FDI and immigration inflows: an empirical analysis for the
Spanish case

Alicia Gómez-Tello*

June 2015
(Preliminary and incomplete: Please do not circulate)

Abstract

The objective of this paper is to investigate whether some specific factors of the local business climate attract the foreign direct investment (FDI). The idea is to analyze to which extent the entry of foreign firms is pulled by the labor market composition, which is defined by the country-born workers and their level of education, and the trade openness. We build an original database for the Spanish case by merging two datasets: SABI and MCVL. We organize information for 7 different origin country-groups, 50 provinces, and 22 sectors, for the overall period 2005-2012. Our results assess that the FDI and immigration inflows in Spain followed different patterns. In addition we detect that the foreign investors established in places with higher competitiveness and human capital with respect to the location of the domestic investors. These results support the idea that Spain displays a clear case of mismatch: it mostly attracts foreign low-skill workers while incoming FDI targets a capital-intensive environment.

*I am particularly grateful to R. Nicolini for her countinous encourage, support, and advice. I also want to thank Y. Fawaz, J.L. Roig-Sabaté, and the participants to the Applied Lunch Seminar at UAB. All errors are my own responsibility. Corresponding to the author: Department of Applied Economics, Universitat Autònoma de Barcelona, Edifici B - Campus de la UAB, 08193 Bellaterra, Barcelona, Spain. Email: alicia.gomez@uab.cat.
1 Introduction

In this paper we create a novel database to study whether two specific factors of the local business climate attract the foreign investors in Spain. In particular we are interested in detecting to which extent the new vacancies created by foreign companies are associated with the local labor market composition, which is defined as the country-born workers and their skill level, and the trade openness, which is a proxy for competitiveness.

Because the globalization and the European economic integration, the movement of factors around the world considerably increased in the second-half of the twentieth century.\(^1\) Spain was historically a net sending of migrants, but since the end of the 1990 (until the beginning of the economic crisis) the country experienced inflows of immigrants at a higher rate than other European countries (Glitz et al., 2014). According to González and Ortega (2001) the foreign-born among the working-age population increased from 4 percent in 1998 to 17 percent in 2009. Among these foreign-born workers, almost 50 percent came from Latin American countries.\(^2\)

Many papers confirm that there is a positive relationship between the movement of people and capital, namely there is a complementarity among factors. Due to the limitation of data there are not many papers studying neither the FDI determinants nor the relationship between immigration and FDI inflows in the Spanish provinces.

In this paper we build an original database by merging information from two datasets: Sistema de Análisis de Balances (SABI), which contains information at firm level, and Muestra Continua de Vidas Laborales (MCVL), which contains information at person level. Despite we handle micro data at firm and individual level, we cannot merge both datasets because they do not have a common identifier. Then, we have to semi-aggregate the available information, which is organized by considering 7 origin country-groups, 50 provinces, and 22 sectors, for overall period 2005-2012. The origin refers to the country from which the capital or the worker come from, and we distinguish between Spain and six foreign country-groups (Africa, Asia-Pacific, EU-15, Rest Europe, Latin America, and North America). The province and sector refer to the place and economic activity where the vacancy is opened or the worker is hired. As our period of analysis (2005-2012) includes years of economic expansion and years of economic crisis, it is important to control for time dummies in the econometric specification.

Taken into account the origin-province-sector identifier, we compute some variables: From

\(^1\) Nevertheless, the increase of the world migration rate was much more moderate than the increase of the ratio of foreign direct investment (FDI) to world output. (Docquier and Rapoport, 2012).

\(^2\) Latin American citizens made profit from some special agreements between Spain and its former colonies (Glitz et al., 2014).
the SABI’s dataset the number of vacancies opened by new firms, while from the MCVL’s dataset the workers that enter in the Spanish job market. Our measure of FDI inflows is the vacancies created by firms that are foreign subsidies or contains a foreign (stake) holder with at least 10 percent of the capital. Our measure of immigrant inflows is the new workers that were born in a foreign country.

We conduct a wide descriptive analysis in order to examine the relationship between our variables of interest. We find that FDI inflows came mainly from EU-15 and North America and was concentrated in services, while the immigrant inflows came mainly from Latin-America, Africa, and Rest Europe and was concentrated in construction and hotel. Then, the regional distribution of immigrants do not pull FDI entry.

