No. of observations	1,760	1,760	1,760	1,760	1,760	1,760	1,760	1,760	1,760	1,760

Note to Table 5: i) Source: Spanish National Immigrant Survey; ii) Heteroskedastic-robust standard deviations are in parenthesis; iii) *** denotes significant at the 1% level, ** denotes significant at the 5% level; * denotes significant at the 10% level; iv) additional controls: years of schooling, potential labor market experience, years since migration, type of contract, marital status, children at home, previous unemployment experience, legal status, immigrant's region of origin and 19 dummies for Spanish Autonomous communities; v) Reference individual: married, has not been unemployed for more than three months in the past, has a non-permanent contract, resides legally in Madrid, comes from Central-Western Europe and has (within sub-sample) average schooling, experience and years since migration.

Source: Own elaboration.

Table 5: The impact of Spanish language ability on hourly earnings, by education groups (continuation)

		IV Combination of instruments				IV Combination of instruments				
	OLS	(a)	(b)	(c)	(d)	OLS	(a)	(b)	(c)	(d)
Panel 2. Workers wih less than upper sec. educ.										
Spanish language proficiency	0.0271	0.1605**	0.1594**	0.1328*	0.1313*	0.0498***	0.2564**	0.2489**	0.1701**	0.1574*
	(0.0239)	(0.0678)	(0.0679)	(0.0690)	(0.0692)	(0.0245)	(0.0789)	(0.0792)	(0.0854)	(0.0861)
Occupation dummies	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Instruments										
Arrived before age 10	—	Yes	Yes	Yes	Yes	—	Yes	Yes	Yes	Yes
Child proficient in Spanish	—	Yes	Yes	No	No	—	Yes	Yes	No	No
plans to stay in Spain	—	Yes	No	Yes	No	—	Yes	No	Yes	No
R^2	0.2533	0.2352	0.2355	0.2419	0.2422	0.1804	0.1357	0.1388	0.1653	0.1683
Sargan-Hansen χ^2 test	—	4.28	4.00**	0.42	—	—	7.72**	5.87**	2.74*	—
F-test	—	50.64***	75.80***	72.80***	145.15***	—	41.89***	61.55***	49.30***	96.34***
R^2 in first stage equation		0.6577	0.6576	0.6554	0.6554		0.6294	0.6292	0.6217	0.6213
Partial R^2	—	0.1554	0.1552	0.1498	0.1496	—	0.1161	0.1115	0.0976	0.0967
Durbin-Wu-Hausman test	—	4.96**	4.87**	3.00*	2.91*	—	8.33***	7.69***	2.32	1.83
No. of observations	1,329	1,329	1,329	1,329	1,329	1,329	1,329	1,329	1,329	1,329

Note to Table 5: i) Source: Spanish National Immigrant Survey; ii) Heteroskedastic-robust standard deviations are in parenthesis; iii) *** denotes significant at the 1% level, ** denotes significant at the 5% level; * denotes significant at the 10% level; iv) additional controls: years of schooling, potential labor market experience, years since migration, type of contract, marital status, children at home, previous unemployment experience, legal status, immigrant's region of origin and 19 dummies for Spanish Autonomous communities; v) Reference individual: married, has not been unemployed for more than three months in the past, has a non-permanent contract, resides legally in Madrid, comes from Central-Western Europe and has (within sub-sample) average schooling, experience and years since migration.

Source: Own elaboration.

Table 6: The impact of Spanish language ability on hourly earnings, by education groups – Restricted sample

