Immigrant Selection over the Business Cycle: the Spanish Boom and the Great Recession

Jesús Fernández-Huertas Moraga*
FEDEA and IAE (CSIC)

Abstract

Spain received more immigrants than any other European country during its boom between 1997 and 2007 but continued receiving them during the recession that followed. This paper documents the selection of these immigrants in terms of their productive characteristics both before and after the peak of the boom. Pre-crisis immigrants were typically positively selected although selection was less positive for some over-represented national groups. Post-crisis immigration became even more positively selected, as well as older and more feminized.

Keywords: *international migration, selection, household survey.*
JEL classification: *F22, O15, J61, D33.*

Resumen

España ha recibido más inmigrantes que ningún otro país europeo durante su periodo de expansión económica entre 1997 y 2007, pero además también ha seguido recibiendo inmigrantes durante la recesión posterior. Este trabajo estudia la selección de estos inmigrantes, en términos de sus características productivas antes y después del pico del periodo expansivo. Los inmigrantes anteriores a la crisis habitualmente tenían una selección positiva, en términos de sus características socioeconómicas, aunque dicha selección resultaba menos positiva para determinados orígenes nacionales sobre-representados. Los inmigrantes posteriores a la crisis han experimentado una selección incluso aún más positiva, ha aumentado su edad media y se ha incrementado el porcentaje de mujeres.

Palabras clave: *emigración internacional, selección, servicio doméstico.*
Clasificación JEL: *F22, O15, J61, D33.*

1. Introduction

The effect of immigration flows on receiving countries' welfare depends both on their magnitude and on their degree of selection (Biavaschi and Elsner, 2013), defined as the difference between the productive characteristics of the immigrants and those of the stayers, approximated by their average wage at origin (Fernández-Huertas Moraga, 2011). Typically, if immigrants are positively selected (more productive than stayers), their impact is calculated to be more positive than if they are negatively selected (Borjas, 1987).

In terms of magnitude, Spain received more immigrants in relative terms than any other OECD country between 1997 and 2007 and was only second to the US in absolute terms during this period[1]. Bertoli and Fernández-Huertas Moraga (2013) studied the determinants of the size of these flows but, to my knowledge, the selection of Spanish migration inflows has only been partially addressed by Cebolla Boado, González Ferrer, Cea D'Ancona, Fernández-Huertas Moraga, Finotelli, González Luna, Moya Malapeira, and Pinyol Jiménez (2013)[2].

However, it is particularly interesting to understand how selection took place in a context of massive immigration inflows during the migration boom and how this selection changed during the ensuing regression.

To this end, this paper uses data from the Spanish Labor Force Survey (Encuesta de Población Activa) gathered by the Spanish Statistics Office (Instituto Nacional de Estadística) every quarter with the main ob jective of measuring the country's unemployment rate (INE, 2014). The survey gathers information about the productive characteristics of immigrants of different origins residing in Spain and also, for most of them, the date at which they entered the country. The Spanish data are combined with educational attainment data from Barro and Lee (2013) that allows the comparison between the education level of immigrants for each origin country of Spanish immigration and the education level of stayers.

There are two main findings in this paper. First, immigrants to Spain are shown to be positively selected both before and after the start of the Great Recession in 2008. Second, the degree of selection improves as the economic conditions in Spain worsen. The first finding confirms the general results presented, for example, by Grogger and Hanson (2011), while the second is more specific to Spain although it echoes the findings in the classical paper by Borjas (1985) on the changes in the quality of the immigrant cohorts in the US over time[3].

These general patterns coexist with a variety of experiences for different countries of origin of immigrants to Spain. The paper also documents the evolution of some stylized facts about immigrants in Spain, like their average age and their gender balance. Immigrants are shown to become younger during the immigration boom and then older with the onset of the economic crisis. In terms of gender, there is a lot of variation in the share of females during the boom but there is a clear feminization of migration to Spain after the crisis.

Of course, this paper is not the first to document many of these facts. There is a huge literature both on the selection of immigrants, briefly reviewed in the next section, and on immigration to Spain. It is out of the scope of this paper to review this literature, where many of the papers use some of the immigrant characteristics described here as inputs with the objective of measuring the effect of immigration

---

[1] Own calculations from data in United Nations, Department of Economic and Social Affairs, Population Division (2013).

[2] BERTOLI, FERNÁNDEZ-HUERTAS MORAGA and ORTEGA (2011, 2013) study the selection of Ecuatorian migration flows to Spain and compare them to Ecuadorian migration flows to the United States.

[3] See BORJAS (2013) for a more recent contribution.

on different Spanish dimensions: employment and wages (Carrasco, Jimeno and Ortega, 2008; González and Ortega, 2011; Farré, Gonzáalez and Ortega, 2011; Simón, Sanromá, and Ramos, 2008), task specialization (Amuedo-Dorantes and De la Rica, 2011), the housing boom (González and Ortega, 2013), the school system (Farré, Ortega, and Tanaka, 2014), immigrant assimilation (Sanromá, Ramos, and Simón, 2009; Izquierdo, Lacuesta and Vegas, 2009; De la Rica and Ortega, 2012; Rodríguez-Planas, 2012), the inflation rate (Bentolila, Dolado and Jimeno, 2008), election outcomes (Méndez and Cutillas, 2014), etc.

The paper proceeds as follows. Section 2 reviews the basic selection theory together with a very brief description of the relevant literature. Section 3 introduces the data, how sources are matched and some interesting descriptive statistics on the evolution of immigration in Spain between 1996 and 2013. Then, section 4 presents the main results on the evolution of selection, selection before and after the Great Recession according to a flow-like concept and some results on the main countries of origin. Finally, section 5 concludes the paper.

2. Theory

Most selection studies (Ambrosini and Peri, 2012; Fernández-Huertas Moraga, 2011, 2013; Chiquiar and Hanson, 2005) concentrate on the difference between wages at origin for future migrants and wages at origin for stayers as a measure of selection into migration. If the wage of the stayers is larger than the wage of the migrants, migrants are said to be negatively selected (Borjas, 1999).

In order to study immigrant selection over the business cycle in Spain, an ideal dataset would require wage data for all countries of origin of Spanish migrants over time for both stayers and migrants to Spain before they take the decision to leave their country. An example of this type of dataset is the Mexican ENET (Fernández-Huertas Moraga, 2011, 2013). However, it is very diffcult to obtain comprehensive data on wages for a large number of relevant countries of origin for migrants to Spain, let alone obtaining wages for the migrants before they leave.

An alternative strategy is to concentrate on selection on observable productive characteristics, such as education, age or gender. Following Fernández-Huertas Moraga (2011), positive selection in education can be defined as a situation in which:

$$E[S|M = 1] > E[S|M = 0] \qquad [1]$$

S denotes years of schooling or another measure of educational achievement, such as having a graduate degree. M = 1 refers to individuals who migrate to Spain whereas M = 0 refers to individuals who choose to remain in their origin countries. The degree of selection in education can then be defined as:

$$DS_{educ} = E[S|M = 1] - E[S|M = 0] \qquad [2]$$

This paper will present different results on the evolution of the degree of selection in terms of education. To complete the analysis, some summary statistics on the evolution of the sex ratio and the age composition of the immigrant population in Spain will also be presented.

The lack of comprehensive data on the origin countries of migration to Spain will necessarily lead to a view of immigrant selection from the point of view of the receiving country: Spain. Immigrant selection is a flow concept so that the appropriate way to measure it is by concentrating on the decision to migrate of individuals locating in the same country at the same point in time, as in Bertoli, Fernández-Huertas Moraga, and Ortega (2013). However, this paper can only approximate this concept.

The approximation is done in two different ways. First, selection will be computed on the stocks of immigrants living in Spain at different points in time. The question that will be asked is: what would the stock of immigrants living in Spain at time t look like if immigrant selection had been neutral at the time the immigrants arrived?

Secondly, an approximation to the flow concept will be attempted by looking at the characteristics of all immigrants living in Spain at a particular point in time having arrived within a particular time frame and not having left since then. The question that the second approximation will be answering is: what would recently arrived immigrants to Spain who did not leave yet look like if selection had been neutral in their origin country?

