

Héctor Bellido, Miriam Marcén\* and Marina Morales

# Spaniards in the wider world: the role of education in the choice of destination country

## Abstract

This paper examines the relationship between the education level of Spanish emigrants and their destination country. Since Spanish emigrants were born under the same laws and institutions, the differences in their destination countries can be due to dissimilarities in their level of education. To explore this, we use census microdata, covering the period from 2000 to 2007, of 21 countries with Spanish emigrants. Results suggest that with low unemployment rates, English- and Spanish-speaking countries are the most likely to become the host countries for more educated individuals, regardless of their location.


**Current version:** April 19, 2021  
**Keywords:** education, migration, Spain  
**JEL codes:** I20; F22  
**Corresponding author:** Miriam Marcén  
mmarcen@unizar.es

---

Department of Economic Analysis, Faculty of Economics and Business, University of Zaragoza, Zaragoza, Spain

---

© The Author(s). 2021. Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

 Cite as: Bellido et al. *IZA Journal of Development and Migration* (2021) 12:06

<https://doi.org/10.2478/izajodm-2021-0006>

## 1 Introduction

In the 2013 Christmas period, almost all Spanish television (TV) channels broadcast an advertisement with the slogan “Become foreigner”.<sup>1</sup> This simply reflected the situation of Spanish migration. From 2009 to 2015, the Spanish expatriate population increased by almost 50% (*Instituto Nacional de Estadística* [INE]). In January of 2009, there were <1.5 million Spanish nationals living in other countries, whereas by January of 2015, there were around 2.2 million Spanish emigrants (*Padrón de Españoles Residentes en el Extranjero* [PERE] 2009, 2015).<sup>2</sup> Despite this increase, the pattern of distribution did not change substantially, with >60% of Spanish emigrants living in the Americas, 30–35% in Europe, and <5% in the rest of the world. The differences are quite remarkable at the country level: 423,006 Spanish nationals live in Argentina, 223,636 in France, 190,600 in Venezuela, 119,662 in Cuba, 113,500 in the US, 91,316 in the UK, around 23,000 in Colombia and Peru, and around 11,000 in Bolivia and Portugal (PERE 2015). This is not simply a consequence of the latest migration process: from the early 20th century to the 1970s, there were several waves of Spanish immigration to South America and Europe (Bover and Velilla 2005).<sup>3</sup> This raises the following question: why do Spanish emigrants settle in so many different countries?

As explained by Bauer et al. (2005), when analyzing patterns of migration, it is not only the decision to emigrate that is relevant, so is the location choice. Several researchers have studied the mechanisms involved in the emigrant’s selection of location, including economic incentives (Eliasson et al. 2003; Nifo and Vecchione 2014), native language of the destination country [see Chiswick and Miller (2014) for a review], immigration policy barriers (Clark et al. 2007), past colonial relationships (Mayda 2010), institutions (Nifo and Vecchione 2014), physical distance (Sjaastad 1962), religious affiliations (Neudörfer and Dresdner 2014), labor market factors (Clemente et al. 2016; Dotti et al. 2013), and cultural links (Belot and Ederveen 2012; Wang et al. 2016), among others. This work explores whether there is a relationship between the destination country and the education level of Spanish emigrants.

The existing literature examining the connection between education and migration generally finds that more-educated individuals are more likely to emigrate, and that those individuals tend to locate in countries with high rewards for skills (Grogger and Hanson 2011). The business cycle also appears to be relevant for the migration choices of more-educated individuals (Venhorst et al. 2011). Other papers studying the educational selectivity of emigrants conclude that cultural similarities and the physical distance between countries have more relevance in the educational selectivity process than wages or immigration policies (Belot and Hatton 2012). In addition, there is a growing literature that analyzes the migration of high-skilled populations from poor to rich countries, which, as researchers suggest, may affect the education incentives in developing countries (Beine et al. 2008; Docquier et al. 2007).

This paper extends prior research by using microdata from the Census related to 21 destination countries of Spanish emigrants in the period 2000–2007. The data come from the

1 In Spanish, the slogan of the advertisement was “*Hazte Extranjero*”. The company was Campofrío.

2 This sample includes individuals whose nationality is Spanish and who are registered in the embassies and consulates located in the countries where they are ordinarily resident. This information is only reported from 2009. See [www.ine.es](http://www.ine.es).

3 Note that this work does not explore another kind of geographical mobility that occurs in Spain, such as the interregional mobility of the Spanish (for this analysis, see Juárez 2000) or the mass waves of immigration that occurred at the beginning of the 21st century in Spain (González and Ortega 2011).

Integrated Public Use Microdata Series–International (Minnesota Population Center 2014). The use of Census data is quite common in the literature, but the majority of papers consider the analysis at the aggregate level. The microdata from the Census do not contain information on why individuals decided to migrate, but the Census does provide valuable information on Spanish emigrants, which allows us to study the relationship between their education level and several characteristics of the destination country: the language, the physical distance from Spain, and the economic circumstances measured by the unemployment level. We first analyze each of them individually, and then we consider them jointly. We add to the literature by studying – at the individual level – the dissimilarities in the location choice of Spanish emigrants depending on their level of education. The novel methodology proposed here allows us to jointly consider the location choice of individuals based on language, migration costs, and job search.

The native language of the host country is one of the characteristics that may be important in the location choice of emigrants. However, researchers have found conflicting results in the relationship between language and choice of destination country. Some conclude that language matters (Adsera and Pytlikova 2015; Beine et al. 2011; Belot and Hatton 2012; Belot and Ederveen 2012; Clark et al. 2007; Grogger and Hanson 2011), while others detect no relationship (Mayda 2010; Ortega and Giovanni 2009). Language skills can be seen as a form of human capital, since they satisfy the following requirements: they are productive (knowing the language of your place of residence facilitates access to a job), they are costly to produce (time and money are needed to acquire language proficiency), and they are embodied in the individual (language skills may not be separated from the person; Chiswick and Miller 2014). If it is assumed that the greater the human capital, the greater is the gift for languages, it would be expected that those Spanish emigrants with a higher education level would be more likely to settle in countries where Spanish is not the official language. However, those who are highly educated also have higher opportunity costs derived from migration to non-Spanish-speaking countries. Since learning a language can be costly, not only because of the direct costs of learning but also because of the lower earnings received in the destination country during the adjustment period before acquiring proficiency, it can be argued that highly educated ones are more likely to settle in countries in which they already know the language; for example, countries with the same language as the home country or countries where emigrants do not have to spend a long time to acquire proficiency, either because the native language of the destination country is a second- language in the schools of their country of origin or because it is a common language in business. Thus, a priori, it is not possible to infer which languages would be preferred for Spanish emigrants with a higher, or lower, level of education, and so, in theory, it is unclear what destination country they would choose.

The physical distance between the home and the host country can also play a role in the location choice of emigrants. It would be expected that low-education-level individuals, who normally have low wage-earning potential, would be less likely to move to distant areas because the travel costs between the home and the destination country are higher. Additionally, physical distance is important to the quality of information about the destination country (job opportunities, income differentials, job market characteristics, and so on), as noted in Sjaastad (1962). If highly educated individuals have greater access to such information, we would expect a less-significant impact of distance on highly educated individuals than on less-educated

individuals. Proximity also facilitates illegal immigration, which may increase the migration of less-educated individuals. In this case, all these possible explanations indicate that, for highly educated individuals, the physical distance between the home and the host country is less important than for low-education-level emigrants, although we will check it here empirically.

Migration can be a consequence of job search (Herzog et al. 1993; Malamud and Wozniak 2012).<sup>4</sup> Thus, the labor market situation of the destination country is a characteristic that emigrants, potentially, take into account when deciding whether or not to move to that specific country. If those who are highly educated are more likely to access and understand the labor market situation of a potential destination country, it would be expected that they would choose countries with lower unemployment rates (URs), where they can find a job more easily. In addition, since the UR is a proxy of a country's business-cycle situation, it would also be expected that those same individuals would locate in countries with low URs, where they can obtain higher rewards for their skills. Once again, we will check this empirically.

After the analysis of the relationship between the educational level of emigrants and each of the previous characteristics of the host countries (language, physical distance, and economic situation), we introduce a novel contribution to the literature by exploring different combinations of host country characteristics, since emigrants may make their location choice considering, for example, all those features simultaneously. Additionally, it is arguable that some of the characteristics of the destination countries may be more important for emigrants than others, so a separate analysis for each would not be sufficient to study the location choice of emigrants. In this work, we propose a methodology that allows us to examine which of the characteristics of the destination country (or which among them) is (are) the most relevant, without imposing any restriction on the location choice process. Results suggest that, with low URs, English- and Spanish-speaking countries are the most likely to become the host countries for more-educated individuals, regardless of their location. However, it is also observed that the greater the proximity, the less important is the economic situation; whereas, the greater the distance, the more central is the labor economic situation of the country for highly educated individuals.

The remainder of the paper is organized as follows: Section 2 presents our empirical strategy, Section 3 describes the data, results are discussed in Section 4, and Section 5 concludes the paper.

## 2 Empirical Strategy

The main purpose of this research is to determine whether a relationship exists between the level of education of Spanish emigrants and certain characteristics of the host country (characteristics considered both separately and jointly). This analysis is necessary because, as explained above, the incentives, costs, and potential gains that emigrants can obtain in the destination country may differ depending on their level of education. We focus on three different characteristics: the language of the destination country, the physical distance from the emigrants' home country, and the economic situation of the host country, measured by its UR, which have been extensively considered in the literature as crucial factors in the emigration choice. In most cases, prior work has analyzed those characteristics separately, using aggregate data.

---

<sup>4</sup> Wage inequality may also play a role, but we do not consider this in our work.

In our case, we suppose that emigrants may decide to settle in a country by considering just one of those characteristics or several of them. Since we have no information on what motivates emigrants to settle in one particular country rather than in another, we are not able to analyze how the location choice process takes place. Nevertheless, we know the outcome of that decision process, that is, the place of residence of emigrants. This gives us information, for example, on whether emigrants with a low level of education tend to locate in countries with certain specific characteristics, whereas those with a high level of education choose other countries with different characteristics. Thus, we can study whether emigrants choose their destination country depending on their level of education. To address this issue, we use a model for nominal outcomes, specifically a multinomial logit model (MNL), in which we estimate a separate binary logit for each pair of outcome categories (Nerlove and Press 1973). This methodology has also been considered in other papers to analyze migration choices (Huber and Nowotny 2013).