In our econometric analysis we consider two different dependent variables: the foreign new vacancies (FDI inflows) and the domestic new vacancies. These dependent variables have some peculiarities, namely they take non-negative values and are over-dispersed. Nevertheless, in this preliminary econometric exercise we do not focus on these issues. We estimate a panel fixed effects model in order to detect the determinants of the new vacancies and to compare whether these determinants are the same for the foreign and domestic vacancies.

As the descriptive analysis, the econometric results confirm that FDI and immigration inflows in Spain followed different patterns. Trade openness had a higher impact on the foreign vacancies than in the domestic ones. Concerning human capital, foreign investors looked for more qualified workers than domestic investors. Therefore, we find that foreign companies wanted to maintain their high productivity level, looking for province-sectors combinations more competitive and with higher human capital.

The rest of the paper is organized as follows. Section 2 provides a brief literature review. Section 3 focus on the data, we explain the structure and the construction of our novel database and we conduct a detailed descriptive analysis in order to investigate the behavior of our variables of interest. Section 4 presents a preliminary empirical exercise, we comment the econometric specifications and discuss the results. Finally, Section 5 concludes.

2 Literature Review

In their seminal paper Helpman et al. (2004) show that only the most productive firms engage in FDI activities. Therefore, the level of productivity is a key determinant to select the firms that invest abroad. The next question is why some firms are more productive than others. The technology and the investment in physical and human capital are important factors to explain
differences in productivity, but they cover only a part of the story.

Some paper stress the importance of the organization of the production process, which is mainly characterized by the task design and the structure of hierarchy (Garicano, 2000 and Caliendo and Rossi-Hansberg, 2007). An adequate organization can solve coordination problems and consequently, increase the productivity and the competitiveness of the firm in the international markets.

The managerial strategy is also a key determinant of the productivity. Manager should be able to solve unexpected problems related to the dynamics of the markets. Bloom and Van Reenen (2007) collect management practice data from 732 medium-sized firms located in Canada, France, Germany, and the United States in order to analyze the effectiveness of the management practices in some variables. They find that management practices are positively correlated with measures of productivity.

There papers that study the relationship between the movement of capital and people and they usually find a positive relationship between the flows of these two factors. Nevertheless, the reason behind this positive relationship is not unanimous: some papers base on networks effects (e.g., Kugler and Rapoport, 2007) while other papers base on agglomeration effects (e.g., Buch et al., 2006).

On the one hand, papers that base on the network effects argue that immigrants create some business and social networks that reduce the information barriers, then the host country takes advantage from this information advantage and invests in the source countries of its immigrants. For instance, Bhattacharya and Groznik (2002) find that the size of the foreign-origin group from a country living in United States is positively correlated with the United States investment in that country. Gao (2003) finds that ethnic Chinese networks in other countries have a positive role on the inward of FDI from these host countries to China.\(^3\) In the same line, some papers stress that the relationship between FDI and migration is stronger from migrants with high degree of education (Kugler and Rapoport, 2007 and Javorcik et al., 2011).\(^4\)

On the other hand, papers that base on the agglomeration forces predict that factors from a specific source country flow toward the same host place because they can complement

---

\(^3\) The networks forces predict that factors follow an opposite direction, namely the stock of immigrants is correlated with the outward FDI or the stock of emigrants is correlated with the inward FDI.

\(^4\) The idea of the network effects was first exploited in studying the relationship between factor flows and bilateral trade. (Gould, 1994; Head and Ries, 1998; Rauch, 2001; Combes et al., 2005; and de la Mata and Llano, 2013). Gould (1994) is one of the first in assessing how immigrants’ ties with their home countries can foster the bilateral trade between their host and home countries. Immigrants’ ties mean knowledge of home-country markets, namely people preferences and business opportunities, which produces an important reduction in the information barriers.
each other. Buch et al. (2006) and Foad (2012) look at the regional distribution of immigrants and inward FDI stocks, the former in Germany and the latter in the United States. Both detect a positive relationship between both variables, namely higher stock of FDI in states holding a large foreign population from the same origin country. In addition, Foad (2012) confirms that immigration leads FDI and that the effect is higher as the skills of the immigrants increase.