The theoretical impact of the business cycle over selection depends heavily on the structure of migration costs. Assuming no effect of the business cycle on migration costs and larger effects on Spanish wages than on wages at origin, the crisis should have led to more negatively selected migration flows if utility was concave (Borjas, 1987) and to more positively selected migration flows if utility was linear (Grogger and Hanson, 2011). With concave utility, relative income differences matter so that more skilled individuals have less incentives to migrate whenever income inequality is larger in the origin than in the destination country (as it is the case between Mexico and the US). With linear utility, absolute income differences matter so that more skilled individuals have more incentives to migrate as the absolute income differences increase with skills (as it is the case between Mexico and the US). Fernández-Huertas Moraga (2013) shows that concave utility is more likely in the case of explaining different selection patterns in rural and urban Mexico[4] and Belot and Hatton (2012) prove that positive selection patterns can be justified even in the case of concave utility due to different effects of migration costs across skill groups.

Also, if multilateral resistance to migration, defined as the confounding effect of alternative destinations on the decision to migrate by Bertoli and Fernández-Huertas Moraga (2013), is taken into account, it becomes very diffcult to derive meaningful

---

[4] The reason is that a linear utility model would predict positive selection out of urban Mexico while the data show negative selection.

predictions about the effect of the business cycle on immigrant selection. Even the effect of economic conditions on the size of migration flows becomes much harder to predict. In a two-country world, with one origin and one destination, an economic crisis with no redistributive consequences at the destination clearly reduces the incentives to migrate. However, if there are multiple destinations, everything depends on whether the crisis hits only one or many alternative destinations. It could even be the case that the economic crisis ends up creating more immigration for an individual destination country if alternative destinations for the immigrants are hit even worse. For example, if Spain and Italy are the main destinations for Romanian immigrants and Spain suffers more than Italy from the economic crisis, Romanian immigration to Italy might even increase. This is particularly relevant for the case of Spain since Bertoli and Fernández-Huertas Moraga (2013) have shown that multilateral resistance to migration had substantial effects on migration flows to Spain during the immigration boom years.

3. Data

The main source of the analysis will be the Spanish Labor Force Survey, henceforth EPA from the Spanish acronym for Encuesta de Población Activa (INE, 2014). The EPA provides quarterly information on the place of birth, year of arrival to Spain[5], education level and other basic demographics for Spanish immigrants living in Spain. In order to compute the degree of selection, the EPA is combined with data on educational achievement from Barro and Lee (2013).

Barro and Lee (2013) define different measures of educational attainment. Out of these, this paper will focus only on two: the number of years of schooling and the share of college graduates among the adult population (aged 15 and older). The former provides a summary measure of human capital accumulation in a given country while the latter reflects the part of the labor force that is typically more valued by receiving societies. For example, having a degree is the definition of high skill in the widely used brain drain dataset of Docquier, Lowell and Marfouk (2009) or in the most recent one by Brücker, Capuano and Marfouk (2013).

There is a possibility that the Barro-Lee dataset and the EPA might have inherent measurement errors. In fact, if we look at the difference between the Barro-Lee dataset and the EPA-implied education levels for all of Spain, we can observe that the Barro-Lee dataset underestimates the share of college graduates in the Spanish population by about 1.1 percentage points on average (Figure 1)[6]. On the contrary, the number of years of schooling is overestimated by an average of half a year (Figure 2).

---

[5] Only available for non-Spanish citizens before 2007. In theory, this could lead to some biases but it will be shown later that these are not relevant empirically.

[6] The biggest difference between both series corresponds to the second quarter of 2005. The Barro-Lee quarterly series is created by linearly interpolating the actual data.

This could lead to estimate a positively biased degree of selection in education when using the college graduates share and a negatively biased degree of selection when using average years of schooling since migrants' education is obtained from EPA while non-migrants' education comes from Barro-Lee[7].

In addition to the Barro-Lee bias with respect to the EPA, Figures 1 and 2 show the evolution of average education levels in Spain and how they have continuously improved during the period. In terms of the share of college graduates aged 15 and older in the Spanish population, their proportion increased from 9.8 percent in the first quarter of 1996 to 17.7 percent in the last quarter of 2013. Average schooling years present a more difficult time comparison due to some methodological changes in the EPA[8] but the raw numbers also denote a significant increase: from 7.9 years to 10.1 between 1996 and 2013.

The Barro-Lee dataset gathers educational data for as many as 146 countries every five years between 1950 and 2010. The EPA dataset distinguishes 94 countries and groups of countries of origin for immigrants to Spain until 1998 and 131 thereafter. There is a perfect match for 80 countries of origin until 1998 and for 99 after that.

Figure 3 shows the evolution of the migration stock in Spain between 1996 and 2013 and which part of it has been matched to origin education in the Barro-Lee dataset.

Three series are presented in Figure 3. First, the solid line represents the evolution of the total number of immigrants, defined as foreign-born individuals, living in Spain in each quarter between 1996 and 2013. This total number increases from 781,000 immigrants in the first quarter of 1996, 2 percent of the resident population, to 6,196,000 in the fourth quarter of 2011, 13.5 percent of the population. The latest number, in the fourth quarter of 2013, lies at 5,824,000 immigrants, down to 12.8 percent of the population.

---

[7] In order to correct for this type of bias, both estimates with the raw data (presented in the paper) and estimates in which the Barro-Lee data are inflated (or deflated) by the observed factor of difference obtained from comparing them to the EPA have been computed. The qualitative results remain the same. Results available from the author upon request.

[8] In particular, the education classification changed in the first quarter of 2000.

FIGURE 1

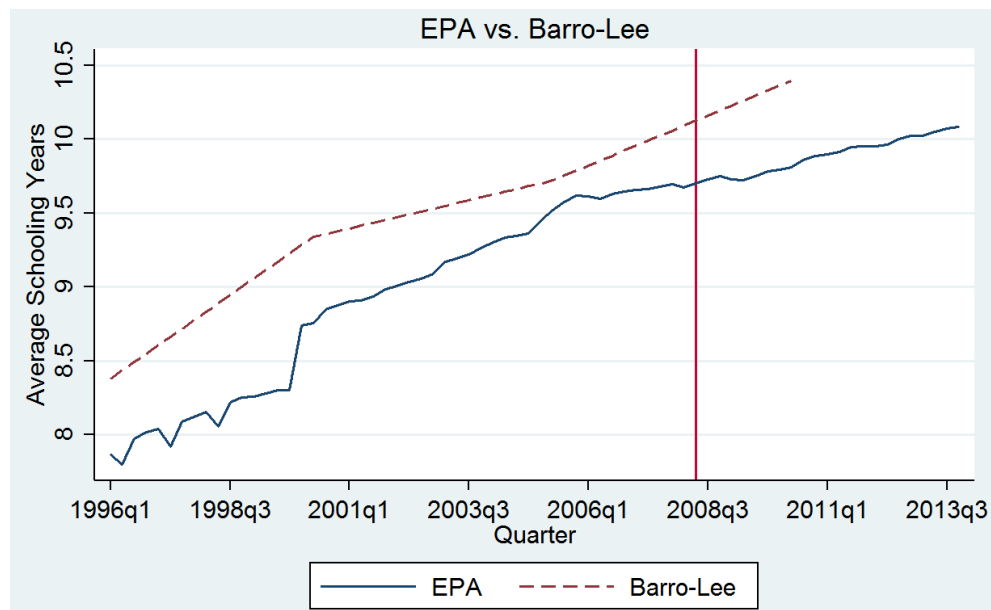**SHARE OF COLLEGE GRADUATES 15 AND OLDER IN SPAIN (1996-2013)**
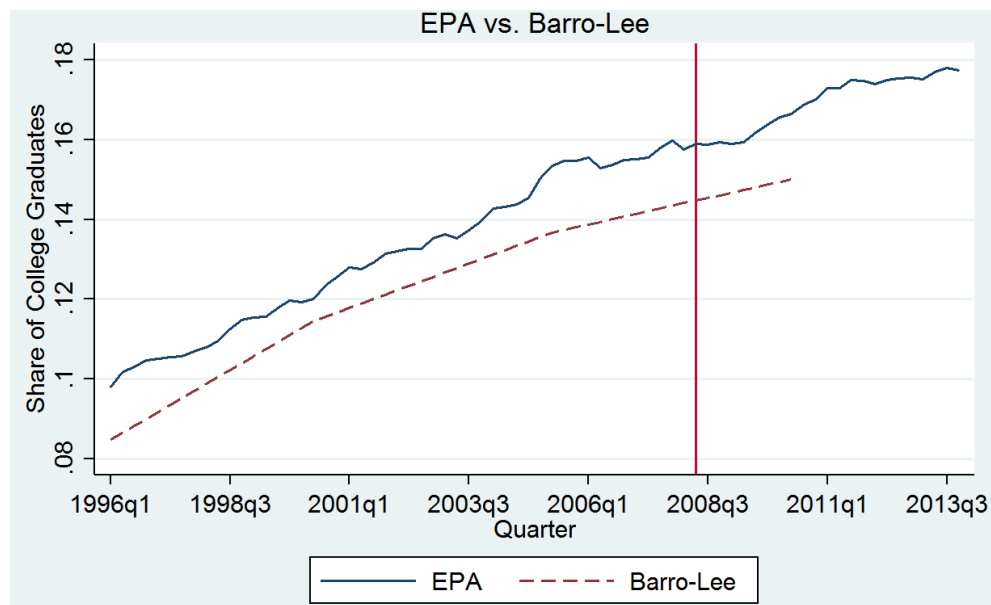
NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2). Quarterly Barro-Lee series are obtained by interpolating the actual five-year data centered at the second quarter of the reference year.