Initially, we examine the relationship between each of the characteristics of the host country and the level of education of the emigrants. Within this framework, we explore whether emigrants with high level of education tend to live in countries with a language other than their native language, or not, relative to those with low levels of education. We also analyze whether emigrants with high levels of education settle in distant countries, that is, far from their home countries, and whether they choose countries with a better economic situation or not. We then consider all the possible combinations of the characteristics of the destination countries and, again, we examine their relationship with the level of education of the emigrants. The results of the joint analysis are compared with the initial separate analysis, to study whether the latter is useful. Note that we do not impose any order in the location choice process, neither in the separate analysis nor in the joint analysis, but with the methodology proposed here, we can observe the characteristics of the destination country that are more (or less) likely to be relevant for emigrants in the choice of a destination country, depending on their level of education. Formally, we estimate the following equation:

$$\ln \phi_{m/b} = \ln \frac{\Pr(y = m/\mathbf{x})}{\Pr(y = b/\mathbf{x})} = \mathbf{x}\beta_{m/b} \text{ for } m = 1 \dots J \quad (1)$$

In Eq. (1),  $b$  is the base category and  $m$  varies from “1” to  $J$ , with  $J$  being the total number of outcome categories (we revisit this issue later). The vector  $\mathbf{x}$  includes not only the educational variables (defined as “secondary education completed”, “some college education or more”, and “primary or less than primary education completed”, which is the omitted measure) but also a range of other socioeconomic characteristics of Spanish emigrants. These characteristics may be considered to have an impact on the location decision for reasons independent of their level of education and include age, gender (whether the emigrant is a man), and employment status (whether the emigrant is employed).<sup>5</sup>

Regarding the outcome categories in the separate analysis, when the dependent variable is the official or the most commonly spoken language of the host country, we consider three categories: English, Spanish, and Others. As explained earlier, it would be expected that more-educated individuals, that is, those with greater human capital (Barro and Lee 2001), would be more easily capable of living in non-Spanish-speaking countries. Nonetheless, because of the

<sup>5</sup> We have repeated the analysis without those controls, and our results do not vary.

opportunity costs to acquire language proficiency, it would be expected that these more-educated individuals choose to live in countries with language in which they have already acquired proficiency, with the most-likely countries being Spanish- and English-speaking countries. As Chiswick and Miller (2014) explain, Spanish is the first language in their country of origin and English is normally the second language in Spanish schools, in addition to being one of the most common languages in business. The outcome categories in the case of physical distance are Europe, the Americas, and Others. For the UR, the categories are as follows: low unemployment (UR <4%), medium unemployment (UR between 4% and 9.9%), and high unemployment (UR ≥10%).<sup>6</sup> As the literature indicates (Belot and Hatton 2012; Clark et al. 2007), the greater the physical distance, the greater are the migration costs and the lower is the possible knowledge of the situation (institutional, social, and economic) of the potential destination country; hence, we would expect that low-education-level emigrants relocate to closer countries (lower travel costs, greater possibilities for illegal migration, more facilities to gain prior knowledge of the country, and ease of return to the country of origin). With respect to the UR, it would be expected that more-educated individuals settle in countries with low URs where they can have greater rewards for their skills. This could be so since they have greater access to information on the economic situation of the destination country, and they are also more capable of understanding it.

We recognize that the use of a model for nominal outcomes in all cases can generate some concerns, since some of the destination country characteristics can be considered as ordered variables. The language of host countries is clearly a nominal variable because all the categories herein are assumed to be unordered. Thus, in principle, the use of a MNL for this analysis would be appropriate. In contrast with the language outcome category, the distance and the UR can be considered as partly ordered – or even ordered – which would imply using an ordinal model in our empirical analysis (McKelvey and Zavoina 1975). However, the parallel regression assumption implicit in the ordinal model is not satisfied in any of these cases. We calculate the approximate likelihood-ratio test of proportionality of odds across outcome categories and detect that the parallel regression assumption can be rejected at the 5% level of significance. In addition, it is possible to argue that these variables are not considered as ordered by emigrants. Because of vacation periods, the schedule, and perhaps even climate, among others, emigrants may see one part of the world as being different from the other. Of course, this can be correlated with physical distance (close countries usually have more similar climates and even customs), but emigrants can observe it as unordered choice. In these cases, the outcomes can be treated as nominal (Marcén and Molina 2012). In the case of the joint analysis, this problem does not exist. The combinations of all these variables (language, distance, and UR) cannot be considered ordered outcomes. Then, it is arguable that the models for nominal outcomes, in particular, the MNL model, are suitable.<sup>7</sup>

6 It has been observed that all these outcome categories are indistinguishable on testing whether any of the independent variables significantly affects the odds of  $m$  versus alternative  $n$  categories (Anderson 1984). We test the following:

$H_0: \beta_{1,m|n} = \dots = \beta_{i,m|n} = \dots = \beta_{i,m|n} = 0$ , where  $\beta_i$  is the coefficient associated with the explanatory variable  $x_i$ . We cannot reject the hypothesis that those outcome categories are indistinguishable when using logistic regression likelihood-ratio test (LR) and Wald tests.

7 It is worth noting that we test the property of independence of irrelevant alternatives. This property is convenient in MNL models, since if alternatives are not truly independent, then the estimates will be inconsistent. Our results indicate that the MNL model is appropriate, using the Hausman test (Hausman and McFadden 1984) and the Small-Hsiao test (Small and Hsiao 1985).

With respect to the outcome categories in the joint analysis, we include all possible combinations without imposing any order on the location process decision. We repeat the analysis by incorporating, as the dependent variable, a combination of the location of the destination country and the official (or the most common) language of that country. Within this framework, we are interested in examining whether emigrants with higher levels of education are more likely to choose a destination country with a specific language, depending on its location. For example, it is possible to surmise that emigrants with low levels of education are more likely to settle in less-distant countries (low travel costs), regardless of the language, since acquiring language proficiency is less important to their low-skilled employment, whereas highly educated emigrants who decide to migrate to a distant country (high travel costs) may prefer those countries with the same language as the home country (or countries where they do not have to spend a lot of time to acquire language proficiency) in order to reduce the migration costs. Similarly, we run the entire analysis defining the dependent variable as a combination of the location of the destination country and its economic situation. As before, in this case, we can explore whether emigrants with high levels of education tend to settle in destination countries with a better economic situation, if they are distant, whereas the economic situation is less important for them when the destination is close to their home country. Because all the characteristics considered here appear to be important, it is logical to propose a combination of all of them in the dependent variable. With this approach, we are able to study whether, depending on the level of education of emigrants, they decide to settle in a country considering just one of those characteristics or several of them.

We analyze the dynamics among the outcome categories by utilizing odds ratios, which is an intuitive method of interpreting the estimates (Greene 2008; Long and Freese 2006). Holding other variables constant, the change factor in the odds of outcome category  $m$  versus outcome category  $n$  when  $x_i$  increased by  $\delta$  is represented as follows:

$$\frac{\phi_{m/n}(\mathbf{x}, x_i + \delta)}{\phi_{m/n}(\mathbf{x}, x_i)} = e^{\beta_{i,m|n}\delta} \quad (2)$$

For a unit change in  $x_i$ ,  $\delta = 1$ , the odds of  $m$  versus  $n$  are expected to change by a factor of  $\exp(\beta_{i,m|n})$ , holding all other variables constant. For a standard deviation change in  $x_i$ ,  $\delta = s_{x_i}$ , the odds of  $m$  versus  $n$  are expected to change by a factor of  $\exp(\beta_{i,m|n} \times s_{x_i})$ . To simplify the odds analysis, odds-ratio plots are also presented (Long and Freese 2006).

### 3 Data

To implement our analysis, we use data from the Integrated Public Use Microdata Series–International of the 2000–2007 Censuses (Minnesota Population Center 2014).<sup>8</sup> Our sample consists of emigrants (first-generation migrants) from Spain who moved to 21 countries in total.<sup>9</sup> The number of countries can be considered reduced, but these are all the countries containing information on the birthplace in the period considered. It is worth noting that almost

<sup>8</sup> This period was chosen since it is possible to get the greatest number of destination countries of Spanish emigrants for those years.

<sup>9</sup> Census data from only one point in time have been used for the different destinations in the main sample. See, for example, the notes to Table 2.

70% of Spanish emigrants choose the countries considered in this work to settle in, representing the majority of their main destinations (PERE 2009, 2015).<sup>10</sup> Individuals between the ages of 25 and 59 years are retained, since everyone in this sample is likely to have completed schooling and is below the retirement age. Our final sample consists of 54,317 Spanish emigrants.<sup>11</sup>

Table 1 shows the summary statistics of the relevant variables by destination country. Our individuals are 45.34 years old on average, with those living in Argentina being older than those in other groups (52 years) and those in Ireland being the youngest (30 years). Almost 50% of emigrants are men, although there are differences, with more men than women living in the countries of South and Central America. Most of the emigrants are married and employed. With respect to the education level of Spanish emigrants, Table 1 presents considerable variations, with less-educated individuals living in Argentina, Portugal, and the Philippines, whereas more highly educated ones are located in Colombia, Peru, and El Salvador. The differences in education may point to a potentially important role of education in the location choice of Spanish emigrants. With respect to the number of observations, we recognize that the large number of observations from the France Census can generate some concerns.<sup>12</sup> We revisit this issue in Section 4.

Columns 7–9 of Table 1 present the summary of the characteristics of the destination countries. Column 7 includes the official or the most common language in each host country, according to data from the US Central Intelligence Agency (CIA). Twelve countries have Spanish as the official language, three have English, and the remaining six countries have other languages (Portuguese, Greek, German, French, and Filipino). As mentioned earlier, we consider three categories (English, Spanish, and Others). In the literature, it is possible to find several classifications of languages when many home and destination countries are analyzed (Chiswick and Miller 2014). However, in this work, we only consider one home country and relatively few host countries, with the number of languages being limited. For this reason, we favor the use of that simple classification.