		IV Combination of instruments					IV Combination of instruments			
	OLS	(a)	(b)	(c)	(d)	OLS	(a)	(b)	(c)	(d)
Panel 1. Workers with upper sec. educ. or more										
Spanish language proficiency	0.0857***	0.5831***	0.5808***	0.6845*	0.6845*	0.1188***	0.5997***	0.6002**	1.118**	1.1541*
	(0.0231)	(0.2201)	(0.2247)	(0.3586)	(0.3719)	(0.0252)	(0.2157)	(0.2179)	(0.6342)	(0.6637)
Occupation dummies	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Instruments										
Arrived before age 10	—	Yes	Yes	Yes	Yes	—	Yes	Yes	Yes	Yes
Child proficient in Spanish	—	Yes	Yes	No	No	—	Yes	Yes	No	No
plans to stay in Spain	—	Yes	No	Yes	No	—	Yes	No	Yes	No
R^2	0.5395	0.3471	0.3491	0.2512	0.2512	0.4921	0.2822	0.2818	0.2801	0.2818
Sargan-Hansen χ^2 test	—	0.19	0.19	0.01	—	—	1.19	1.19	.06	—
F-test	—	12.37***	18.17***	15.05***	29.28***	—	8.42***	12.52***	6.53***	12.74***
R^2 in first stage equation	—	0.3496	0.3494	0.3415	0.3412	—	0.3147	0.3146	0.3014	0.3012
Partial R^2		0.0251	0.0247	0.0131	0.0124		0.0243	0.0241	0.0053	0.005
Durbin-Wu-Hausman test	—	6.55**	6.18**	3.48*	3.19*	—	6.33**	6.19**	3.54***	3.44*
No. of observations	914	914	914	914	914	914	914	914	914	914

Note to Table 6: i) Source: Spanish National Immigrant Survey; ii) Heteroskedastic-robust standard deviations are in parenthesis; iii) *** denotes significant at the 1% level, ** denotes significant at the 5% level; * denotes significant at the 10% level; iv) additional controls: years of schooling, potential labor market experience, years since migration, type of contract, marital status, children at home, previous unemployment experience, legal status, immigrant's region of origin and 19 dummies for Spanish Autonomous communities; v) Reference individual: married, has not been unemployed for more than three months in the past, has a non-permanent contract, resides legally in Madrid, comes from Central-Western Europe and has (within sub-sample) average schooling, experience and years since migration.

Source: Own elaboration.

Table 6: THE IMPACT OF SPANISH LANGUAGE ABILITY ON HOURLY EARNINGS, BY EDUCATION GROUPS – RESTRICTED SAMPLE (continuation)

		IV Combination of instruments					IV Combination of instruments			
	OLS	(a)	(b)	(c)	(d)	OLS	(a)	(b)	(c)	(d)
Panel 2. Workers wih less than upper sec. educ.										
Spanish language proficiency	0.0221	0.1056	0.1038	0.0393	0.0384	0.0509*	0.2319**	0.2241**	0.1158**	0.1071*
	(0.0272)	(0.0801)	(0.0806)	(0.0814)	(0.0820)	(0.0271)	(0.0948)	(0.0956)	(0.1044)	(0.1056)
Occupation dummies	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Instruments										
Arrived before age 10	—	Yes	Yes	Yes	Yes	—	Yes	Yes	Yes	Yes
Child proficient in Spanish	—	Yes	Yes	No	No	—	Yes	Yes	No	No
plans to stay in Spain	—	Yes	No	Yes	No	—	Yes	No	Yes	No
R^2	0.2681	0.2956	0.2961	0.3037	0.3037	0.2451	0.2027	0.2062	0.2397	0.2411
Sargan-Hansen χ^2 test	—	2.67	2.65	0.04	—	—	2.17	1.67	0.69	—
F-test	—	25.21***	37.52***	32.99***	65.39***	—	24.03***	35.21***	21.55***	42.05***
R^2 in first stage equation		0.4947	0.4942	0.4853	0.4851		0.4409	0.4401	0.4149	0.4146
Partial R^2	—	0.1207	0.1199	0.1043	0.1041	—	0.0981	0.0968	0.0562	0.0557
Durbin-Wu-Hausman test	—	1.22	1.16	0.05	0.04	—	3.99**	3.63*	0.41	0.31
No. of observations	825	825	825	825	825	825	825	825	825	825

Note to Table 6: i) Source: Spanish National Immigrant Survey; ii) Heteroskedastic-robust standard deviations are in parenthesis; iii) *** denotes significant at the 1% level, ** denotes significant at the 5% level; * denotes significant at the 10% level; iv) additional controls: years of schooling, potential labor market experience, years since migration, type of contract, marital status, children at home, previous unemployment experience, legal status, immigrant's region of origin and 19 dummies for Spanish Autonomous communities; v) Reference individual: married, has not been unemployed for more than three months in the past, has a non-permanent contract, resides legally in Madrid, comes from Central-Western Europe and has (within sub-sample) average schooling, experience and years since migration.