The dashed line in Figure 3 refers to the same concept restricted to those individuals who were 15 years old and older at the time of the EPA survey. The total numbers are obviously smaller but the overall evolution is very similar since this age group accounts on average for 93.5 percent of the total immigrant population in Spain.

Finally, the line with dash and dots in Figure 3 shows the evolution of the part of the immigrant population aged 15 and older that will be taken into account to compute immigrant selection measures in this paper, that is, the part of the EPA immigrant sample that has been matched to the Barro-Lee dataset. The matched sample accounts on average for 97.4 percent of the whole population of immigrants 15 and older.

FIGURE 2

**AVERAGE SCHOOLING YEARS FOR INDIVIDUALS 15 AND OLDER IN SPAIN
(1996-2013)**

NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2). Quarterly Barro-Lee series are obtained by interpolating the actual five-year data centered at the second quarter of the reference year.

The match between the EPA and the Barro-Lee dataset is based on two variables. The first one is the country of birth, explained above, and the second is the year of arrival. The EPA contains information about the year of arrival in Spain for all non-citizens before the first quarter of 2007 and for almost all foreign-born after that date. The EPA year of arrival is then matched to the closest 5-year date available in Barro-Lee. For example, if there is a Moroccan-born individual who declared in the EPA arriving in Spain in 1998, her actual education level is that reflected in the EPA while her counterfactual education level will be the education level of the typical Moroccan resident in 2000 according to Barro-Lee. If she declared arriving in Spain in 1997, her counterfactual education level will be the 1995 education level in Barro-Lee for all Moroccans. This way, the actual EPA-declared education level of all immigrants in Spain can be compared to the education level of the typical Moroccan inhabitant by the time they left their country.

FIGURE 3

EVOLUTION OF THE STOCK OF IMMIGRANTS, IMMIGRANTS AGED 15 AND OLDER, AND IMMIGRANTS AGED 15 AND OLDER MATCHED TO BARRO-LEE. SPAIN (1996-2013)
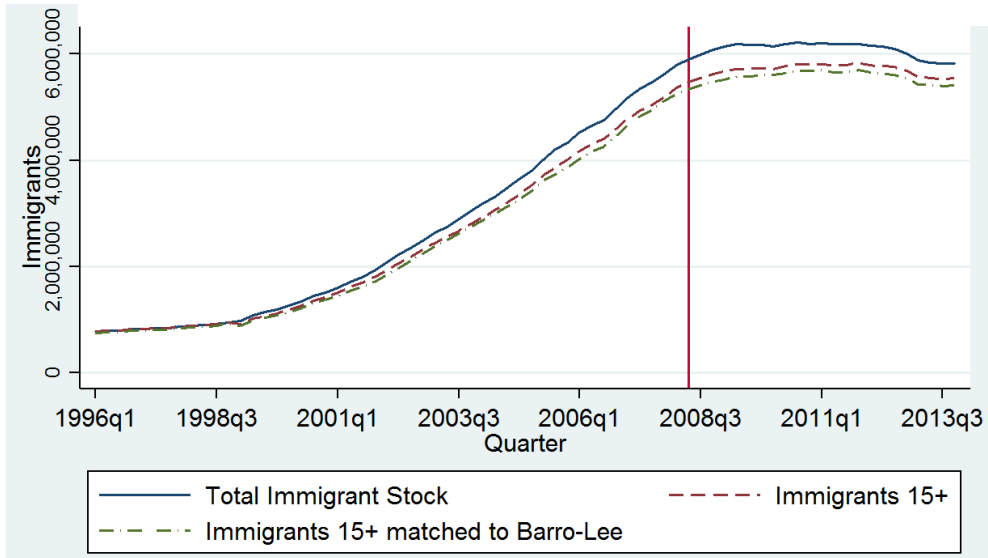
NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2).

For the very few immigrants arriving in Spain at any date before 1950, 1950 is taken as their counterfactual level of education. The main mismatches in the data correspond either to countries not included in Barro-Lee, countries grouped in the EPA or foreign-born Spanish citizens. The first two problems are negligible and they jointly represent 2.6 percent of the observations[9]. The third problem is potentially more serious, given the large number of naturalized immigrants in Spain, 12.7 percent of the immigrant population in the last quarter of 2013, for example. In fact, if naturalized immigrants for whom there is no recorded date of arrival before 2007 are treated as non-matched, the overall matching rate drops from 97.4 percent to 87.8 percent. To avoid missing this almost 10 percent of observations and the sudden jump in the representativeness of the sample that takes place in 2007, the date of arrival of naturalized immigrants is assumed to be the average date of arrival for non-naturalized immigrants for each country of birth and in each of the countries of origin[10].

---

[9] The largest immigrant stocks excluded are those from Equatorial Guinea and those from Nigeria. For example, they accounted for 44 and 36 observations in the last EPA: fourth quarter of 2013.

[10] None of the results of the paper hinges on this assumption and everything is robust to treating naturalized immigrants without year of arrival as missing observations. Using the data from 2007 onwards, it can be seen that naturalized immigrants arrived on average 10 years before non-naturalized ones. Correcting the matching by country-specific differences does not alter the results of the paper. Robustness analysis available upon request.

The representativeness of the matched sample can be further examined in the following graphs. First, Figure 4 shows the evolution of the share of college graduates among the immigrant population in Spain aged 15 and older between 1996 and 2013 both for the full EPA sample and for the EPA sample matched to the Barro-Lee dataset. As it could be expected, there is hardly any difference between the full and the matched sample in terms of average education. It moves around 23 percent until 2003 and then goes down to around 15 percent at the beginning of the recession. It looks like it has stabilized somehow after that. For comparison purposes, this evolution is shown together with that of the Spanish-born population during the same period and for the same age group. It can be seen that the share of college graduates among Spanish natives aged 15 and older has steadily increased during the period: from a low of 9.6 percent in 1996 to a high of 18.1 percent in 2013.

The difference between the immigrant and the native series, together with its 95 percent confidence interval, is represented in Figure 5. First of all, it must be emphasized that Figure 5 does not refer to immigrant selection. Figure 5 just shows that the average education level of the immigrants to Spain relative to the education level of Spanish nationals went down during the immigration boom: from around 10 percentage points to basically 0. However, Figure 4 also shows that this had more to do with the continuous increase in the average education level of the Spanish population, also clear in Figure 1, rather than with the evolution of the education level of the stock of immigrants in Spain itself.

FIGURE 4
EVOLUTION OF THE SHARE OF COLLEGE GRADUATES AMONG SPANISH IMMIGRANTS AND NATIVES 15+ (1996-2013)
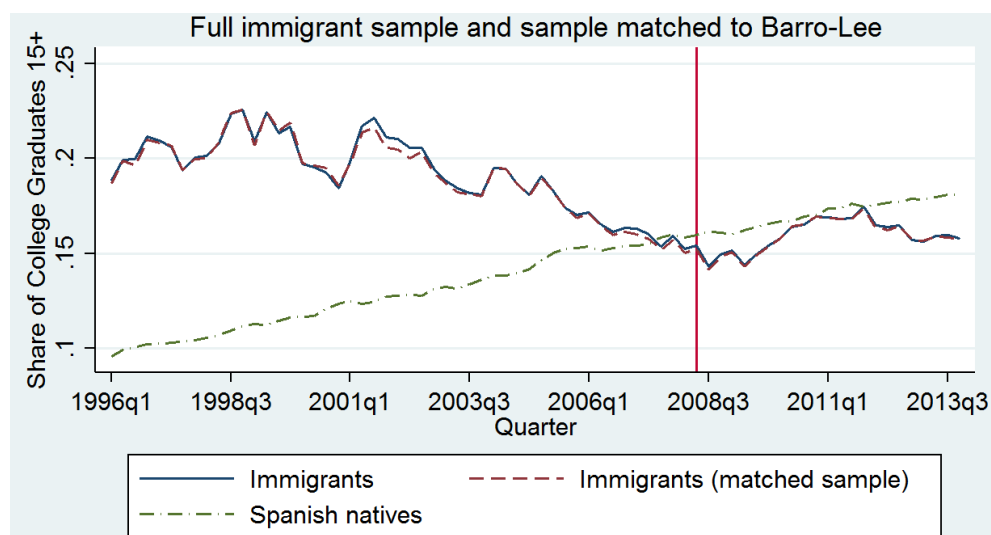
NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2).