With respect to the UR, the data come from the World Data Bank. Column 8 only includes information on the UR for the year in which the Census was collected in each country. We do

---

10 There are three exceptions: the UK, Germany, and Switzerland. With respect to the UK, we repeat the analysis by including information on that country, since this is available for 1991. Results do not vary (see below). Unfortunately, this could not be rerun to incorporate Germany and Switzerland because of the lack of information in the Census. Spanish emigrants who were born in Spain and live in Germany and Switzerland represent around 8% and 7% of the total Spanish emigrants, respectively (PERE 2009, 2015). Other destination countries that could not be included in the analysis, because of the lack of data, received <1% of the Spanish emigrants each (PERE 2009, 2015). In this setting, it is possible to argue that any of the destination countries incorporated in our analysis is driving the results. To tackle this issue, we have reestimated all the specifications by excluding one country at a time, and the results are maintained. Moreover, we have repeated the analysis by including information on the year of migration for countries with that information, and in this case, we are able to include and exclude the years of major waves of migrations to the omitted destination countries, and the results do not vary.

11 We only have information on those individuals who decide to migrate. In this framework, it is possible to argue that there can be a selection problem, that is, emigration could be nonrandomly determined. Since we are comparing individuals with different levels of education, the potential selection problem would be a concern if, for example, among those individuals with a high level of education, only those who are less risk-averse decide to migrate to a specific country, whereas among those with a low level of education, only those who are more risk-averse decide to do so. If this were the case, our estimated coefficients would be capturing the effect of differences in the risk-averse behavior, rather than the impact of the level of education on their location choice. Although we recognize that this is difficult to explore with our data set, to mitigate the concerns, we note that our results are maintained over time, which may reinforce our findings, since even during the periods in which the migration decision could be motivated by dictatorship or by the Spanish Civil War, the location choice of the individuals by level of education was quite similar to that of those individuals who were not forced to migrate. We have also repeated the analysis excluding those emigrants living in each destination country, one at a time and by age group. Results do not vary.

12 All the estimates are obtained using weights from the Census. Results without weights do not change.



**Table 1** Summary statistics by destination country: countries sorted alphabetically

Country	1	2	3	4	5	6	7	8	9	10
	Age, years	Males	Secondary education completed	College or more	Married	Employed	Official/most common language	Unemployment rate, %	Continent	Observations, n
Argentina	51.97	0.49	0.31	0.17	0.73	0.58	Spanish	18.3	America	3,372
Austria	38.80	0.37	0.28	0.34	0.37	0.73	German	3.6	Europe	132
Bolivia	41.06	0.56	0.32	0.64	0.47	0.64	Spanish	5.4	America	77
Brazil	50.44	0.58	0.27	0.36	0.80	0.65	Portuguese	9.5	America	737
Chile	41.65	0.57	0.24	0.69	0.69	0.70	Spanish	8.9	America	379
Colombia	42.15	0.60	0.12	0.85	0.62	0.76	Spanish	12.0	America	96
Costa Rica	42.55	0.57	0.21	0.67	0.64	0.70	Spanish	5.1	America	86
Cuba	50.36	0.45	0.40	0.33	0.67	0.55	Spanish	3.3	America	55
El Salvador	42.24	0.52	0.16	0.72	0.84	0.80	Spanish	6.3	America	25
France	47.01	0.47	0.12	0.21	0.66	0.74	French	8.8	Europe	41,193
Greece	39.87	0.22	0.31	0.50	0.72	0.54	Greek	10.2	Europe	68
Ireland	30.43	0.36	0.30	0.66	0.18	0.79	English	4.2	Europe	343
Mexico	43.65	0.59	0.30	0.51	0.82	0.70	Spanish	2.6	America	850
Nicaragua	41.56	0.44	0.26	0.70	0.67	0.78	Spanish	5.6	America	27
Panama	43.52	0.58	0.26	0.49	0.76	0.79	Spanish	13.5	America	96
Peru	42.42	0.61	0.22	0.74	0.57	0.57	Spanish	4.5	America	195
The Philippines	38.68	0.40	0.10	0.41	0.66	0.00	Filipino	11.2	Asia	19
Portugal	39.56	0.44	0.20	0.34	0.75	0.69	Portuguese	4.0	Europe	259
South Africa	47.07	0.57	0.60	0.21	0.79	0.65	English	25.4	Africa	23
The United States	38.80	0.48	0.23	0.66	0.66	0.73	English	4.0	America	3,248
Venezuela	48.53	0.53	0.34	0.29	0.76	0.66	Spanish	12.8	America	3,037
Mean	45.34	0.49	0.21	0.36	0.68	0.70		8.75		
Standard deviation	9.69	0.50	0.41	0.48	0.47	0.46		4.47		

Note: Data come from the Integrated Public Use Microdata Series—International microdata for the period 2000–2007.

not detect a clear regional pattern, with the highest rates being for Argentina and South Africa and the lowest for Mexico, Cuba, and Austria. The analysis has been repeated using the average UR from 1990 to 2000, to overcome differences in the years in which the information is taken. Results are maintained. Regarding the physical distance between the home country and the host country, we separate the countries by continent. There are five countries from Europe, 14 from the Americas, and just two from other parts of the world (Column 9). Although not included in this document, we have also used the physical distance measured in kilometers from the capital of the country of origin to the capital of the destination country, as in Belot and Hatton (2012). Results do not vary.

## 4 Results

### 4.1 Destination countries classified by language

Table 2 shows the factor changes in the odds of one outcome category versus the others. The dependent variable is the official or the most common language, and the specifications incorporate all the controls (age, gender, marital status, and employment status). Regarding the relationship between the education level and the location choice, as can be seen in Column 1,

**Table 2** Location choice: destination countries classified by language (dependent variable: English, Spanish, and Others)

Variables	1	2	3	4	5	6	7	8
<b>Secondary education completed</b>								
English–Others	8.797***	2.328***	2.862***	2.328***	2.881***	2.328***	2.117***	2.328***
English–Spanish	2.027***	1.396**	2.047***	1.396**	2.100***	1.396**	1.509***	1.396**
Others–Spanish	0.230***	0.600***	0.715***	0.600***	0.729***	0.600***	0.713***	0.600***
<b>Some college/university completed</b>								
English–Others	11.937***	2.752***	4.127***	2.752***	4.159***	2.752***	2.628***	2.752***
English–Spanish	3.680***	1.548***	3.720***	1.548***	4.152***	1.548***	1.606***	1.548***
Others–Spanish	0.308***	0.563***	0.901	0.563***	0.998	0.563***	0.611***	0.563***
Year dummies	No	Yes	No	Yes	No	Yes	Yes	Yes
McFadden's adjusted $R^2$	0.161	0.713	0.144	0.471	0.154	0.460	0.517	0.735
Observations	54,317	54,317	13,124	13,124	12,781	12,781	12,865	54,518

*Notes:* Sample consists of Spanish emigrants in Argentina 2001, Austria 2001, Bolivia 2001, Brazil 2000, Chile 2002, Colombia 2005, Costa Rica 2000, Cuba 2002, El Salvador 2007, France 2006, Greece 2001, Ireland 2002, Mexico 2000, Nicaragua 2005, Panama 2000, Peru 2007, Philippines 2000, Portugal 2001, South Africa 2001, United States 2000, and Venezuela 2001. The dependent variable is the official or the most common language of the destination countries. Languages are classified into the following categories: English, Spanish, and Others. This table shows the factor change in the odds ratios,  $\exp(\beta_{i,m|n})\delta$ , with  $\delta=1$ . Columns 1–8 include the following as controls: age, age square/100, gender (male), being married, and being employed. Columns 3 and 4 exclude France. Columns 5 and 6 include information on those countries collecting their censuses from 2000 to 2002. Column 7 excludes France and Portugal. Column 8 includes UK 1991. Estimates are weighted. Standard errors are robust.

\*Significant at the 10% level.

\*\*Significant at the 5% level.

\*\*\*Significant at the 1% level.

those with at least some college education are more likely to settle in English-speaking countries, since the odds of living in an English-speaking country versus Others or versus Spanish-speaking countries are almost 12 and 3.7 times greater, respectively, holding the remaining variables constant. This result may suggest that the capacity to learn a language is important when choosing a country of residence. However, for those with at least some college education, the results reveal that Spanish-speaking countries are more likely to be chosen than countries with other languages, since the factor change is  $<1$ , whence it may be inferred that not only is the capacity of those more educated relevant, but also the opportunity costs. Results do not vary when controls for year dummies are added in Column 2. Their incorporation can be considered as being necessary, since not all the Censuses were collected in the same year. In this setting, it is possible to argue that those differences are affecting our estimates, if, for example, in a specific year, there had been a massive migration to a destination country due to changes in the migration laws or in the economic situation of both the home country and the potential host country. Year dummies are supposed to pick up the effect of those unobserved characteristics. It should be noted that McFadden's adjusted  $R^2$  is quite high in Column 2, which may indicate some problems of collinearity after the incorporation of the year dummies.<sup>13</sup>

Another concern with these estimates is the inclusion of Spanish emigrants living in France, since the number of observations is quite high in that Census. This should not be surprising since the main destination of Spanish emigrants, for almost 18% of them, is France (PERE 2009). To tackle that problem, we repeat the analysis without the Spanish emigrants in France, in Columns 3 and 4 (with year dummies). Results are unchanged. We also find no changes in the estimates when we only consider in the sample those Censuses that were collected from 2000 to 2002, as shown in Columns 5 and 6 (with year dummies). The 2000–2002 sample is chosen since most of the Censuses, 16 of 21, were collected between 2000 and 2002, and thus any impact of the differences in the year of the Census collection should be mitigated. To provide even more evidence that these results are not driven by the sample of destination countries, we have repeated the analysis by excluding Spanish emigrants living in the two neighboring countries, namely, France and Portugal (Column 7). Results are the same.

Table 2 also includes the estimates after adding Spanish emigrants living in the UK. Although there are no observations for the period considered (2000–2007) for UK, the 1991 UK Census has information on Spanish emigrants. Despite the concerns that may be generated by the incorporation of such information because of the differences in the periods in which data were collected, we check whether the results are maintained after incorporating this destination (see Column 8). Adding year dummies (in this case, it is even more necessary to capture the unobservable characteristics, due to the significant differences in the years of the Census), the location choices are again English, Spanish, and Others, for both Spanish emigrants with secondary education completed and with at least some college education. As before, we recognize some concerns with the high values for McFadden's adjusted  $R^2$ . Although not included in the table, we observe lower  $R^2$  once observations from France are excluded in all specifications, and the findings do not vary substantially.