Source: Own elaboration.

of alternative instruments. When immigrants whose mother language is Spanish are excluded from the sample, the estimated return is a little lower, about 16.5 pp. All in all, these figures are slightly above Dustmann & Fabbri's (2003) results for the returns to English language proficiency among immigrants in the UK and a little above the 13-18 pp range reported by Di Paolo (2011) and Di Paolo & Raymond (2012) for Catalan proficiency in the Spanish region of Catalonia. It is in our current research agenda to use a common specification and restrict our attention to Spain's bilingual regions in order to assess whether the returns to Spanish language proficiency are actually higher than the returns to regional languages in bilingual regions.

The results reported in the present paper may help policy makers devise strategies and immigration policies that promote and guarantee economic and social stability. It would be advisable to provide language courses for immigrants upon arrival especially in the current economic downturn. Still, there are profound differences in the earnings premium between immigrants with diverse levels of educational attainment. Most of the return found in the pooled sample is due to the highly educated. When we discriminate between education levels, we find that immigrants with less than upper secondary education reap substantially lower returns from Spanish proficiency than highly educated immigrants. The observed differential under the benchmark specification, 31.8 pp (47.9 against 16.1 pp), can be seen as a lower bound, given that, under other specifications and using a restricted sample, the low-educated fail to reap a statistically significant premium. This result warns us that indiscriminate language training policies may worsen economic inequalities among immigrants. In this respect, the provision of complementary schemes among the less qualified, including affordable education, professional training programs and guidance related to the Spanish labor market legislation and characteristics in order to ensure familiarity with work-related terms and usages might prove beneficial.

Due to the cross-sectional nature of our data set, it is imperative to emphasize the need for some caution in the interpretation of the findings. The patterns outlined in the present paper deserve scrutiny in longitudinal data. Controlling for individual time-invariant effects will allow us to establish more accurate causal relationships. Following people over time and exploring how the earnings of the same individuals change when they improve their language skills will allow us to attain this objective. An immigrant-based panel data survey in Spain would allow us to analyze the heterogeneous effects of divergent Spanish language fluency on earnings.

REFERENCES

Adsera, A. and Chiswick, B. (2007): "Are there gender and country of origin differences in immigrant labor market outcomes across European destinations?", *Journal of Population Economics,* vol. 20(3), pp. 495-526.

Angrist, J. and Pischke, S. (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2009.

Aldashev, A., Gernandt, J. and Thomsen, S. (2009): "Language usage, participation, employment and earnings: evidence for foreigners in West Germany with multiple sources of selection", *Labor Economics,* vol. 16(3), pp. 330-341.

Amuedo-Dorantes, C. and de la Rica, S. (2007): "Labor market assimilation of recent immigrants in Spain", *British Journal of Industrial Relations,* vol. 45(2), pp. 257-284.

Beenstock, M., Chiswick, B.R. and Paltiel, A. (2010): "Testing the Immigrant Assimilation Hypothesis with Longitudinal Data", *Review of Economics of the Household,* vol. 8(1), pp. 7-27.

Berman, E., Lang K. and Siniver E. (2003): "Language-skill complementarity: returns to immigrant language acquisition", *Labor Economics,* vol. 10(3), pp. 265-290.

Blázquez, M. and S. Rendón (2012): "Over-education in multilingual economies: Evidence from Catalonia", forthcoming in *International Migration*.

Bleakley, H. and Chin, A. (2004): "Language skills and earnings: evidence from childhood immigrants", *The Review of Economics and Statistics,* vol. 86(2), pp. 481-496.

Bleakley, H. and Chin, A. (2010): "Age at Arrival, English Proficiency, and Social Assimilation among US Immigrants", *American Economic Journal: Applied Economics*, American Economic Association, vol. 2(1), pp. 165-92.

Bound, J., Jaeger, D. and Baker, R. (1995): "Problems with instrumental variable estimation when the correlation between instruments and the endogenous explanatory variables is weak", *Journal of the American Statistical Association*, vol. 90(430), pp. 443-450.

Budría, S. (2005): "Education and Inequality: Evidence from Spain", in Asplund, R. and E. Barth (eds), *Education and Wage Inequality in Europe: A literature review*. ETLA, Series B209, Helsinki.