Still, even though it does not represent immigrant selection, Figure 5 has a useful interpretation in terms of whether the average immigrant to Spain contributes to increasing or decreasing the average level of human capital in the Spanish population. The graph shows that immigrants to Spain contributed to increasing the average level of human capital at a decreasing rate during the immigration boom. At the time the recession hit, this contribution was basically 0 and it has continued to be close to 0 with some significantly small negative contributions during the second half of 2008, 2009 and again since 2012: the difference became negative by as much as 2 percentage points at the end of the period.

The same type of analysis can be repeated using years of schooling as the relevant human capital measure. Figure 6 reflects the evolution of the average years of schooling for all immigrants aged 15 and older living in Spain between 1996 and 2013. Again, the difference between the full immigrant sample and the sample matched to the Barro-Lee dataset is hardly noticeable. The comparison with the average years of schooling for Spanish natives is more interesting. Differently from the evolution of the share of college graduates shown above, the average years of schooling of immigrants are always above those of natives and remain quite stable around 10.4-10.5. In the meantime, average years of schooling for Spanish natives aged 15 and older rose relentlessly from 7.8 to 10 at the end of the period.

FIGURE 5

DIFFERENCE BETWEEN THE SHARE OF COLLEGE GRADUATES FOR SPANISH IMMIGRANTS AND NATIVES 15+ (1996-2013)
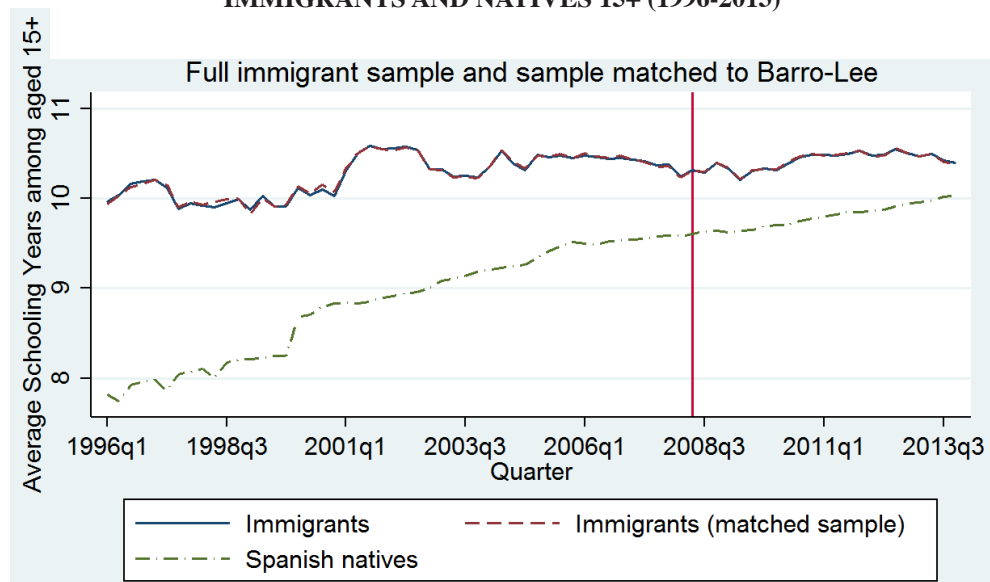

NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2).

The difference between schooling years for immigrants and natives is represented alone with its corresponding confidence interval in Figure 7. Figure 7 confirms that the difference is always significantly positive. However, the time pattern is analogous to that observed in Figure 5. The education level of immigrants relative to that of natives decreases steadily during the immigration boom and then stabilizes once the recession starts. The only distinctive feature in this case is that the stabilization takes place at positive levels: around half a year of schooling.

Finally, this section ends with similarly built graphs on the evolution of immigrant and native ages and gender composition. As far as average ages are concerned, Figure 8 shows the evolution of the average age of immigrants aged 15 and older living in Spain between 1996 and 2013. This average age was around 41 years old at the beginning of the period. It then went down to around 37 by 2003 and stayed there until the end of 2009. The last four years of the sample show an increase leading again to an average age of around 40 years old by 2013. During the same period, the average age of the Spanish population aged 15 and older increased continuously from 44 years in 1996 to almost 49 at the end of 2013.

FIGURE 6

EVOLUTION OF THE SCHOOLING YEARS AMONG SPANISH IMMIGRANTS AND NATIVES 15+ (1996-2013)

NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2).

FIGURE 7

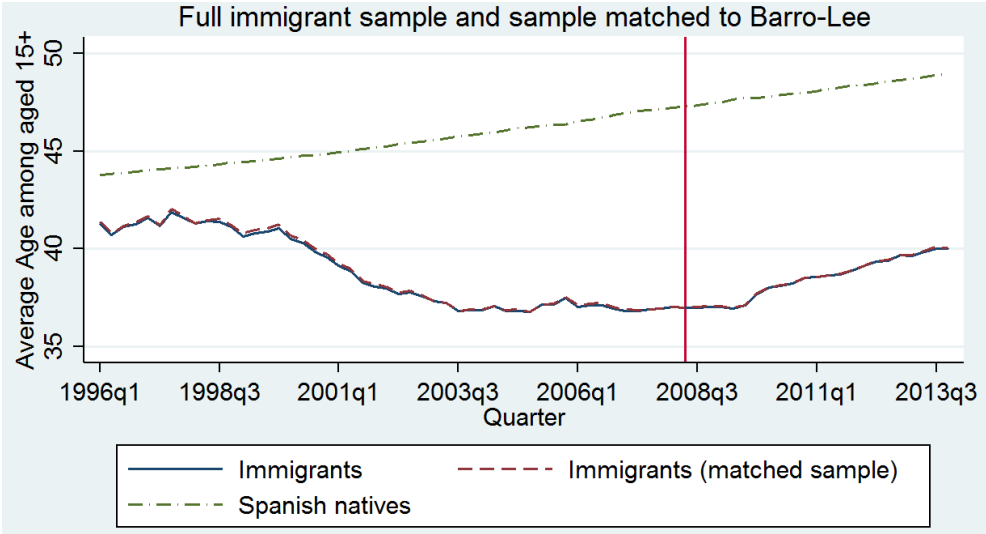**DIFFERENCE BETWEEN SCHOOLING YEARS FOR SPANISH IMMIGRANTS AND NATIVES 15+ (1996-2013)**

NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2).

Figure 9 shows the evolution of the difference between the average ages of Spanish immigrants and Spanish natives aged 15 and older between 1996 and 2013. Contrary to the education case, the resulting U-shaped pattern comes mostly from the U-shaped evolution of the immigrant average age although the ageing of the Spanish native population implies that the slope of the U is much steeper in the first few years than in the later ones. The difference goes from a minimum absolute level of 3 years of age at the beginning of the period to a maximum of 10.7 years reached in the second quarter of 2009, even after the recession had already started. It is at this point when immigrants to Spain were contributing the most towards decreasing the average age of the Spanish population. This contribution has continued to be positive although the difference has lately decreased to around 9 years of age.

In terms of gender composition, Figure 10 presents the evolution of the share of females among immigrants to Spain and Spanish natives between 1996 and 2013. While the native ratio has remained pretty much constant around 51 percent, the immigrant one has seen a lot of variation, typically above the native one before 2000 and also from 2010 on and at the same level in between.

The difference between the female share among immigrants and the female share among natives, with its corresponding 95 percent confidence interval, is graphed in the last few years to around 2 percentage points after being statistically equal to 0 between 2000 and 2010.