In this paper we aim to examine some determinants of the FDI inflows in Spain, and we consider that one of these determinants could be the local labor market composition, which is defined as the country-born workers and their skill level.

All the previous studies use bilateral data about stock of FDI and migrants. Nevertheless, the availability of this type of information is limited (for this reason most of the studies focus on the United States). In this paper, however, we overcome this data limitation and we create some original variables in order to measure the FDI and immigration inflows in Spain.

3 Data

One of the novelties of this study is the construction of a novel database that allows us to investigate selected determinants of the FDI inflows in Spain during the period 2005-2012. This section is divided into two parts. First, we focus on the structure of the final database and the construction of the key variables, namely FDI and immigration inflows. Then, we perform a descriptive analysis in order to study the behavior of these two variables.

3.1 Database structure and relevant variables

In this paper we create an original database that allows us to study the connection between the FDI and immigration inflows. We use two main data source. On the one hand, the Sistema de Análisis de Balances (SABI)’s dataset provides information on firm activity and allows us to define the variable FDI inflows. On the other hand, the Muestra Continua de Vidas Laborales (MCVL)’s dataset provides information on workers and allows us to define our variable immigration inflows.

Although we have micro data at firm and worker level, we cannot merge this information. Then, following other studies (e.g., Buch et al., 2006) we semi-aggregate our raw data. The level of aggregation that allow us to merge the information from the two datasets is origin-province-sector. The origin refers to the home place, the country where investors or workers come from. There are seven different origins: Spain and six foreign origin country-groups.
As host place, we consider 50 Spanish provinces (Ceuta and melilla are excluded). Finally, we create an *ad-hoc* classification by sector to make information in SABI (NACE-93) comparable with the details included in the MCVL (NACE-93 and NACE-2009). Our own classification is at 2-digit level and distinguishes 22 production activities (See Table A.1 for detailed information).

The SABI’s **dataset** compiles information at a firm level by gathering data extracted from the firms’ balance sheets (according to the availability of information disclosed by *Registro Mercantil*).

In order to build our key variable we consider the following selection criterion. First we select the new firms located in Spain during the period 2005-2012 with at least two employees (in this way we excluded self-employment). Then, we distinguish domestic firms from foreign firms. A foreign firm have to fulfill one of the following two conditions: having a parent company located abroad or and accounting for a foreign (stake) holders with at least 10 percent of the capital. If neither of these two conditions are fulfilled, then the firm is considered domestic or local. Once the new firms have been classified according to its country of origin, we identify the number of vacancies (as the number of employees) created by each firm. Then, after aggregating (for each year) at origin-province-sector level, we obtain two interesting variables: new foreign vacancies (our definition of FDI inflows) and new domestic vacancies.

One of the advantages of our variable FDI inflows is that it is a real variable unaffected by the units of measure. According to Markusen (2002), the multinationals should be considered as real production units in the economy. Foad (2012) uses the number of foreign affiliates located in United States in order to study the relationship between the stock of immigrants and the inwards FDI in the United States. Instead of the number of vacancies we could use number of workplaces created by each firm (namely, the headquarter and its delegation(s) if any). Nevertheless, taking into account the number of vacancies allow us to control better for the size effect.

The second main dataset, the **MCVL**, is a set of individual and anonymous data extracted and compile by the Spanish Social Security. They record information about gender, age, civil status, nationality, country of birth, education, employment, among others.

---

5 The group ‘Rest Europe’ includes the other EU countries (namely, non-EU15) and European countries that are not members of the EU. The group ‘North America’ includes mainly Canada and the United States.

6 This sample is not random or stratified but its size (more than one million of Spanish and Portuguese firms) makes it a reliable reference for economic studies at a national level.

7 According to the OECD/IMF definition, FDI is an investment in a foreign company where the foreign investor owns at least 10 percent of the ordinary shares, undertaken with the objective of establishing a lasting interest in the country, a long-term relationship and significant influence on the management of the firm.