---

<sup>13</sup> The analysis has been repeated including, and excluding, all controls, and the findings do not change.

To further test the robustness of the results, we have also considered different subsamples, varying the age of the Spanish emigrants. Older Spanish emigrants could have been forced to migrate to their destination country as a consequence of the Spanish Civil War (1936–1939) and the subsequent dictatorship period (1939–1975). But, even in this case, the final destination country of those Spanish emigrants could be related to their level of education. To address this issue, we have considered different age groups. The samples consist of individuals aged 25–64, 25–49, 30–49, and 25–39 years.<sup>14</sup> The results are presented in Table A1. Regardless of the age groups analyzed, the results point to an important role of the level of education of emigrants in the location choice. In all the samples considered, we observe that the most-likely choice for those with at least some college education is an English-speaking country. Next, they prefer Spanish-speaking countries, with countries with another official language being their least-likely choice. Another way to test this issue would be to select individuals by year of migration. Unfortunately, this information is only available for 10 destination countries (Bolivia, Chile, Colombia, Costa Rica, El Salvador, France, Greece, Ireland, Nicaragua, and the United States). This is somewhat problematic, since that group of countries cannot be considered a random sample of Spanish emigrants' destination countries. It does not incorporate the main destinations of Spanish emigrants, such as Argentina, Cuba, or Venezuela. In addition, these Spanish-speaking countries have a greater proportion of Spanish emigrants with at least some college education. In Table 1, we observe that around 70% of Spanish emigrants in these Spanish-speaking countries have at least some college education, whereas in, for example, Argentina, only 17% of Spanish emigrants have at least some college education. When the year of migration is used to select the sample of those arriving in the destination country, from the Spanish Civil War until 1960 (and even until 1970), we observe that more-educated individuals are more likely to settle in Spanish-speaking countries. However, when we repeat the analysis with only those 10 destination countries and use the main sample, that is, without considering the year of migration, we obtain the same results: Spanish-speaking countries are the most-likely choices. Additionally, the year of migration information is not particularly useful, due to the significant deficiencies in the sample.

The decision to migrate for Spanish emigrants could have been for reasons of education. For example, emigrants could move to a specific country to obtain a University degree, or they could have arrived in the destination country at a young age, as a consequence of their parents' migration decision, and so all their education would have taken place in the destination country.<sup>15</sup> For those Spanish emigrants, the level of education that they have when the Census was collected in their destination countries does not correspond to the level of education at the time of migration, possibly biasing our results. Although this should, in part, be mitigated by the incorporation of individuals older than 25 years (they are more likely to have finished schooling), we could have used information on the age at migration of the Spanish emigrants to deal with this potential problem, but the data are not available. The use of the information on the year of migration is a little problematic, as we explained earlier. Another way to address this problem is to redefine the education variables (Grogger and Hanson 2011). All the analyses

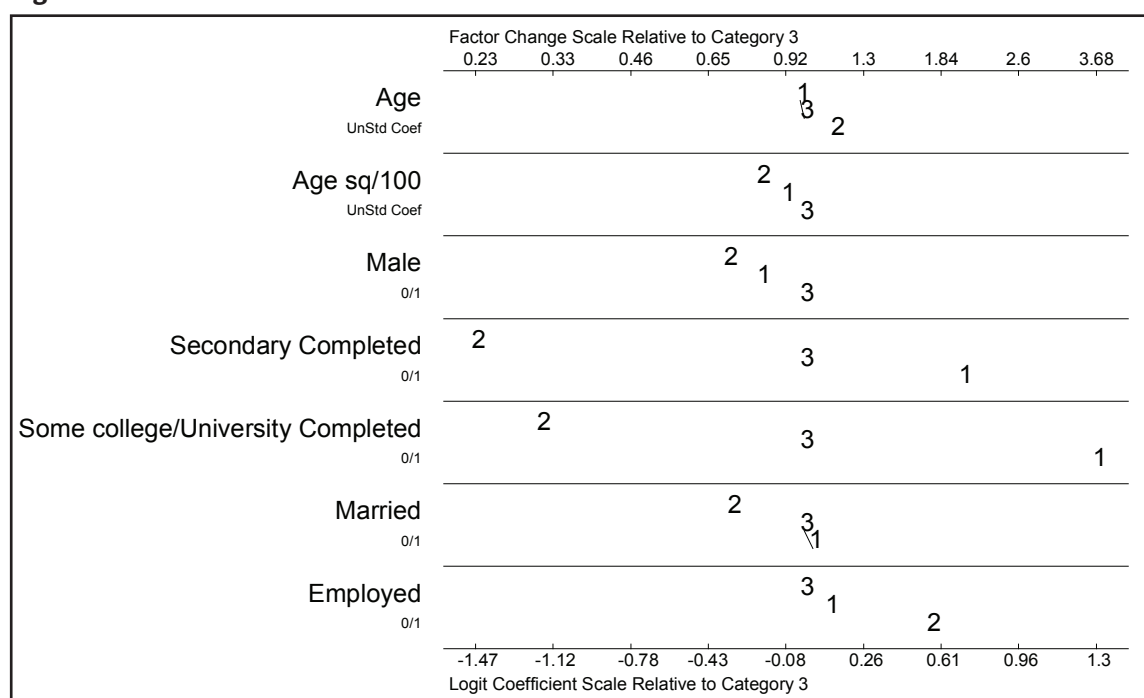
<sup>14</sup> This analysis also allows us to mitigate the concerns that can be generated by the lack of information on the number of emigration experiences. If older individuals are more likely to experience multiple emigrations, then by using young Spanish emigrants aged 25–39 years, the possibility of multiple emigrations is reduced.

<sup>15</sup> Using information from PERE 2009, we see that <3% of the Spanish emigrants are younger than 16 years of age. This represents a quite small part of the Spanish emigrant population.

have been rerun by incorporating an education variable that takes the value “1” if the Spanish emigrant reports a secondary level of completed education (or at least some college education) and the value “0” otherwise. This minimizes the possible bias if tertiary-educated emigrants obtain at least a secondary education in their country of origin. Results do not vary. This is not surprising, since previous estimates show that individuals with at least some college education and those with only secondary level of completed education are more likely to live in English-speaking countries. Next, they prefer to settle in Spanish-speaking countries, and those countries with other languages are the least-likely choice as a place of residence.

Empirical evidence is also presented in an odds-ratio plot, which can be easier to interpret. Figure 1 shows the estimates for the main specification, Column 1 of Table 2. In the odds-ratio plot, the independent variables are represented in separate rows. The horizontal axis indicates the relative magnitude of the coefficients associated with each outcome category. The numbers correspond to the outcome categories: “1” denotes English as official language, “2” indicates other languages, and “3” corresponds to Spanish, which is the base category in that figure. The distance between a given pair of outcome categories indicates the magnitude of the effect, and statistical significance is shown by drawing a line between categories for which there is no statistically significant coefficient at the 5% level of significance. Results suggest that the education level of emigrants is important in the choice of destination country. In this figure,

**Figure 1**



*Notes:* Dependent variable: languages (English, Spanish, and Others). Standard errors are robust. With respect to the sample, see the notes in Table 2. Countries are classified using the official or the most common language. The numbers correspond to the outcome categories: “1” indicates English, “2” indicates Others, and “3” indicates Spanish. The additive scale on the bottom axis measures the value of  $\beta_{i,m|n} \delta$ . The multiplicative scale on the top axis measures  $\exp(\beta_{i,m|n}) \delta$ . The statistical significance is shown by drawing a line between categories for which there is no significant coefficient at the 5% level. UnStd Coef means Unstandardized Coefficient and 0/1 means a dummy variable.

it is easy to observe the similar behavior of those with secondary and tertiary levels of education, relative to those with a primary level of education (or less than primary), although those with at least some college education appear to be more likely to be living in English-speaking countries. Thus, it appears that the location choice considering the language of the destination country varies between the primary level of education and the secondary (and higher) level of education.

#### 4.2 Destination countries classified by continent

Table 3 reports the results considering the continent of the destination countries as the dependent variable. Those Spanish emigrants with at least some college education are less likely to settle in Europe (the factor change is <1; Column 1). There are no statistically significant differences between America and continents other than Europe. Similarly, those reporting having completed secondary education are also less likely to settle in Europe than in the Americas or elsewhere. This result again points to two different responses by level of education. The low-education-level individuals (with primary school or less) are less likely to live in distant countries, whereas physical distance appears to be less important for those with secondary or tertiary level of education because they prefer to live in the Americas or elsewhere. Results do not change substantially when the year dummies are included in Column 2, although there are no statistically significant differences between Europe and other continents (except for the Americas) when individuals report having at least some

**Table 3** Location choice: destination countries classified by continent (dependent variable: Europe, America, and Others)

Variables	1	2	3
<b>Secondary education completed</b>			
Europe–America	0.168***	0.413***	0.413***
Europe–Others	0.139***	0.333**	0.333**
America–Others	0.828	0.806	0.806
<b>Some college/university completed</b>			
Europe–America	0.151***	0.252***	0.252***
Europe–Others	0.288**	0.457	0.457
America–Others	1.908	1.810	1.811
Year dummies	No	Yes	Yes
McFadden's adjusted $R^2$	0.134	0.823	0.834
Observations	54,317	54,317	54,518

*Notes:* The dependent variable is the location of the destination countries. This location is classified into the following categories: Europe, America, and Others (Philippines and South Africa). This table shows the factor change in the odds ratios,  $\exp(\beta_{i,m|n})\delta$ , with  $\delta = 1$ . With respect to the sample, see notes in Table 2. Columns 1–3 include the following as controls: age, age square/100, gender (male), being married, and being employed. Column 3 includes UK 1991. Estimates are weighted. Standard errors are robust.

\*Significant at the 10% level.

\*\*Significant at the 5% level.