FIGURE 8

EVOLUTION OF THE AGE AMONG SPANISH IMMIGRANTS AND NATIVES 15+ (1996-2013)
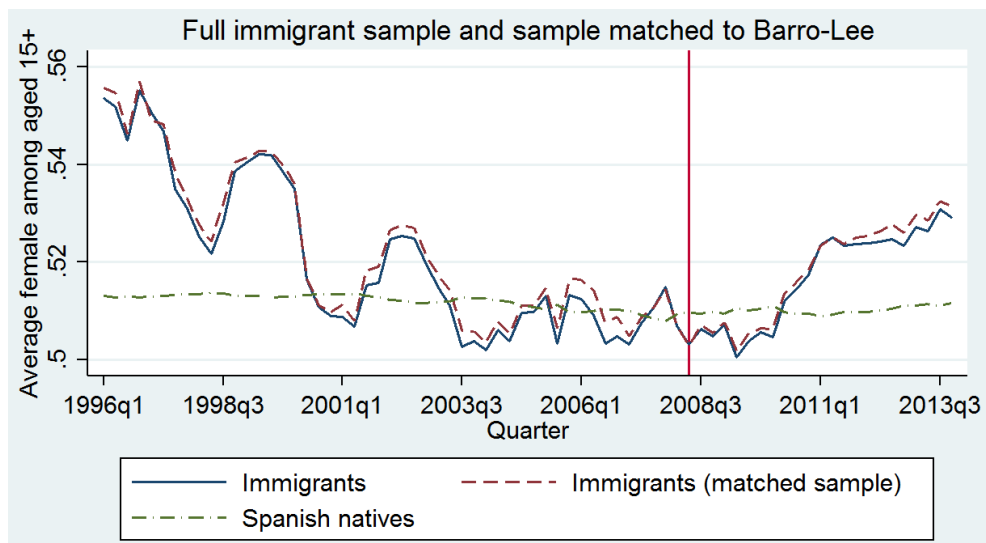

NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2).

FIGURA 9

DIFFERENCE BETWEEN AGE FOR SPANISH IMMIGRANTS AND NATIVES 15+ (1996-2013)
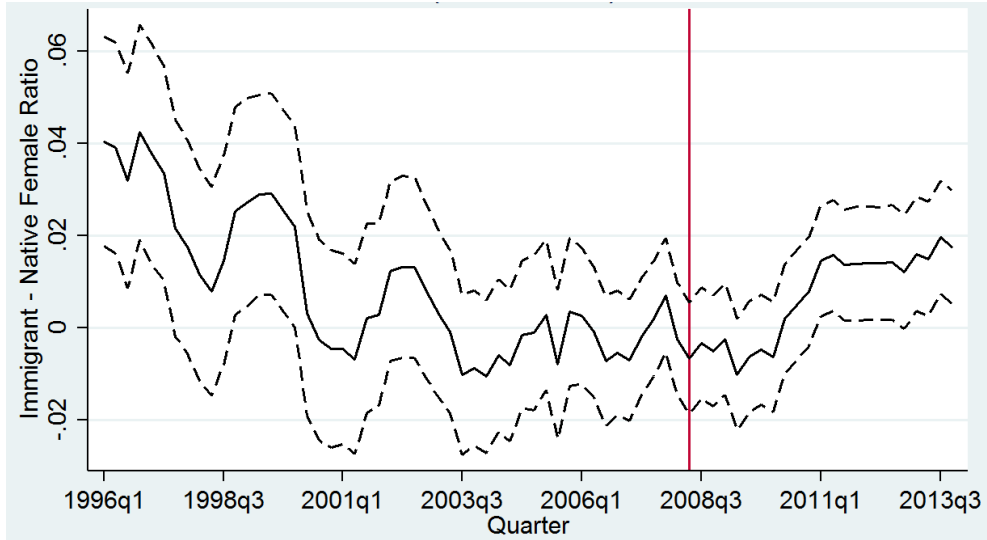

NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2).

FIGURA 10

**EVOLUTION OF THE FEMALE RATIO AMONG SPANISH IMMIGRANTS
AND NATIVES 15+ (1996-2013)**

NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2).

FIGURA 11

**DIFFERENCE BETWEEN FEMALE RATIOS FOR SPANISH IMMIGRANTS
AND NATIVES 15+ (1996-2013)**

NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2).

4. Results

4.1. Evolution

The evolution of the overall degree of selection for immigrants to Spain between 1996 and 2013 is reported in Figure 12. The solid line represents the evolution of the share of college graduates among immigrants to Spain, already shown in Figure 4. The new information in this figure corresponds to the dashed line, which represents the evolution of the counterfactual share of college graduates among immigrants to Spain. It is counterfactual in the sense that this is what the share would look like if immigrants to Spain were not selected, that is, if immigrants from each origin at each point in time were representative of the average education level in their country of origin at that time. The figure shows how the evolution of this counterfactual share did not have many noticeable changes during the period. It started at its minimum level of 4.8 percent in the first quarter of 1996 and reached a maximum of 6.7 percent in the second quarter of 2004. After that, it moved up and down between 6 and 6.4 percent to end the sample period at the latter. Despite the changes in the composition of immigrant flows to Spain, both inflows and outflows, the stock of immigrants has not shown remarkable changes in its average counterfactual education level.

The difference between the actual and the counterfactual share replicates the definition of the degree of selection in education presented in equation (2). This degree of selection in terms of the share of college graduates is shown in Figure 13. The first interesting remark that comes from observing Figure 13 is that the degree of selection of the stock of immigrants to Spain has been positive throughout the boom and the recession. The maximum level is registered in the third quarter of 1998 at 17 percentage points but it can be said that the degree of selection remains relatively constant around 15 percentage points between 1996 and 2002. After 2002, the degree of selection goes down as the Spanish immigration boom unfolds. The minimum level is recorded in the third quarter of 2008 at 8 percentage points, right after the beginning of the Great Recession in Spain. However, the degree of selection does not stay long at those low levels and after the second half of 2009 it has remained around 10 percentage points.

The evolution of the degree of selection for Spanish immigration shows that Spain is not an outlier in the typical positive selection pattern observed in most migration corridors around the world (Grogger and Hanson, 2011) even though there might be very relevant exceptions[11]. However, similarly to the US experience (Borjas, 1985),

FIGURA 12
EVOLUTION OF THE ACTUAL AND COUNTERFACTUAL SHARE OF COLLEGE GRADUATES AMONG SPANISH IMMIGRANTS 15+ (1996-2013)
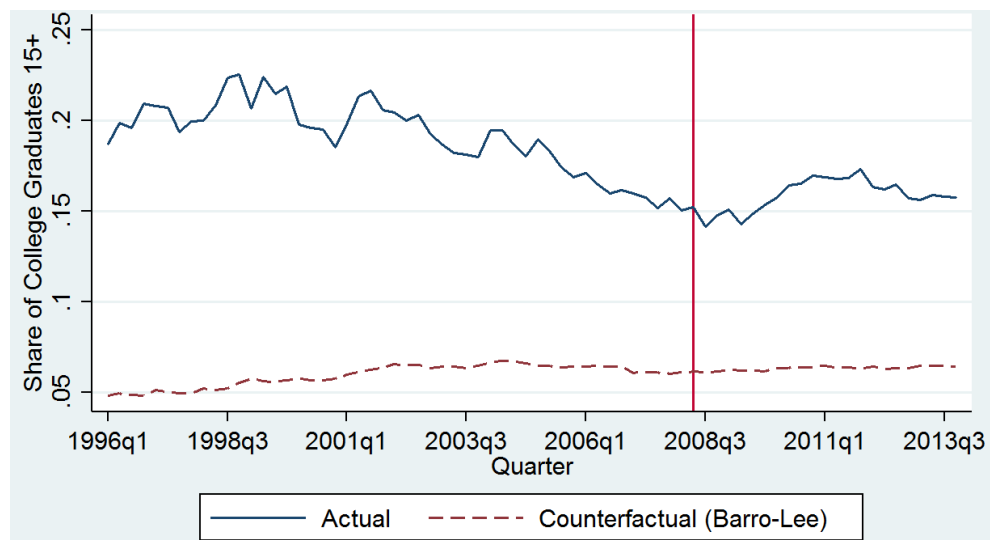
NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2).

---

[11] FERNÁNDEZ-HUERTAS MORAGA (2011) shows that migration flows from Mexico to the United States, the largest migration corridor in the world, were negatively selected between 2000 and 2004.

FIGURA 13

EVVOLUTION OF THE DEGREE OF SELECTION MEASURED BY THE SHARE OF COLLEGE GRADUATES (1996-2013)

NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2).

the selection of immigrants to Spain worsened as the numbers increased. The arrival of the Great Recession reverted this trend and the average education of the immigrant pool appears to have somehow improved as a consequence of the economic crisis although this late trend has not consolidated yet.