8 The Social Security merges its information with the population census extracted from the *Instituto Nacional de Estadística* (INE) and the Personal Income Tax extracted from the Spanish Tax Agency (AEAT).
A person will be part of the MCVL in a specific year if he/she fulfills two independent criteria. The first condition is to have a personal identifier number (identificador de persona física or IPF), namely the DNI for natives and the NIE for foreigners. The second condition is being part of the reference population, which is defined as people that had a relationship with the Social Security in the year of reference, namely affiliated in any regimen or receiving a contributive pension. Considering these two requirements, it is conducted a simple random sample: people from the reference population are selected whether their IPF contains some specific digits that were randomly selected.\textsuperscript{10} The MCVL is able to follow the same person along time (as long as he/she maintains the same IPF and be part of the reference population), but also is able to include new people following an automatic procedure (MCVL, 2015).

We are interested in people entering in the Spanish labor market, namely people that were hired in the local labor market. As before, we have to aggregate the information (for each year) at origin-province-sector, and we get two interesting variables: new foreign workers (our definition of immigration inflows) and new native workers.

Therefore, our final database has information about new vacancies and new workers at origin-province-sector during the period 2005-2012.\textsuperscript{11} Our measure of FDI inflows is the foreign new vacancies, while our measure of immigrant inflows is the foreign new workers.

### 3.2 Descriptive Analysis

In this subsection we develop a descriptive analysis in order to examine some key variables of this research, namely new vacancies and new workers.

First, we focus on the average number of vacancies opened by the new firms, which indirectly reflects the size. We distinguish between domestic or local firms and the firms that were created with FDI. Figure 1 (panel A) depicts the huge difference in the average size of these two groups of firms. The new domestic firms had 8 employees, on average, while the firms created with FDI had 48 employees, on average. This result supports the stylized fact that multinational are larger than national firms (Barba Navaretti and Venables, 2004, p. 3). If we compute the average size by sector (panel B), this difference is still higher. In the case of the firms created with FDI, some sectors had more than 80 employees, on average.\textsuperscript{12} In the case of the domestic firms, only one sector (‘800 - energy’) had more than 20 employees, on average.

\textsuperscript{9}DNI is the Documento Nacional de Identidad and NIE is the No de Identificación de Extranjeros.

\textsuperscript{10}For instance, in 2006 the reference population was 29.3 million people and the sample 1.17 million people (approximately 4 percent of the reference population).

\textsuperscript{11}The final database contains additional variables, which are are explained in the econometric specification.

\textsuperscript{12}These sectors were the following: ‘100 - Food, beverages, and tobacco,’ ‘400 - Chemical, plastic, and petroleum transformation,’ ‘500 - Metallurgy and mechanical equipment manufacture,’ and ‘1100 - Hotel.’
Figure 2 depicts the **new vacancies**. Panel A describes its trend over time. Due to the high scale differences, we have to distinguish between domestic values (left y-axis) and foreign values (right y-axis). If we observe the new domestic vacancies, which are represented with grey bars, we observe a decreasing trend: from 228,905 new vacancies in 2005 to only 63,783 in 2012. This great decrease could be consequence of the economic crisis. The foreign vacancies were mainly created with capital from EU-15. The behavior of the foreign vacancies is better described in panel B, where foreign vacancies for a specific year are normalized to 100 and it is represented the relative importance of each foreign origin country-group. The graph confirms the large relative importance of EU-15 as foreign investor: except in 2010 more than 50 percent of the foreign vacancies were created with capital from EU-15. Vacancies created from North America represented around 17 percent, while the new vacancies created from Asia-Pacific and Africa were insignificant.

Figure 3 depicts the **new hiring**. Panel A described its trend over time. New native workers (represented with grey bars) experienced a decreasing trend (from 257,927 in 2005 to 185,063 in 2012), but it was not so strong as in the case of the domestic vacancies. In the case of the new foreign hiring, the most important origin country-group was Latin America, although it also suffered a decreasing trend since 2008. Panel B focuses only on the foreign new hiring. Workers from Latin America countries represented almost 50 percent of the new foreign workers. The other important groups were Africa and Rest Europe with 16.6 and 17.7 percent of the foreign new workers, respectively. This result is consistent with other studies about the immigration patterns in Spain. Glitz et al. (2014) indicate that the most important immigrant incoming in 2011 were from Romania, Morocco and Ecuador, representing 12.4, 11.5, and 7 percent of the foreign-born population, respectively. The relative importance of the other origin country-groups was quite constant over time: the EU-15 represented 10.6 percent, Asia-Pacific 6.1 percent, and North America only 1 percent.