\*\*\*Significant at the 1% level.

college education. The same result is obtained when Spanish emigrants living in the UK are incorporated in the sample (Column 3).<sup>16</sup>

### 4.3 Destination countries classified by unemployment level

Up to this point, we have omitted the analysis of the economic situation of the destination countries as a relevant characteristic of the destination country. There are several ways to measure this economic situation, and in this study, we use the UR as a common proxy that reflects the labor market situation. Unfortunately, we cannot introduce the unemployment level in the year of migration of Spanish emigrants. As explained above, the information on the year of migration is not available for all the destination countries considered. Though recognizing that this is not the best measure of the economic situation, we measure this variable in the year in which the Census was available.

In Table 4, we observe the main results of this analysis. For those reporting secondary education completed, we find that the odds of living in a country with a low UR, compared to those with medium or high URs, are around 3.8 and 1.5 times greater, respectively, holding

**Table 4** Location choice: destination countries classified by economic situation (dependent variables: high UR [HU], medium UR [MU], and low UR [LU])

Variables	1	2	3
<b>Secondary education completed</b>			
LU–MU	3.783***	1.014	1.014
LU–HU	1.505***	1.571***	1.571***
MU–HU	0.3978***	1.549***	1.549***
<b>Some college/university completed</b>			
LU–MU	2.898***	0.712**	0.712**
LU–HU	2.347***	2.622***	2.622***
MU–HU	0.810***	3.683***	3.683***
Year dummies	No	Yes	Yes
McFadden's adjusted $R^2$	0.078	0.373	0.389
Observations	54,317	54,317	54,518

*Notes:* The dependent variable is the UR of the destination country (source: World Bank). This UR is classified into the following categories: low unemployment (LU: UR <4%), medium unemployment (MU: UR between 4% and 9.9%), and high unemployment (HU, UR ≥10%). This table shows the factor change in the odds ratios,  $\exp(\beta_{i,m|n})\delta$ , with  $\delta = 1$ . With respect to the sample, see notes in Table 2. Columns 1–3 include the following as controls: age, age square/100, gender (male), being married, and being employed. Column 3 includes UK 1991. Estimates are weighted. Standard errors are robust.

UR = unemployment rate.

\*Significant at the 10% level.

\*\*Significant at the 5% level.

\*\*\*Significant at the 1% level.

<sup>16</sup> Our findings are unchanged when we consider different subsamples, as in Subsection 4.1. The same occurs with the remainder of the analysis.

the remaining variables constant (Column 1). This is not so clear for those with at least some college education, although without year dummies, the behavior is similar; after adding these controls (Column 2), we detect that those with at least some college education are more likely to settle in countries with medium URs than with low URs, and that the high-unemployment countries are the least-likely option. The findings are the same when the sample of Spanish emigrants living in the UK is added to the main sample in Column 3. Consequently, although, for high-educated individuals, the less-likely countries are those with high URs, mixed results appear when comparing those with URs <10% (medium and low URs). These mixed results may be due to the fact that Spanish emigrants consider not only the economic situation of a country when they decide to migrate but also other characteristics. We revisit this issue later.

#### 4.4 Destination countries classified using a combination of their characteristics

When emigrants decide where to move, they may consider more than one characteristic of the potential host countries. As the empirical evidence indicates, all the characteristics considered in this paper appear to play a role. Therefore, the analysis has been repeated using different combinations of those characteristics. First, we consider the combination of the physical distance and the language.<sup>17</sup> The sample is limited to those living in Europe and America.<sup>18</sup> The official or the most common languages are classified as before (English, Spanish, and Others). Figure 2 shows the main findings in an odds-ratio plot. In all cases, it is observed that the education level of emigrants is significant in choosing a destination country. Results suggest the following sequence of choices for Spanish emigrants with secondary education completed: Europe (English), the Americas (English), the Americas (Spanish), the Americas (Others), and Europe (Others). For those with some college education, the results suggest that the preferences of location are as follows: the Americas (English), Europe (English), the Americas (Others), the Americas (Spanish), and Europe (Others).<sup>19</sup> It is observed that more-educated Spanish emigrants are more likely to live in English-speaking areas.<sup>20</sup> Thus, with the combination of these characteristics, we do not detect substantial differences with the individual analysis of the languages of the destination countries. However, these findings appear to reveal that language is a more important characteristic than physical distance, since, regardless of the continent in which the destination countries are located, the sequence established by the language prevailed over that for the physical distance.

We also combine the economic situation and the location of the destination countries.<sup>21</sup> The results are presented in Table A3. Irrespective of the sample used and the controls included, it can be concluded that countries in the Americas with low URs are the most-likely choices when individuals have at least some college education. This is also seen in the case of those with secondary education completed, although the estimates are not statistically significant

17 Again, we want to clarify here that we are not imposing any order in the location choice process.

18 These are the main destinations of Spanish emigrants. Individuals born in Spain and living in those areas of the world represent 96% of the total Spanish emigrants (PERE 2009). Note that adding multiple categories to the dependent variable can be problematic for the convergence of the estimates. The number of observations in this case is 54,275.

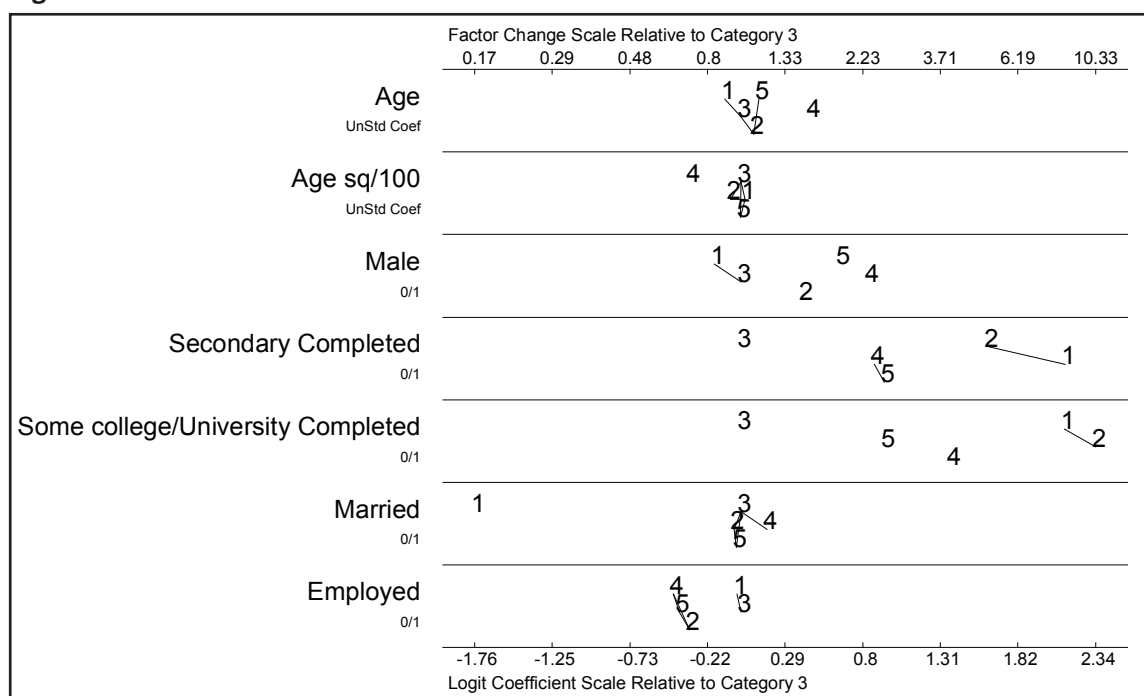
19 We recognize that among some of those outcome categories, there are no statistically significant differences. In any case, we have described the more relevant findings here.

20 The estimates are displayed in Table A2 of the Appendix (with/without year dummies and with/without UK).

21 The analysis is limited to Europe and the Americas.



**Figure 2**

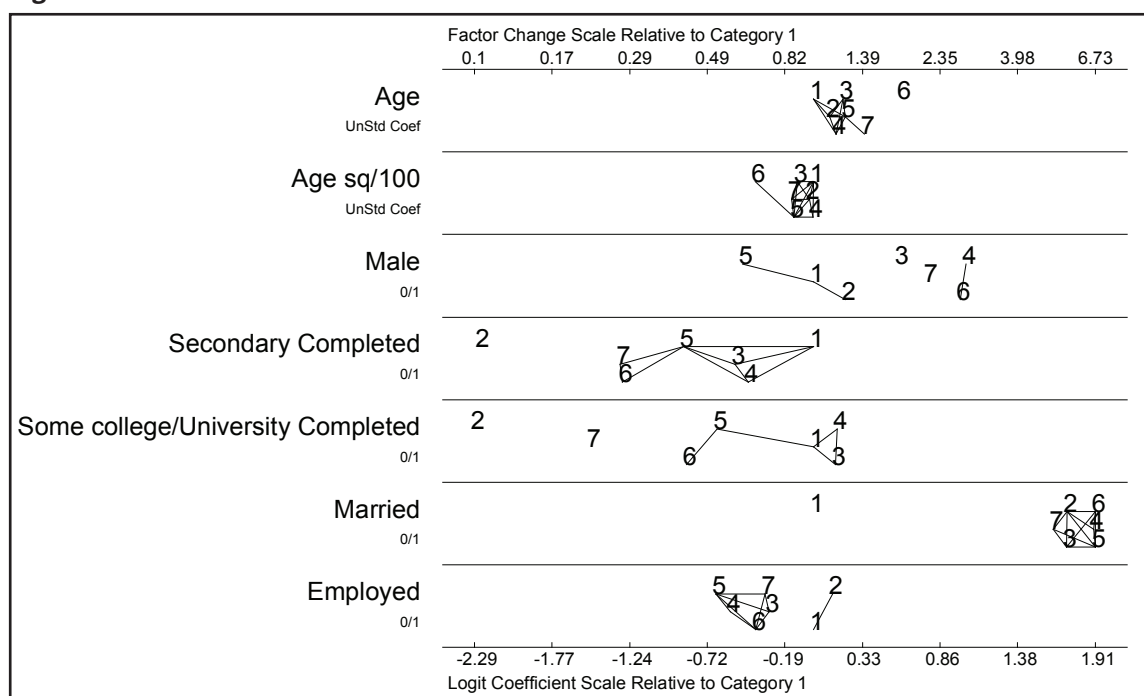


Notes: Dependent variables: languages and continent. This corresponds to Column 3 of Table A2. The numbers correspond to the following outcome categories: “1” indicates Europe (English), “2” is for America (English), “3” represents Europe (Others), “4” corresponds to America (Others), and “5” means America (Spanish). UnStd Coef means Unstandardized Coefficient and 0/1 means a dummy variable.

between the categories “American countries with low unemployment” and “European countries with high unemployment”. European countries with high URs are also more likely to be the location choice of highly educated individuals, versus American countries with equally high URs. It is striking that between European countries with low and high URs, individuals with secondary education completed and those with at least some college education are more likely to choose the latter. This appears to reflect that the shorter the distance, the less important is the better economic situation of the destination country. The combination of these characteristics suggests that the economic situation is only relevant if highly educated individuals decide to move to a distant country.