Next, the evolution of the degree of selection in terms of years of schooling is presented. This allows both to check the consistency of the results using a summary measure that considers all levels of education jointly and to see whether the results hold when using a measure that under-estimates positive selection in the Barro-Lee dataset in the same way that the share of college graduates was likely to over-estimate it.

Figure 14 documents first the evolution of the average years of schooling for immigrants to Spain, already shown in Figure 6, together with the evolution of the counterfactual average years of schooling. Both series appear to be similarly stable during the period. The counterfactual years of schooling do increase from around 6.5 years in 1996 to around 7.5 years by 2004 but there have been no significant changes during the last 10 years.

Figure 15 shows a bit more action. There is a significant decrease in the degree of selection in terms of years of schooling from around 3.5 years in 1996 to little less than 3 years in 2004. No more significant differences are observed thereafter.

FIGURA 14

EVOLUTION OF THE ACTUAL AND COUNTERFACTUAL SCHOOLING YEARS AMONG SPANISH IMMIGRANTS 15+ (1996-2013)
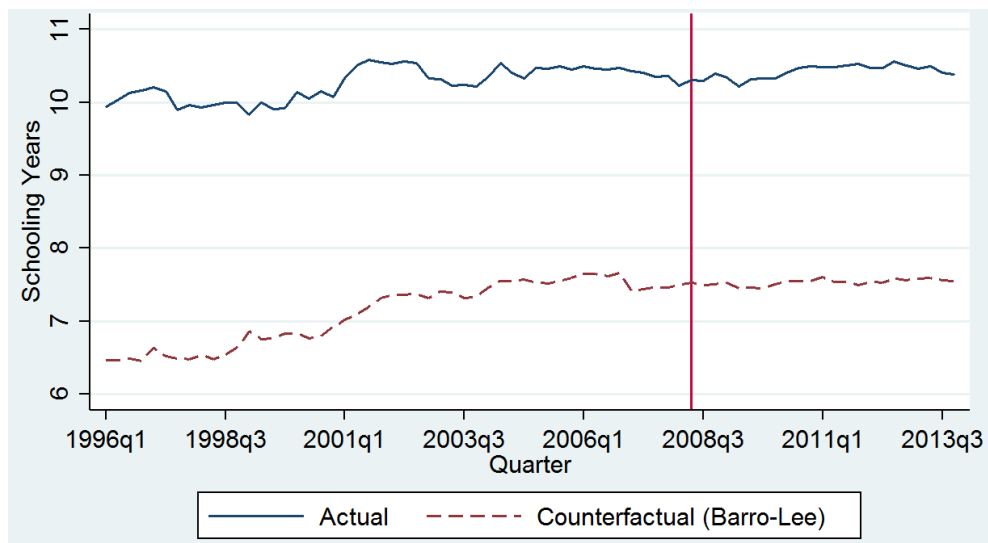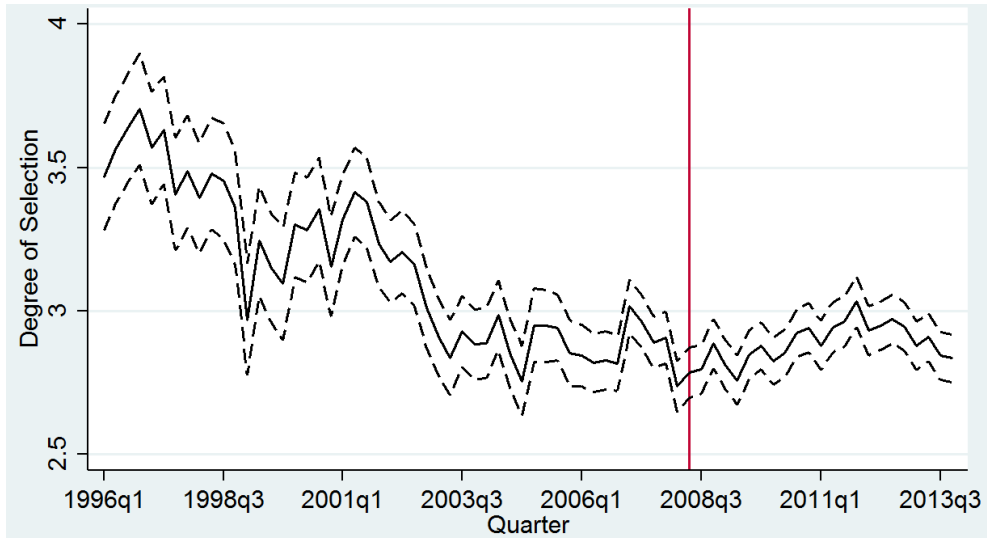

NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2).

The minimum 2.7 level is recorded in the first quarter of 2008 and there is an uptick in the degree of selection that peaks at 3 schooling years at the end of 2011 but the differences are not particularly striking.

Nevertheless, the series on years of schooling confirm the main patterns reflected in the evolution of the actual and counterfactual share of college graduates. The stock of immigrants to Spain was positively selected both before and after the recession. However, the degree of selection somehow worsened as the immigration boom developed and this worsening was stopped and maybe even reversed after the crisis hit Spain.

FIGURA 15

EVOLUTION OF THE DEGREE OF SELECTION MEASURED BY THE YEARS OF SCHOOLING (1996-2013)

NOTE: The vertical line marks the beginning of the Great Recession in Spain (2008q2).

4.2. Flows

As mentioned in the theory section, the closest one can get to the correct flow concept of selection is to look at the characteristics of immigrants who arrived in Spain at approximately the same time, without aggregating stocks of individuals who arrived during many different years.

In order to concentrate on the potential effect of the Great Recession, this subsection takes two pictures of recently arrived immigrants in Spain. The pre-Recession one corresponds to the EPA from the fourth quarter of 2007 and concentrates on those immigrants who arrived during the previous six years, that is, between 2002 and 2007. The post-Recession picture will reflect the arrival of immigrants once the crisis hit. It corresponds to the EPA from the fourth quarter of 2013 and concentrates on immigrants arrived between 2008 and 2013.

The main problem with this type of analysis is the different magnitude of return migration processes in both periods. Although it is argued that selection on return migration tends to strengthen initial selection patterns (Borjas and Bratsberg, 1996; Lacuesta, 2010), there is no guarantee that this needs to be the case in Spain. As a result, the view presented here will be quite a Spain-centric one, since the observed outcomes will be the consequence not just of initial selection at origin but also of

subsequent selection into return migration, which further complicates the already diffcult task of performing meaningful tests of the theory.

With this in mind, it is nonetheless useful to compare immigrant selection and other observable characteristics of the immigrants who arrived and chose to remain in Spain during the six years that preceded the Great Recession (2002-2007) and during the six years that succeeded the Great Recession (2008-2013). This exercise is performed in Table 1.

Table 1 presents first in its first line the evolution of the magnitude of migration flows to Spain in the six years preceding and succeeding the economic crisis. It is clear that many more immigrants aged 15 and older chose to go and remain in Spain between 2002 and 2007, more than 2.4 million, than between 2008 and 2013, barely 0.8 million. The post-crisis figure is not even one third of the pre-crisis one.

This smaller magnitude of the flows was accompanied by an increased feminization. The female share of the immigrants went up significantly from 53 percent before the crisis to 59 percent after the crisis. In terms of age, the post-crisis flows were 1.6 years older than the pre-crisis ones: from 33.2 years of average age in 2002-2007 to 34.8 in 2008-2013.

Regarding human capital accumulation, the Great Recession brought an improvement on the average education of immigrant inflows into Spain. The share of college graduates among immigrants went up significantly from 14 percent before the crisis to 23 percent after it: 9 percentage points more or an impressive 60 percent increase in this measure in relative terms. The result is not so impressive when looking at the evolution of the average years of schooling among immigrants. In this case, the increase is barely of 0.5 years of schooling after the crisis, from 10.3 average years in 2002-2007 to 10.7 in 2008-2013.

Finally, the two measures of the degree of selection introduced earlier in the paper, the difference between the educational level of immigrants to Spain and that of stayers in the respective origin countries measured either through the share of college graduates or through the average years of schooling, again tell the same consistent story. First, immigrant selection was positive both before and after the crisis using this peculiar flow concept of comparing immigrants who both came to Spain over a six-year period and did not leave with their typical origin-country populations. Second, the degree of selection improved significantly after the crisis, both in terms of the share of college graduates among the immigrants and in terms of their average years of schooling.