Finally, we refer to the **sectoral dimension** (Figure 4). More than 67 percent of new domestic vacancies were created in 5 sectors. In the case of the foreign vacancies, the concen-

---

13Glitz et al. (2014) point out that in 2008 there were more than two million immigrants coming from Latin America, representing almost 50 percent of the foreign-born working population. For this specific year, we find that the Latin American workers represented 49.4 percent of the new foreign workers, which indicates that our database is consistent that study.

14The relative importance Africa decreased over time (from 19 percent in 2005 to 12.9 percent in 2012) while the relative importance of the Rest Europe increased over time (from 13.2 percent in 2005 to 20.1 percent in 2012).

15These sectors were the following: ‘900 - Construction,’ ‘1000 - Wholesale, retail sale, and vehicle motor repair,’ ‘1100 - Hotel,’ ‘1900 - Administrative and support activities,’ and ‘2200 - Services (health, leisure, sports, and culture).’
This page contains a detailed analysis of the new foreign worker and FDI inflows in Spain, highlighting the sectors that experienced significant increases in immigration and investment. The text discusses how two sectors, construction and hotel, accounted for a large portion of new foreign vacancies, with the construction sector representing 1.2% of vacancies and 23.5% of new foreign workers, and the hotel sector representing 2.4% of vacancies and 20% of new foreign workers. The text also references the findings of González and Ortega (2001) that suggest a noticeable rise in immigrant employment in construction, services, and domestic help from 1997 to 2007.

Furthermore, the text notes that foreign direct investment (FDI) inflows came mainly from EU-15 and North America and were concentrated in services, while immigrant inflows came mainly from Latin-America, Africa, and Rest Europe and were concentrated in construction and hotel. The text concludes by stating that the FDI and immigration inflows in Spain followed a different pattern, with FDI inflows mainly from EU-15 and North America, and immigrant inflows mainly from Latin-America, Africa, and Rest Europe. The two sectors mentioned in the text were: `1000 - Wholesale, retail sale, and vehicle motor repair' and `1900 - Administrative and support activities.'
Figure 1: Average firm size. New firms (2005-2012)

Panel A: By capital origin

Panel B: By capital origin and sector

Source: Own elaboration based on data from SABI.
Figure 2: New vacancies (2005-2012)

Panel A: Trend by origin

Panel B: Foreign vacancies: Distribution by origin

Source: Own elaboration based on data from SABI.
Source: Own elaboration based on data from MCVL.
Figure 4: New vacancies versus new hiring. Relative importance by sector (2005-2012)

Panel A: New vacancies

Panel B: New hiring

Source: Own elaboration based on data from SABI (Panel A) and MCVL (Panel B).
4 Preliminary Empirical Exercise

This section, which develops a preliminary empirical exercise, is divided into two parts. In the first part, we present the econometric specification and in the second part we discuss the results.

4.1 Econometric Specification

The empirical model that we estimate is designed to explain the creation of vacancies in Spain. The baseline model is as follows:

\[
\text{vacancy}_{opst} = \alpha_1 + \alpha_2 \text{vacancy}_{other}^{opst-1} + \alpha_3 \text{hiring}_{opst}^{opst-1} + \alpha_4 \text{hiring}_{other}^{opst-1} + \alpha_5 \text{humancapital}_{pt}^{pt-1} + \alpha_6 \text{topeness}_{opst}^{opst-1} + \beta Z_{pst}^{pst-1} + \gamma X_{pt}^{pt-1} + \mu_{ops} + \theta_t + \varepsilon_{opst},
\]

where \( o = 1, \ldots, 7 \) is the origin country-group, \( p = 1, \ldots, 50 \) is province, \( s = 1, \ldots, 22 \) is sector, and \( t = 2006, \ldots, 2012 \) is year.