Carliner, G. (1981): "Wage differences by language group and the market for language skills in Canada", *The Journal of Human Resources,* vol. 16(3), pp. 384-399.

Casale, D. and Posel, D. (2011): "English language proficiency and earnings in a developing country: the case of South Africa", *Journal of Socio-Economics,* vol. 40, pp. 385-393.

Charete, M. and Meng, R. (1994): "Explaining language proficiency. Objective versus self-assessed measures of literacy", *Economics Letters,* vol. 44, pp. 313-321.

Chiswick, B. (1991): Speaking, Reading, and Earnings among Low-Skilled Immigrants. *Journal of Labor Economics,* vol. 9(2), pp.149-170.

Chiswick, B. (1998): "Hebrew language usage: Determinants and effects on earnings among immigrants in Israel", *Journal of Population Economics,* vol. 11, pp. 253-271.

Chiswick, B. and Miller, P. (1995): "The Endogeneity between language and earnings: international analyses", *Journal of Labor Economics*, vol. 13(2), pp. 246-288.

Chiswick, B. and Miller, P. (1999): Language skills and earnings among legalized aliens. *Journal of Population Economics,* vol. 12, pp. 63-89.

Chiswick, B. and Repetto, G. (2001): "Immigrants Adjustment in Israel: The Determinants of Literacy and Fluency in Hebrew and the Effects on Earnings" *International Migration: Trends, Policy and Economic Impact,* New York: Routledge, 2001, pp. 204-228.

Chiswick, B. and Miller, P. (2003): 'The complementarity of language and other human capital: immigrant earnings in Canada"*, Economics of Education Review*, vol. 22(5), pp. 469-480.

Chiswick B. and Miller, P. (2008): "A test for critical period hypothesis for language learning". *Journal of Multilingual and Multicultural Development,* vol. 29(1), pp. 16-29.

Chiswick, B. and Miller, P. (2010): "Occupational language requirements and the value of English in the US labor market", *Journal of Population Economics,* vol. 23, pp. 353-372.

Di Paolo, A. (2011): "Knowledge of Catalan, public/private sector choice and earnings: evidence from a double sample selection model", *Hacienda Pública Española,* vol. 197-(2), pp. 9-35.

Di Paolo, A. and Raymond, J.L. (2012): "Language knowledge and earnings in Catalonia", *Journal of Applied Economics,* vol. 15(1), pp. 89-118.

Dustmann, C. (1994): "Speaking fluency, writing fluency, and earnings among migrants", *Journal of Population Economics* vol. 7(2), pp. 133-156.

Dustmann, C. and van Soest, A. (2001): "Language fluency and earnings: estimation with misclassified language indicators", *The Review of Economics and Statistics*, vol. 83(4), pp. 663-674.

Dustmann, C. and van Soest, A. (2002): "Language and the earnings of immigrants", *Industrial and Labor Relations Review,* vol. 55(3), pp. 473-492.

Dustmann, C. and Fabbri, F. (2003): "Language proficiency and labor market performance of immigrants in the UK", *The Economic Journal*, vol. 113, pp. 695-717.

Fernández, C. and Ortega, C. (2008): "Labor market assimilation of immigrants in Spain: employment at the expense of bad job-matches?", *Spanish Economic Review*, vol. 10, pp. 83-107.

Friedberg, R. (2000): "You can't take it with you? Immigrant assimilation and the portability of human capital", *Journal of Labor Economics,* vol. 18(2), pp. 221-251.

Gao W. and Smyth, R. (2011): "Economic returns to speaking "standard Mandarin" among migrants in China's urban labor market", E*conomics of Education Review,* vol. 30(2), pp. 342-352.

Grenier, G. (1984): "The Effects of language characteristics on the wages of Hispanic-American males", *The Journal of Human Resources,* vol. 19, pp. 33-52.

Hayfron, J. (2001): "Language training, language proficiency and earnings of immigrants in Norway", *Applied Economics,* vol. 33, pp. 1971-1979.

Hu, W-Y.(2000): "Immigrant earnings assimilation: estimates from longitudinal data. *American Economic Review, Papers and Proc.* vol. 90 (May): 368-372.