In the case of the degree of selection calculated by comparing the actual share of college graduates among the immigrants aged 15 and older who arrived and remained in Spain in 2002-2007 and the counterfactual share of the typical resident of the different origin countries, its value was 8 percent. This degree of selection increased to 15 percent for immigrants arrived in 2008-2013, which gives a total increase in the degree of selection of 7 percentage points: an 88 percent relative increase.

Using years of schooling to measure immigrant selection leads to the same conclusion in a less spectacular way. The degree of selection increases from 2.0 to 2.3 years of schooling between 2002-2007 and 2008-2013: a 0.3 years increase or a 14 percent relative increase.

TABLE 1

IMMIGRANT SELECTION IN SPAIN BEFORE AND AFTER THE GREAT RECESSION

	2002-2007	2008-2013	Difference
Inflows	2,418,839 [30,655]***	786,765 [23,332]***	–1,632,074 [46,646]***
Share Female	0.529 [0.009]***	0.585 [0.016]***	0.056 [0.019]***
Average Age	33.179 [0.214]***	34.768 [0.436]***	1.589 [0.486]***
Share of College Graduates	0.141 [0.006]***	0.227 [0.014]***	0.086 [0.016]***
Degree of Selection (share of college graduates)	0.078 [0.006]***	0.147 [0.014]***	0.069 [0.016]***
Average Schooling Years	10.262 [0.068]***	10.719 [0.130]***	0.458 [0.147]***
Degree of Selection (schooling years)	2.015 [0.067]***	2.297 [0.113]***	0.282 [0.131]**
Observations	4,005	1,498	

NOTE: Standard errors in brackets. *** significant at the 99 percent level, ** significant at the 95 percent level, * significant at the 90 percent level. 2002-2007 refers to immigrants aged 15 or older at the time of the survey who arrived six years before the crisis according to the 2007q4 EPA and 2008-2013 refers to immigrants aged 15 or older at the time of the survey who arrived six years after the crisis according to the 2013q4 EPA. Degree of selection computed matching the EPA data to the Barro-Lee dataset as described in the text. Statistics computed only on matched observations with non-missing data.

Despite the relatively loose language used in this subsection, it must be emphasized that Table 1 does by no means represent an attempt at causally identifying the effects of the Great Recession on immigrant selection. The exercise performed here is merely descriptive although still relevant to understand more clearly selection patterns over the business cycle and to document some stylized facts that models should be able to explain.

4.3. Main Origin Countries

The above analysis could in principle be replicated for each of the countries of origin of immigrants to Spain. However, the number of available observations in the different EPA's is a limiting factor. Thus, this subsection carries out a comparative analysis of the evolution of the degree of selection for those countries for which there are at least 50 available observations both in the fourth quarter of 2007 and in the fourth quarter of 2013. There are only 8 countries for which this condition is satisfied and they are presented in Table 2.

These 8 countries jointly account for 56 percent of the total migration flows to
Spain reflected in Table 1 both during the period 2002-2007 and the period 2008-2013.
Table 2 reflects a great deal of heterogeneity in terms of selection experiences. Still,
some common patterns emerge. The most robust one is the prevalence of positive
selection for both the pre-crisis and the post-crisis period. There is only one case of
significant negative selection among the 16 computed degrees of selection: negative
selection in education for emigrants from Ecuador to Spain in the post-crisis period
(2008-2013). However, this case is not robust to the choice of the relevant education
variable since this negative selection result turns significantly positive when using
average schooling years instead of the share of college graduates.

TABLE 2
SELECTION OF IMMIGRANTS TO SPAIN FOR MAIN ORIGIN COUNTRIES BEFORE
AND AFTER THE GREAT RECESSION

	Degree of Selection: College Graduates			Degree of Selection: Schooling Years		
Country	2002-2007	2008-2013	Difference	2002-2007	2008-2013	Difference
Romania (1,023)	0.040 [0.012]***	0.058 [0.034]*	0.018 [0.036]	0.417 [0.135]***	0.105 [0.284]	–0.311 [0.314]
Morocco (630)	–0.004 [0.010]	0.011 [0.019]	0.015 [0.022]	2.638 [0.231]***	2.391 [0.358]***	–0.247 [0.426]
Ecuador (464)	–0.020 [0.013]	–0.054 [0.019]***	–0.034 [0.023]	1.935 [0.185]***	1.906 [0.299]***	–0.028 [0.352]
Colombia (392)	0.052 [0.026]**	0.172 [0.049]***	0.119 [0.055]**	3.132 [0.278]***	3.928 [0.380]***	0.796 [0.471]*
Peru (244)	0.053 [0.035]	0.098 [0.052]*	0.046 [0.062]	2.868 [0.244]***	3.072 [0.310]***	0.204 [0.394]
Dominican R. (160)	0.123 [0.045]***	0.106 [0.052]**	–0.017 [0.069]	3.409 [0.402]***	3.734 [0.464]***	0.325 [0.614]
Cuba (155)	0.323 [0.062]***	0.042 [0.044]	–0.281 [0.076]***	2.562 [0.367]***	1.176 [0.304]***	–1.385 [0.477]***
Paraguay (118)	0.127 [0.047]***	0.100 [0.065]	–0.026 [0.080]	3.653 [0.328]***	1.679 [0.560]***	–1.974 [0.649]***

NOTE: Available observations in parentheses. Standard errors in brackets. *** significant at the
99 percent level, ** significant at the 95 percent level, * significant at the 90 percent level. 2002-
2007 refers to immigrants aged 15 or older at the time of the survey who arrived six years before the
crisis according to the 2007q4 EPA and 2008-2013 refers to immigrants aged 15 or older at the time
of the survey who arrived six years after the crisis according to the 2013q4 EPA. Degree of selection
computed matching the EPA data to the Barro-Lee dataset as described in the text. Origin countries are
those for which there were at least 50 valid observations available both in the 2007q4 EPA and in the
2013q4 EPA.

When it comes to the evolution of the degree of selection, no such dominant pattern emerges. Despite the fact that Table 1 shows a significant improvement on the degree of selection both in terms of the college share and in terms of the average years of schooling, this result is only confirmed on a origin-by-origin basis in Table 2 for the case of immigrants to Spain coming from Colombia. The rest of the main countries of origin show mostly non-significant evolutions of the degree of selection with varying signs. The only exceptions are Paraguay, which shows significant worsening selection when computed from schooling years, and Cuba, where the worsening selection pattern is robust across the share of college graduates and average years of schooling. The Cuban exception comes from the fact that their immigrants in Spain were extremely positively selected before the crisis (38 percent of them were college graduates) while this went down to a more modest 14 percent after the crisis. This may have had to do with less restrictive Cuban emigration policies after Fidel Castro's resignation in 2008. It is also remarkable that Ecuador shows significant negative selection in terms of the share of college graduates for the 2008-2013 interval even though the difference is not statistically significant with respect to the pre-crisis period. This is not surprising since Bertoli, Fernández-Huertas Moraga, and Ortega (2011) had already shown that Ecuadorians immigrants in Spain were negatively sorted with respect to Ecuadorian immigrants in the US and hence a less protracted economic crisis there could lead to even more negative sorting due to the substitutability between Spain and the US as destinations for Ecuadorian potential migrants (Bertoli, Fernández-Huertas Moraga, and Ortega, 2013). This might be an example of the possibly unintuitive effect of multilateral resistance to migration on the selection of migrants to a particular destination.

5. Conclusion

The current paper has documented the evolution of immigrant selection over the business cycle using the specific example of immigration to Spain before the Great Recession, when Spain became home to the fastest growing immigrant population in the world, and after the crisis, when net migration flows to Spain reversed even while Spain continued to receive substantial gross migration flows.

Two different ways of looking at selection have been presented, both imperfect but representing good approximations given the data constraints. The first view has looked at immigrant stocks living in Spain over time and asked what their composition would have been if immigrants had not been selected from their origin countries. This exercise has shown that Spanish immigration stocks were quite positively selected on average before the crisis arrived although the degree of selection worsened as the immigration boom progressed. Once the Great Recession hit Spain, the process of worsening selection halted or even reversed according to some measures.

The second view has focused on a broad concept of migration flows, rather than stocks. More precisely, the flow of migrants who arrived and remained in Spain

during the six-year period going from 2002 to 2007 has been compared to the flow of migrants who arrived and remained in Spain during the six-year period going from 2008 to 2013. The result of this exercise has been consistent with the first one. Firstly, migration flows to Spain have been shown to be positively selected both before and after the Great Recession. Secondly, the selection of migrants has improved from the first period to the second.