The dependent variable is the new vacancies generated from origin country-group \( o \) to province \( p \) in sector \( s \) and year \( t \). It depends on a constant, some explanatory variables, fixed effects (\( \mu_{ops} \)), year dummies (\( \theta_t \)), and the error term (\( \varepsilon_{opst} \)). The explanatory variables are lagged one year (\( t - 1 \)) because we assume that the act of investing is carried out one year after taking the decision.\(^{17}\)

The explanatory variables included in the econometric specification are the following: vacancies generated by other origin country-groups in the same province-sector (\( \text{vacancy}_{other}^{opst-1} \)), new workers that come from the same origin country-group and were hired in the same province-sector (\( \text{hiring}_{opst}^{opst-1} \)), new workers that come from different origin country-groups and were hired in the same province-sector (\( \text{hiring}_{other}^{opst-1} \)), the human capital in the province of destination (\( \text{humancapital}_{pt}^{pt-1} \)), and trade openness between the origin country-group and the province of destination (\( \text{topeness}_{opst}^{opst-1} \)). The human capital is the average years of education of the working-age population and it is extracted from the Bancaja and Ivie’s dataset. Trade openness is sometimes used as proxy for competitiveness (Bloom and Van Reenen, 2007). It is computed as total trade flows (sum of exports and imports flows) over gross domestic product (GDP).\(^{18}\)

In addition, we include other controls in the econometric specification (\( Z_{pst}^{pst-1} \) and \( X_{pt}^{pt-1} \)). From SABI’s dataset we compute the number of firms that shut down in a province-sector.

\(^{17}\)In the same line, Kugler and Rapoport (2007) consider that the migration effects on FDI can be contemporaneous or dynamic.

\(^{18}\)Information about trade flows is extracted from DataComex and it is available at origin-province level. The GDP at province level is extracted from the INE.
combination. This variable could be interpreted as the risk to set up a firm in a specific province-sector. Using data from INE, we create some controls at provincial level: GDP per capita and percentage of the working-age population (population aged between 15 and 64 years old). The GDP per capita is transformed into real values using the Penn World Table (pwt)’s dataset.\footnote{It is available online at: http://citaotest01.housing.rug.nl/FebPwt/Dmn/AggregateXs.mvc/PivotShow. Specifically, we use the price level of GDPo (pl,gdpo). The price level of USA GDPo in 2005 used as the reference (it is equal to one).} In the case of the working-age population, we distinguish between natives and foreigners.\footnote{As the INE includes all the EU countries in a group, it is not possible to divide the foreigners between our six origin country-groups.} In the Appendix A, Table A.2 provides detailed definition of all the variables.

We propose an alternative measure of human capital. From the new hiring we are able to identify the skill of the workers: We distinguish between high (tertiary), medium (secondary) and low (primary) level of education. In order to include these new variables into the econometric specification, we have to modify the baseline model and we get the following specification:

\[
vacancy_{opst} = \alpha_1 + \alpha_2 vacancy_{opst-1} + \alpha_3 hiring_{opst}^{highedu} + \alpha_4 hiring_{opst}^{mededu} + \alpha_5 hiring_{opst}^{lowedu} + \alpha_6 openness_{opst-1} + \beta X_{opst-1} + \gamma X_{opst} + \mu_{opst} + \theta_t + \epsilon_{opst},
\]

where \( o = 1, \ldots, 7 \) is the origin, \( p = 1, \ldots, 50 \) is province, \( s = 1, \ldots, 22 \) is sector, and \( t = 2006, \ldots, 2012 \) is year. The variables \( hiring_{opst}^{highedu} \), \( hiring_{opst}^{mededu} \), and \( hiring_{opst}^{lowedu} \) are the number of new workers in a specific province-sector with high, medium, and low level of education, respectively. In order to maintain a clear interpretation of all the variables, we have to exclude the variables \( hiring_{opst-1} \) and \( hiring_{opst-1} \) from this new specification.\footnote{Both groups of variables include all the new workers by province-sector, that is: \( hiring_{opst} = \sum_{o=1}^{7} hiring_{opst} = hiring_{opst}^{highedu} + hiring_{opst}^{mededu} + hiring_{opst}^{lowedu} \). Then, each worker comes from a specific origin and has a specific level of education. As the same worker is included in both groups, the interpretation of the coefficients will be very ambiguous.} Then, in this new specification we focus on the human capital and trade openness as determinants of the FDI inflows.