Following a similar framework, we have rerun the analysis, incorporating – in the dependent variable – combinations of the characteristics of the language (English, Spanish, and Others), the distance (Europe or the Americas), and the unemployment level of the destination countries (below-average UR, or above-average UR). We focus on the odds-ratio plot represented in Figure 3. Those individuals with at least some college education are more likely to settle in English- or Spanish-speaking American countries with low URs, rather than living in countries with high URs, located in Europe or the Americas, and having Spanish or languages other than English as the official language. The least-likely destination choice for highly educated individuals is a European country with low URs and with languages other than Spanish or English. In addition, even when including the official language, we find that the shorter the distance, the less likely are the highly educated to consider the better economic situation of the destination country, with the exception of the English-speaking European countries. However,

**Figure 3**



Notes: Dependent variables: languages, unemployment level, and continent. This corresponds to Column 3 of Table A4. The numbers correspond to the outcome categories: “1” indicates Europe, unemployment rate (UR) below the average, and English as the most common language; “2” indicates Europe, UR below the average, and Others as the most common language; “3” stands for America, UR below the average, and English as the most common language; “4” corresponds to America, UR below the average, and Spanish as the most common language; “5” means Europe, UR above the average, and Others as the most common language; “6” indicates America, UR above the average, and Others as the most common language; and “7” represents America, UR above the average, and Spanish as the most common language. UnStd Coef means Unstandardized Coefficient and 0/1 means a dummy variable.

the greater the distance between the country of origin and the destination country (American countries), the more likely are the highly educated to take into account the economic situation and so to choose low-unemployment countries. With respect to those with a secondary level of education, the behavior is quite similar, although results suggest that a greater proportion prefer to choose the English-speaking European countries – but again there are no statistically significant differences with English- and Spanish-speaking American countries having low URs. American countries with high URs are less-likely choices than those with low URs. As before, the least-likely destination for those with secondary education completed are European countries with low URs and other languages, while European countries with high URs and other languages are more-likely choices. The results are maintained with/without year dummies, with/without data from the UK (see Table A4 in the Appendix), and with/without data from Spanish emigrants living in France (see Figure A1 in the Appendix). All in all, we can conclude that there are few dissimilarities between the Spanish emigrants with secondary and tertiary levels of education. For the more-educated emigrants, English- and Spanish-speaking countries are preferred, but with low URs. The emigrants not only focus on the language of the destination country, they also consider the economic situation of that country. For highly

educated individuals, we observe that being in a better economic situation plays a role when the distance between the country of origin and the destination country is greater.

#### 4.5 Analysis over time

We extend our previous analysis to study whether our findings are maintained over time. We recognize that the results obtained in this subsection should be treated with caution, since the number of destination countries is quite limited in certain periods and, in several cases, the destination is concentrated only in the Americas, with no data from Europe. For this reason, we only analyze the relationship between the official (or the most common) language of each destination country and the education level of Spanish emigrants (including all controls). The estimates have been repeated every 10 years since 1960 to 2010. Data come from the Integrated Public Use Microdata Series–International of the Censuses (Minnesota Population Center 2014).

Results are reported in Table 5. The estimates indicate that English-speaking countries are much more likely to be the place of residence of Spanish emigrants than countries with Spanish as the official language, in all decades considered, for those with secondary and tertiary levels of education. Results on the relationship between Spanish-speaking countries and countries with languages other than English are mixed. Note that the results for countries with Spanish and Other languages vary dramatically from one decade to another, so the results cannot be conclusive with those samples. In any case, being aware of the problems that the sample of destination countries can generate, it is comforting to note that the pattern of behavior of Spanish emigrants with both secondary and tertiary levels of education is quite similar over time.

### 5 Conclusion

This paper examines the relationship between the education level of Spanish emigrants and their destination country. Since Spanish emigrants were born under the same laws, economic conditions, and institutions, the differences in their location choice can be due to dissimilarities in their level of education. To study this issue, our work uses microdata from the Censuses of 21 countries, for the period from 2000 to 2007. Destination countries are classified using three characteristics: the official (or the most common) language, the physical distance between the country of origin and the destination, and the level of unemployment. In this way, we take into consideration the migration decisions of individuals based on language, migration costs, and job search.

Our results suggest that Spanish emigrants with secondary and tertiary levels of education have greater preferences for English-speaking countries, more-distant countries, and countries with low URs, when each of these factors is analyzed separately. Our findings are quite robust to the inclusion of several sociodemographic characteristics and to the use of different subsamples.

By combining these characteristics, since Spanish emigrants may consider more than one characteristic of the potential host country when deciding the location, we draw interesting conclusions. There are still similarities in the location choice between those with secondary

**Table 5** Analysis over time: decades 1960–2000 (dependent variable: languages – English, Spanish, and Others)

<b>Decade 1960</b>	<b>1</b>	<b>2</b>	<b>Decade 1970</b>	<b>3</b>	<b>4</b>
<b>Secondary education completed</b>					
English–Others	16.073***	17.345***		10.797***	10.744***
English–Spanish	4.734***	3.257**		9.623***	10.847***
Others–Spanish	0.295***	0.188***		0.891	1.010
<b>Some college/university education completed</b>					
English–Others	14.586***	15.379***		10.240***	10.180***
English–Spanish	4.752***	1.626		10.505***	10.406***
Others–Spanish	0.326***	0.106***		1.026	1.022
Year dummies	No	Yes		No	Yes
McFadden's adjusted $R^2$	0.078	0.546		0.052	0.178
Observations	8,803	8,803		24,226	24,226
<b>Decade 1980</b>			<b>Decade 1990</b>		
<b>Secondary education completed</b>					
English–Others	3.701***	3.879***		1.071	1.056
English–Spanish	5.469***	6.050***		1.473***	1.533***
Others–Spanish	1.478***	1.560***		1.376***	1.451***
<b>Some college/university education completed</b>					
English–Others	4.175***	4.411***		1.069	1.062
English–Spanish	7.312***	7.992***		1.598***	1.816***
Others–Spanish	1.752***	1.812***		1.495***	1.710***
Year dummies	No	Yes		No	Yes
McFadden's adjusted $R^2$	0.053	0.084		0.078	0.094
Observations	27,499	27,499		17,368	17,368
<b>Decade 2000</b>					
<b>Secondary education completed</b>					
English–Others	4.130***	4.130***			
English–Spanish	3.879***	3.879***			
Others–Spanish	0.939	0.939			
<b>Some college/university education completed</b>					
English–Others	4.599***	4.599***			
English–Spanish	8.628***	8.628***			
Others–Spanish	1.876**	1.876**			
Year dummies	No	Yes			
McFadden's adjusted $R^2$	0.120	0.120			
Observations	2,718	2,718			

Notes: *Decade 1960*: Sample consists of emigrants in Brazil 1960, Colombia 1964, Costa Rica 1963, Ecuador 1962, Mexico 1960, United States 1960, and Uruguay 1963. *Decade 1970*: Argentina 1970, Bolivia 1976 Brazil 1970, Chile 1970, Colombia 1973, Costa Rica 1973, Mexico 1970, Romania 1977, United States 1970, Uruguay 1975, and Venezuela 1971. *Decade 1980*: Argentina 1980, Brazil 1980, Chile 1982, Costa Rica 1984, Ireland 1986, Panama 1980, Portugal 1981, United States 1980, Uruguay 1985, and Venezuela 1981. *Decade 1990*: Argentina 1991, Brazil 1991, Chile 1992, Colombia 1993, Ireland 1996, Mexico 1990, Nicaragua 1995, Peru 1993, Portugal 1991, Romania 1992, United Kingdom 1991, United States 1990, and Venezuela 1990. *Decade 2000*: Argentina 2010, Brazil 2010, Ecuador 2010, Mexico 2010, Panama 2010, United States 2010. The dependent variable is the official (or the most common) language of the destination countries. Languages are classified into the following categories: English, Spanish, and Others. This table shows the odds ratios,  $\exp(\beta_{i,m|n})\delta$ , with  $\delta = 1$ . All columns include the following as controls: age, age square/100, gender (male), being married, and being employed. Estimates are weighted. Standard errors are robust.

\*Significant at the 10% level.

\*\*Significant at the 5% level.

\*\*\*Significant at the 1% level.

and tertiary levels of education. English-speaking countries, regardless of the location (Ireland, the US, and the UK), and Spanish-speaking countries with low URs are the most-likely residence choices for more educated emigrants. The highly educated are more likely to choose to live in distant countries (with the exception of the UK and Ireland), but if it is a Spanish-speaking country, the economic situation is a significant factor. However, when those individuals decide to move to a European country, they appear to pay less attention to the economic situation, since the least-likely option for them is a European country with a low UR (excluding the UK and Ireland).

To sum up, our results reveal that the level of education plays a role in the location choice of those who decide to migrate, but that the differences appear to be more significant between those with a primary level of education and those with a secondary or higher level of education. Our findings denote the importance of considering several of the characteristics of the destination country, to determine how potential emigrants choose a country to live in. The conclusions provided by this work can also be useful for potential destination countries, in order to implement policies to attract specific groups of emigrants.