These general patterns have coexisted with some exceptions that have emerged once selection in individual countries of origin has been considered. For example, the degree of selection for immigrants from Ecuador to Spain has been shown to be negative by some measures after the crisis.

Finally, the paper has also provided some descriptive results on the evolution of other observable characteristics for immigrants in Spain, such as the feminization and ageing of the flows after the beginning of the crisis.

It would be interesting to deepen the analysis presented here by relating it to the different migration policies followed by the Spanish government before and after the economic crisis and also by going beyond descriptive statements and trying to establish causal relationships. Both of these elements constitute relevant topics for future research.

The welfare impact of immigration can crucially depend on the selection of migrants at origin but this must be coupled with an analysis of how they assimilate in their destinations. This paper has neglected this very relevant element but it must be taken into account that not only the educational level of immigrants matters but also how this educational level is translated into the receiving labor market, as many immigrants can end up working in jobs for which they are over-qualified, particularly in Spain and particularly once the recession hit (Rodríguez-Planas and Nollenberger, 2014).

References

[1] AMBROSINI, J. W. and PERI, G. (2012): "The Determinants and the Selection of MexicoUS Migrants", *The World Economy,* 35(2), 111-151.

[2] AMUEDO-DORANTES, C. and DE LA RICA, S. (2011): "Complements or substitutes? Task specialization by gender and nativity in Spain", *Labour Economics,* 18(5), 697-707.

[3] BARRO, R. J. and LEE, J. W. (2013): "A new data set of educational attainment in the world, 1950-2010", *Journal of Development Economics,* 104(C), 184-198.

[3] BELOT, M. V. and HATTON, T. J. (2012): "Immigrant Selection in the OECD", *Scandinavian Journal of Economics,* 114(4), 1105-1128.

[4] BENTOLILA, S.; DOLADO, J. J. and JIMENO, J. F. (2008): "Does immigration affect the Phillips curve? Some evidence for Spain", *European Economic Review,* 52(8), 1398-1423.

[5] BERTOLI, S. and FERNÁNDEZ-HUERTAS MORAGA, J. (2013): "Multilateral resistance to migration", *Journal of Development Economics,* 102(C), 79-100.

[6] BERTOLI, S.; FERNÁNDEZ-HUERTAS MORAGA, J. and ORTEGA, F. (2011): "Immigration Policies and the Ecuadorian Exodus", *World Bank Economic Review,* 25(1), 57-76.

[7] BERTOLI, S.; FERNÁNDEZ-HUERTAS MORAGA, J. and ORTEGA, F. (2013): "Crossing the border: Self-selection, earnings and individual migration decisions", *Journal of Development Economics*, 101(C), 75-91.

[8] BIAVASCHI, C. and ELSNER, B. (2013): "Let's Be Selective about Migrant Self-Selection", *IZA Discussion Papers* 7865, Institute for the Study of Labor (IZA).

[9] BORJAS, G. J. (1985): "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants", *Journal of Labor Economics,* 3(4), 463-89.

[10] BORJAS, G. J. (1987): "Self-Selection and the Earnings of Immigrants", *American Economic Review,* 77(4), 531-553.

[11] BORJAS, G. J. (1999): "The economic analysis of immigration", in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, vol. 3 of Handbook of Labor Economics, chap. 28, pp. 1697-1760. Elsevier.

[12] BORJAS, G. J. (2013): "The Slowdown in the Economic Assimilation of Immigrants: Aging and Cohort Effects Revisited Again", *NBER Working Papers* 19116, National Bureau of Economic Research, Inc.

[13] BORJAS, G. J. and BRATSBERG, B. (1996): "Who leaves? The Outmigration of the Foreign Born", *Review of Economics and Statistics*, 78(1), 165-176.

[14] BRÜCKER, H.; CAPUANO, S. and MARFOUK, A. (2013): "Education, gender and international migration: insights from a panel-dataset 1980-2010", mimeo.

[15] CARRASCO, R.; JIMENO, J. and ORTEGA, A. (2008): "The effect of immigration on the labor market performance of native-born workers: some evidence for Spain", *Journal of Population Economics*, 21(3), 627-648.

[16] CEBOLLA BOADO, H.; GONZÁLEZ FERRER, A.; CEA D'ANCONA,M.; FERNÁNDEZ-HUERTAS MORAGA, J.; FINOTELLI, C.; GONZÁLEZ LUNA, L.; MOYA MALAPEIRA, D. and PINYOL JIMÉNEZ, G. (2013): *Inmigración. ¿Integración sin Modelo?* Colección España: Política y Sociedad. Alianza Editorial. Madrid.

[17] CHIQUIAR, D. and HANSON, G. H. (2005): "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States", *Journal of Political Economy,* 113(2), 239-281.

[18] DE LA RICA, S. and ORTEGA, F. (2012): "Economic and Cultural Gaps among Foreign-born Minorities in Spain", in *Cultural integration of Immigrants in Europe*, ed. by Y. Algan, A. Bisin, Manning and T. Verdier, Studies of Policy Reform, pp. 148-171. Oxford Economic Press.

[19] DOCQUIER, F.; LOWELL, B. L. and MARFOUK, A. (2009): "A Gendered Assessment of Highly Skilled Emigration", *Population and Development Review,* 35(2), 297-321.

[20] FARRÉ, L.; GONZÁLEZ, L. and ORTEGA, F. (2011): "Immigration, Family Responsibilities and the Labor Supply of Skilled Native Women", *The B.E. Journal of Economic Analysis & Policy,* 11(1), 1-48.

[21] FARRÉ, L.; ORTEGA, F. and TANAKA, R. (2014): "Immigration and School Choices in the midst of the Great Recession", mimeo.

[22] FERNÁNDEZ-HUERTAS MORAGA, J. (2011): "New Evidence on Emigrant Selection", *The Review of Economics and Statistics,* 93(1), 72-96.

[23] FERNÁNDEZ-HUERTAS MORAGA, J. (2013): "Understanding different migrant selection patterns in rural and urban Mexico", *Journal of Development Economics*, 103(C), 182-201.

[24] GONZÁLEZ, L. and ORTEGA, F. (2011): "How do very open economies adjust to large immigration flows? Evidence from Spanish regions", *Labour Economics,* 18(1), 57-70.

[25] GONZÁLEZ, L. and ORTEGA, F. (2013): "Immigration And Housing Booms: Evidence From Spain", *Journal of Regional Science,* 53(1), 37-59.

[26] GROGGER, J. and HANSON, G. H. (2011): "Income maximization and the selection and sorting of international migrants", *Journal of Development Economics*, 95(1), 42-57.

[27] INE (2014): *Encuesta de Población Activa. www.ine.es.*

[28] IZQUIERDO, M.; LACUESTA, A. and VEGAS, R. (2009): "Assimilation of immigrants in Spain: A longitudinal analysis", *Labour Economics,* 16(6), 669-678.

[29] LACUESTA, A. (2010): "A Revision of the Self-selection of Migrants Using Returning Migrants Earnings", *Annales d' Économie et de Statistique*, 97/98, 235-260.

[30] MÉNDEZ, I. and CUTILLAS, I. (2014): "Has immigration affected Spanish presidential elections results?", *Journal of Population Economics,* 27(1), 135-171.

[31] RODRÍGUEZ-PLANAS, N. (2012): "Wage and occupational assimilation by skill level: migration policy lessons from Spain", *IZA Journal of European Labor Studies*, 1:8.

[32] RODRÍGUEZ-PLANAS, N. and NOLLENBERGER, (2014): "A Precarious Position: The Labor Market Integration of New Immigrants in Spain", *MPI discussion paper,* Migration Policy Institute and International Labour Office, Washington, DC and Geneva.

[33] SANROMÀ, E.; RAMOS, R. and SIMÓN, H. (2009): "Immigrant wages in the Spanish labour market: does the origin of human capital matter?", *IEB Working Paper* 2009/8, Institut d'Economia de Barcelona (IEB).

[34] SIMÓN, H.; SANROMÀ, E. and RAMOS, R. (2008): "Labour segregation and immigrant and native-born wage distributions in Spain: an analysis using matched employeremployee data", *Spanish Economic Review,* 10(2), 135-168.

[35] UNITED NATIONS, DEPARTMENT OF ECONOMIC AND SOCIAL AFFAIRS, POPULATION DIVISION (2013): *Trends in International Migrant Stock: The 2013 Revision - Migrants by Age and Sex (Nations database, POP/DB/MIG/Stock/Rev.2013/ Age).* New York.