Isphording, I.E. (2013). "Returns to Foreign Language Skills of Immigrants in Spain". *Labor*, vol. 27(4), pp. 443-461.

Lacuesta, A., Izquierdo, M. and Vegas, R. (2009). "Assimilation of immigrants in Spain: A longitudinal analysis", *Labor Economics*, vol. 6(6), pp. 669-678.

Lubotsky, D. (2007): "Chutes or Ladders: A Longitudinal Analysis of Immigrant Earnings", *Journal of Political Economy*, vol. 115(5), pp. 820-867.

Lui, H.K. (2007): "The returns to language ability in Hong Kong: before and after the handover", *Applied Economic Letters*, vol. 14(2), pp. 121-125.

McManus, W., Gould, W. and Welch, F. (1983): "Earnings of Hispanic men: the role of English language proficiency", *Journal of Labor Economics,* vol. 1(2), pp. 101-130.

Ministry of Labor and Social Affairs (*Ministerio de Trabajo y Asuntos Sociales*) (2007): "Strategic plan for citizenship and integration", *Subdirección General de Información Administración y Publicaciones.* Available at: http://extranjeros.empleo.gob.es/es/IntegracionRetorno/Plan_estrategico/pdf/PECIingles.pdf

OECD (2012): "Migration policy developments, *International Migration Outlook 2012, OECD Publishing.* Available at: http://dx.doi.org/10.1787/migr_outlook-2012-6-en

Quella, N. and Rendón, S. (2012): Occupational Selection in Multilingual Labor Markets: The Case of Catalonia, *International Journal of Manpower*, vol. 33(8), pp. 918-937.

Rendon, S. (2007): "The Catalan premium: language and employment in Catalonia", *Journal of Population Economics*, vol. 20, pp. 669-686.

Rivera-Batiz, F. (1992): "English proficiency and the earnings of young immigrants in the U.S. Labor Markets", *Review of Policy Research,* vol. 11(2), pp. 165-175.

Rodríguez-Planas, N. (2012): "Wage and occupational assimilation by skill level: migration policy lessons from Spain", *IZA Journal of European Labor Studies*, vol. 1(8).

Simón, H., Sanromá, E. and Ramos, R. (2008): "Labor segregation and immigrant and native-born wage distributions in Spain: an analysis using matched employer–employee data", *Spanish Economic Review*, vol. 10(2), pp. 135-168.

Stock, J.H., Wright, J. and Yogo, M. (2002): *"*A survey of weak instruments and weak identification in generalized method of moments"*, Journal of Business & Economic Statistics, American Statistical Association*, vol. 20(4), pp. 518-529.

Tainer, E. (1988): "English Language Proficiency and the Determinants of Earnings among Foreign-Born Men". *The Journal of Human Resources,* vol. 23(1), pp. 108-122.

Wang, C. and Wang, L. (2011): "Language Skills and the Earnings Distribution among Child Immigrants", *Industrial Relations*, vol. 50(2), pp. 297-322.

Zhen, Y. (2013): "The Effects of English Proficiency on Earnings of U.S. Foreign-Born Immigrants: Does Gender Matter?" *Journal of Economics and Finance*, vol. 1(1), pp. 27-41.

Fecha de recepción del original: agosto, 2013
Versión final: julio, 2014

RESUMEN

Este artículo utiliza microdatos de la Encuesta Nacional de Inmigrantes con el fin de analizar el impacto del aprendizaje del Castellano sobre los ingresos de los inmigrantes. Los resultados basados en Variables Instrumentales (VI) apuntan a un retorno al Castellano de alrededor del 20%. La cifra varía en gran medida entre los inmigrantes con niveles de educación alto y aquellos con niveles de educación bajos. Aquellos con un nivel de estudios alto obtienen una prima del 50%, mientras que los inmigrantes con bajos niveles de educación no alcanza a ser estadísticamente significativo bajo algunas especificaciones. Esta evidente complementariedad entre educación y competencias lingüísticas es un desafío para las políticas de inmersión lingüísticas dirigidas a inmigrantes.

Palabras clave: inmigración, competencia en idiomas, ingresos, variable instrumentales.

Clasificación JEL: F22, J24, J61.