## References

- Adsera, A.; M. Pytlikova (2015): The Role of Language in Shaping International Migration. *The Economic Journal* 125, F49-F81.
- Anderson, J. (1984): Regression and Ordered Categorical Variables. *Journal of the Royal Statistical Society Series B* 46, 1-30.
- Barro, R.; J. Lee (2001): International Data on Educational Attainment: Updates and Implications. *Oxford Economic Papers* 53, 541-563.
- Bauer, T.; G. Epstein; I. Gang (2005): Enclaves, Language, and the Location Choice of Migrants. *Journal of Population Economics* 18, 649-662.
- Beine, M.; F. Docquier; H. Rapoport (2008): Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers. *Economic Journal* 118, 631-652.
- Beine, M., F. Docquier; Ç. Özden (2011): Diaspora. *Journal of Development Economics* 95, 30-41.
- Belot, M.; S. Ederveen (2012): Cultural Barriers in Migration Between OECD Countries. *Journal of Population Economics* 25, 1077-1105.
- Belot, M.; T. Hatton (2012): Immigrant Selection in the OECD. *The Scandinavian Journal of Economics* 114, 1105-1128.
- Bover, O.; P. Velilla (2005): Migration in Spain: Historical background and current trends European Migration: What Do We Know? Oxford University Press: Oxford/New York.
- Clark, X.; T. J. Hatton; J. G. Williamson (2007): Explaining U.S. Immigration, 1971-1998. *Review of Economics and Statistics* 89, 359-373.
- Clemente, J.; G. Larramona; L. Olmos (2016): Interregional Migration and Thresholds: Evidence from Spain. *Spatial Economic Analysis* 11(3), 1-18.
- Chiswick, B. R.; P. W. Miller (2014): International Migration and the Economics of Language. *Handbook of the Economics of International Migration*, 1A: The Immigrants, 211-269.
- Docquier, F.; O. Lohest; A. Marfouk (2007): Brain Drain in Developing Countries. *World Bank Economic Review* 21, 193-218.
- Dotti, N.; U. Fratesi; C. Lenzi; M. Percoco (2013): Local Labour Markets and the Interregional Mobility of Italian University Students. *Spatial Economic Analysis* 8, 443-468.
- Eliasson, K., U. Lindgren; O. Westerlund (2003): Geographical Labour Mobility: Migration or Commuting? *Regional Studies* 8, 827-837.
- Gonzalez, L.; F. Ortega (2011): How do Very Open Economies Adjust to Large Immigration Flows? Evidence from Spanish Regions. *Labour Economics* 18, 57-70.
- Greene, W. H. (2008): *Econometric Analysis*. Upper Saddle River, New Jersey: Granite Hill Publishers.
- Grogger, J.; G. H. Hanson (2011): Income Maximization and the Selection and Sorting of International Migrants. *Journal of Development Economics* 95, 42-57.
- Hausman, J.; D. McFadden (1984): Specification Tests for the Multinomial Logit Model. *Econometrica* 52, 1219-1240.

- Herzog Jr., H. W.; A. M. Schlottmann; T. P. Boehm** (1993): Migration as Spatial Job–Search: A Survey of Empirical Findings. *Regional Studies* 27, 327-340.
- Huber, P., K. Nowotny** (2013): Moving Across Borders: Who Is Willing to Migrate or to Commute? *Regional Studies* 47, 1462-1481.
- Juarez, J. P.** (2000): Analysis of Interregional Labor Migration in Spain Using Gross Flows. *Journal of Regional Science* 40, 377-399.
- Long, J.; J. Freese** (2006): *Regression Models for Categorical Dependent Variables Using Stata*, 2nd ed. Stata Press Publication.
- Malamud, O.; A. Wozniak** (2012): The Impact of College on Migration Evidence from the Vietnam Generation. *Journal of Human Resources* 47, 913-950.
- Marcén, M.; J. A. Molina** (2012): Informal Caring-Time and Caregiver Satisfaction. *The European Journal of Health Economics* 13, 683-705.
- Mayda, A. M.** (2010): International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows. *Journal of Population Economics* 23, 1249-1274.
- Mckelvey, R. D.; W. Zavoina** (1975): A Statistical Model for the Analysis of Ordinal Level Dependent Variables. *Journal of Mathematical Sociology* 4, 103-120.
- Minnesota Population Center** (2014): *Integrated Public Use Microdata Series. International: Version 6.3* [Machine-Readable Database]. Minneapolis: University of Minnesota.
- Nerlove, M.; S. Press** (1973): *Univariate and Multivariate Log-Linear and Logistic Models*, 1306, Santa Monica, California: Rand.
- Neudörfer, P.; J. Dresdner** (2014): Does Religious Affiliation Affect Migration? *Papers in Regional Science* 93, 577-594.
- Nifo, A.; G. Vecchione** (2014): Do Institutions Play a Role in Skilled Migration? The Case of Italy. *Regional Studies* 48, 1628-1649.
- Ortega, F.; P. Giovanni** (2009): *The Causes and Effects of International Migrations: Evidence from OECD Countries 1980-2005*, Working Paper No. 09.6, Department of Economics, University of California, USA.
- Sjaastad, L.** (1962): The Costs and Returns of Human Migration. *Journal of Political Economy* 70, 80-93.
- Small, K.; C. Hsiao** (1985): Multinomial Logit Specification Tests. *International Economic Review* 26, 619-627.
- Venhorst, V.; J. Van Dijk; L. Van Wissen** (2011): An Analysis of Trends in Spatial Mobility of Dutch Graduates. *Spatial Economic Analysis* 6, 57-82.
- Wang, Z.; T. Graaff; P. Nijkamp** (2016): Cultural Diversity and Cultural Distance as Choice Determinants of Migration Destination. *Spatial Economic Analysis* 11, 176-200.

## Appendix

**Table A1** Robustness check: considering different subsamples varying the age of Spanish emigrants (dependent variable: English, Spanish, and Others)

Variables	1	2	3	4	5	6	7	8	9	10
<b>Secondary education completed</b>										
English–Others	8.797***	2.328**	9.407***	2.704***	9.028***	2.2033***	9.930***	2.089***	1.444***	3.174***
English–Spanish	2.027***	1.396*	2.352***	1.587***	1.747***	1.204	1.955***	1.262	0.732***	1.244
Others–Spanish	0.230***	0.600***	0.250***	0.587***	0.194***	0.5463***	0.197***	0.604***	0.507***	0.392***
<b>Some college/university education completed</b>										
English–Others	11.937***	2.752***	12.257***	3.268***	12.73***	2.4242***	14.727***	2.276***	8.223***	4.229***
English–Spanish	3.680***	1.548**	4.148***	1.630***	2.940***	1.4963***	3.404***	1.705***	1.509**	1.568**
Others–Spanish	0.308***	0.563***	0.338***	0.499***	0.231***	0.6172***	0.231***	0.749**	0.184***	0.371***
Year dummies	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
McFadden's adjusted $R^2$	0.161	0.713	0.152	0.713	0.161	0.698	0.160	0.696	0.084	0.674
Observations	54,317	54,317	65,045	65,045	30,743	30,743	27,235	27,235	12,259	12,259

Notes: With respect to the sample, see notes in Table 2. The dependent variable is the official language or the most common language of the destination countries. Languages are classified into the following categories: English, Spanish, and Others. This table shows the factor change in the odds ratios,  $\exp(\beta_{i(im)})\delta$ , with  $\delta = 1$ . Columns 1–10 include the following as controls: age, age square/100, gender (male), being married, and being employed. Columns 3 and 4 include individuals aged 25–64 years. Columns 5 and 6 include individuals aged 25–49 years. Columns 7 and 8 include individuals aged 30–49 years. Columns 9 and 10 include individuals aged 25–39 years. Standard errors are robust.

\*Significant at the 10% level.

\*\*Significant at the 5% level.

\*\*\*Significant at the 1% level.

**Table A2** Location choice: destination countries classified by languages and continents

Variables	1	2	3
<b>Secondary education completed</b>			
Europe (English)–America (English)	1.491	1.660*	1.660*
Europe (English)–Europe (Others)	15.209***	8.468***	8.468***
Europe (English)–America (Others)	3.076***	3.516***	3.516***
Europe (English)–America (Spanish)	2.854***	3.297***	3.297***
America (English)–Europe (Others)	10.199***	5.101***	5.101***
America (English)–America (Others)	2.063***	2.118***	2.118***
America (English) –America (Spanish)	1.914***	1.986***	1.986***
Europe (Others)–America (Others)	0.202***	0.415***	0.415***
Europe (Others)–America (Spanish)	0.188***	0.389***	0.389***
America (Others)–America (Spanish)	0.928	0.938*	0.938
<b>Some college/university education completed</b>			
Europe (English)–America (English)	0.700	0.815	0.815
Europe (English)–Europe (Others)	10.150***	8.417***	8.417***
Europe (English)–America (Others)	1.800**	2.118**	2.118**
Europe (English)–America (Spanish)	2.521***	3.265***	3.265***
America (English)–Europe (Others)	14.503***	10.331***	10.332***
America (English)–America (Others)	2.572***	2.600***	2.600***
America (English)–America (Spanish)	3.602***	4.007***	4.008***
Europe (Others)–America (Others)	0.177***	0.252***	0.252***
Europe (Others)–America (Spanish)	0.248***	0.388***	0.388***
America (Others)–America (Spanish)	1.401***	1.541***	1.541***
Year dummies	No	Yes	Yes
McFadden's adjusted $R^2$	0.161	0.570	0.632
Observations	54,275	54,275	54,476

Notes: With respect to the sample, see notes in Table 2. The dependent variable is a combination of the location of the destination countries and the official or the most common language. This location is classified into the following categories: Europe and America. The language classification is as follows: English, Spanish, and Others. The languages of the destination countries are in parentheses. This table shows the factor change in the odds ratios,  $\exp(\beta_{i,m|n})\delta$ , with  $\delta = 1$ . Columns 1–3 include the following as controls: age, age square/100, gender (male), being married, and being employed. Column 3 includes UK 1991. Estimates are weighted. Standard errors are robust.

\*Significant at the 10% level.

\*\*Significant at the 5% level.

\*\*\*Significant at the 1% level.



**Table A3** Location choice: destination countries classified by URs and continents

Variables	1	2	3
<b>Secondary education completed</b>			
Europe/LU–America/LU	0.099***	0.255***	0.255***
Europe/LU–Europe/HU	0.148***	0.363***	0.363***
Europe/LU–America/HU	0.209***	0.548***	0.548***
America/LU–Europe/HU	1.498	1.424	1.424
America/LU–America/HU	2.109***	2.148***	2.148***
Europe/HU–America/HU	1.407	1.508	1.508
<b>Some college/university education completed</b>			
Europe/LU–America/LU	0.069***	0.133***	0.133***
Europe/LU–Europe/HU	0.161***	0.292***	0.292***
Europe/LU–America/HU	0.307***	0.590***	0.590***
America/LU–Europe/HU	2.344**	2.189**	2.189**
America/LU–America/HU	4.479***	4.425***	4.425***
Europe/HU–America/HU	1.911*	2.022**	2.022**
Year dummies	No	Yes	Yes
McFadden's adjusted $R^2$	0.178	0.630	0.646
Observations	54,275	54,275	54,476

Notes: With respect to the sample, see notes in Table 2. The dependent variable is a combination of the location of the destination country and the UR. This location is classified into the following categories: Europe and America. The UR is divided into two categories: above-average UR (HU); and below-average UR (LU). This table shows the factor change in the odds ratios,  $\exp(\beta_{i,m|n})\delta$ , with  $\delta = 1$ . Columns 1–3 include the following as controls: age, age square/100, gender (male), being married, and being employed. Column 3 includes UK 1991. Estimates are weighted. Standard errors are robust.

HU = high UR; LU = low UR; UR = unemployment rate.

\*Significant at the 10% level.

\*\*Significant at the 5% level.

\*\*\*Significant at the 1% level.

**Table A4** Location choice: destination countries classified by languages, URs, and continents

Variables	1	2	3
<b>Secondary education completed</b>			
EU/LU/English–EU/LU/Others	15.571***	9.662***	9.595***
EU/LU/English–AM/LU/English	1.510	1.712*	1.701*
EU/LU/English–AM/LU/Spanish	1.364	1.667	1.555
EU/LU/English–EU/HU/Others	2.203*	2.431*	2.411*
EU/LU/English–AM/HU/Others	3.130***	3.689***	3.647***
EU/LU/English–AM/HU/Spanish	3.189***	3.741***	3.700***
EU/LU/Others–AM/LU/English	0.097***	0.177***	0.177***
EU/LU/Others–AM/LU/Spanish	0.088***	0.173***	0.162***
EU/LU/Others–EU/HU/Others	0.142***	0.252***	0.251***
EU/LU/Others–AM/HU/Others	0.201***	0.382***	0.380***

(continued)

**Table A4** Continued

<b>Variables</b>	<b>1</b>	<b>2</b>	<b>3</b>
EU/LU/Others–AM/HU/Spanish	0.205***	0.387***	0.386***
AM/LU/English–AM/LU/Spanish	0.903	0.974	0.914
AM/LU/English–EU/HU/Others	1.459	1.419	1.418
AM/LU/English–AM/HU/Others	2.073***	2.155***	2.144***
AM/LU/English–AM/HU/Spanish	2.112***	2.185***	2.175***
AM/LU/Spanish–EU/HU/Others	1.615	1.458	1.550
AM/LU/Spanish–AM/HU/Others	2.295***	2.213***	2.345***
AM/LU/Spanish–AM/HU/Spanish	2.338***	2.244***	2.379***
EU/HU/Others–AM/HU/Others	1.421	1.518	1.513
EU/HU/Others–AM/HU/Spanish	1.447	1.539	1.535
AM/HU/Others–AM/HU/Spanish	1.019	1.014	1.015
<b>Some college/university education completed</b>			
EU/LU/English–EU/LU/Others	10.706***	9.888***	9.912***
EU/LU/English–AM/LU/English	0.723	0.854	0.861
EU/LU/English–AM/LU/Spanish	0.712	0.967	0.853
EU/LU/English–EU/HU/Others	1.677	1.911	1.923
EU/LU/English–AM/HU/Others	1.926**	2.325**	2.361**
EU/LU/English–AM/HU/Spanish	3.700***	4.460***	4.522***
EU/LU/Others–AM/LU/English	0.068***	0.086***	0.087***
EU/LU/Others–AM/LU/Spanish	0.067***	0.098***	0.086***
EU/LU/Others–EU/HU/Others	0.157***	0.193***	0.194***
EU/LU/Others–AM/HU/Others	0.180***	0.235***	0.238***
EU/LU/Others–AM/HU/Spanish	0.346***	0.451***	0.456***
AM/LU/English–AM/LU/Spanish	0.985	1.131	0.991
AM/LU/English–EU/HU/Others	2.320**	2.236**	2.235**
AM/LU/English–AM/HU/Others	2.665***	2.721***	2.744***
AM/LU/English–AM/HU/Spanish	5.119***	5.219***	5.254***
AM/LU/Spanish–EU/HU/Others	2.357**	1.976*	2.256**
AM/LU/Spanish–AM/HU/Others	2.707***	2.405***	2.770***
AM/LU/Spanish–AM/HU/Spanish	5.199***	4.613***	5.304***
EU/HU/Others–AM/HU/Others	1.148	1.217	1.228
EU/HU/Others–AM/HU/Spanish	2.206**	2.334**	2.351**
AM/HU/Others–AM/HU/Spanish	1.921***	1.918***	1.915***
Year dummies	No	Yes	Yes
McFadden's adjusted $R^2$	0.159	0.529	0.581
Observations	54,275	54,275	54,476

*Notes:* With respect to the sample, see notes in Table 2. The dependent variable is a combination of the location of the destination country, the official (or the most common) language, and the UR. This location is classified into the following categories: Europe (EU) and America (AM). The language classification is as follows: English, Spanish, and Others. The UR is divided into two categories: above-average UR (HU) and below-average UR (LU). This table shows the factor change in the odds ratios,  $\exp(\beta_{i,m|n})\delta$ , with  $\delta=1$ . Columns 1–3 include the following as controls: age, age square/100, gender (male), being married, and being employed. Column 3 includes UK 1991. Standard errors are robust.

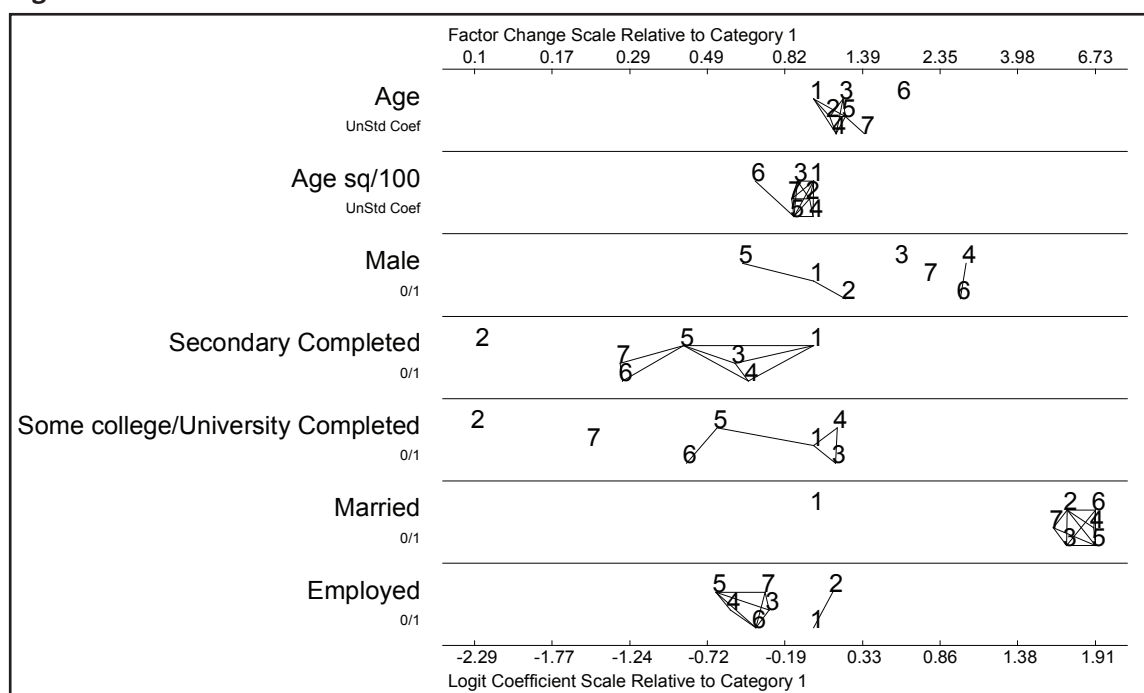
HU = high UR; LU = low UR; UR = unemployment rate.

\*Significant at the 10% level.

\*\*Significant at the 5% level.

\*\*\*Significant at the 1% level.

**Figure A1**



Notes: Dependent variables: languages, unemployment level, and continent. This corresponds to Column 3 of Table A4. The sample consists of Spanish emigrants in Argentina 2001, Austria 2001, Bolivia 2001, Brazil 2000, Chile 2002, Colombia 2005, Costa Rica 2000, Cuba 2002, El Salvador 2007, Greece 2001, Ireland 2002, Mexico 2000, Nicaragua 2005, Panama 2000, Peru 2007, Portugal 2001, United States 2000, UK 1991, and Venezuela 2001. The dependent variable is a combination of the location of the destination countries, the official language (or the most common language), and the UR. This location is classified into the following categories: Europe and America. The language classification is as follows: English, Spanish, and Others. The UR is divided into two categories: above-average UR and below-average UR. The numbers correspond to the following outcome categories: “1” indicates Europe, UR below the average, and English as the most common language; “2” indicates Europe, UR below the average, and Others as the most common language; “3” indicates America, UR below the average, and English as the most common language; “4” corresponds to America, UR below the average, and Spanish as the most common language; “5” means Europe, UR above the average, and Others as the most common language; “6” indicates America, UR above the average, and Others as the most common language; and “7” represents America, UR above the average, and Spanish as the most common language. The additive scale on the bottom axis measures the value of  $\beta_{i,m|n} \delta$ . The multiplicative scale on the top axis measures  $\exp(\beta_{i,m|n}) \delta$ . Estimates are weighted. Standard errors are robust. The statistical significance is shown by drawing a line between the categories for which there is no significant coefficient at the 5% level. UR = unemployment rate. UnStd Coef means Unstandardized Coefficient and 0/1 means a dummy variable.