4.2 Results

In this exploratory analysis we want to study the determinants of new vacancies and whether or not these determinants are the same for the foreign and domestic vacancies. For this reason, we consider two different dependent variables, one is for the new foreign vacancies \((vacancy_{opst})\) and the other for the new domestic vacancies \((vacancy^{domestic}_{opst})\).\footnote{When the dependent variables is the foreign new vacancies, the explanatory variables \( vacancy_{opst} \) and \( hiring_{opst-1} \) of specification (1) are divided into two: other foreign country-groups \((other = other_o)\) and domestic \((other = domestic)\).}
Our dependent variables have some peculiarities, they take non-negative values and are over-dispersed (the standard deviation is higher than the mean). In addition, in the case of the new foreign vacancies, there is an excess of zeroes. At the moment, we run a panel fixed effects model that does not consider these important issues, but for the next version of the article, we are working on improving the econometric model.

We begin with using the new foreign vacancies as a dependent variable. Results are presented in Table 1. Columns (1) and (2) refer to specification (1) and columns (3) and (4) refer to specification (2). We directly focus on the sign and significance of the variables of columns (2) and (4) because they include all the explanatory variables.\footnote{The potential number of groups is 6,600 (namely, 6 $\cdot$ 50 $\cdot$ 22). In order to reduce the number of zeroes, we exclude all the province-sector combinations in which no foreign origin country-group set up a firm in the period 2005-2012. Then, we work with 2,094 groups.}

Results in column (2) assess that new foreign vacancies established from other origin countries had a negative effect on the number of vacancies. This negative sign means that the foreign investors had different objectives. Concerning the new hirings, new foreign worker (from the same group) had a negative effect on the number of vacancies, while new domestic workers had a positive effect. The former sign confirms what we find in the descriptive analysis: FDI and immigration inflows in Spain followed a different pattern (See Figure 4).

Trade openness had a very large positive effect, meaning that an increase of the trade flows between an origin country-group and a province had a positive impact on the number of vacancies opened by this origin country-group in this province. The amount of firms that closed down had a negative effect on the number of new vacancies, what means that foreign investors had incentives to locate in a province-sector combination where the companies do not have to leave the market.

Finally, we observe that the percentage of native working-age population had a positive effect on the number of foreign vacancies, what means that foreign investor had incentives to locate in provinces with larger availability of workers. Although we identify this quantitative effect, we do not isolate any effect related to the quality of the working-age population: the variable human capital (measured as the average years of education of the working-age population) had no impact on the opening of foreign vacancies.

Results in column (4) provides additional information about human capital. We observe that the hiring of high educated workers had a negative effect (\textit{a priori} it was not expected), the hiring of medium-educated workers had a positive effect, and the hiring of low-educated workers had no effect. In addition, the sign and significance of the variables that appear in column (2)
As a benchmark for the previous results we focus on the **new domestic vacancies**. Results are presented in Table 2. Columns (1) and (2) refer to specification (1) and columns (3) and (4) refer to specification (2). As before, we only comment columns (2) and (4), which include
all the explanatory variables.\footnote{In order to reduce the number of zeroes, we exclude the province-sector combinations in which there were any domestic vacancy during the period 2005-2012. There are 1,100 potential groups (namely, 50 \cdot 22), but after this procedure we work with 1,076 groups.}

From column (2) we detect that new workers (natives and foreigners) had a positive impact on the number of new domestic vacancies, while the amount of firms that close down had a negative impact on the number of new domestic vacancies. The percentage of the native working-age population had a very large positive impact, but the quality of the working-age population had no impact.

From column (4) we stress that the hiring of high educated workers had a negative effect, while the hiring of medium- and low-educated workers had a positive effect on the number of new domestic vacancies. In this econometric specification, the trade openness also had a positive impact, indicating that domestic vacancies were located in those provinces more competitive.

Despite this is a preliminary exploratory analysis, we are able to identify some interesting results. First, the econometric results support the descriptive analysis: FDI and immigration inflows in Spain followed different patterns. Second, trade openness had a positive effect both on foreign and domestic new vacancies, but the effect was higher in the foreign vacancies. Finally, the new foreign vacancies were located in province-sector combinations with abundant medium-educated workers, while the domestic vacancies also were located in province-sector combinations with abundant low-educated workers. Therefore, foreign investors looked for more qualified workers than domestic
Table 2: Determinant factors in the new domestic vacancies

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>vacancy&lt;sub&gt;domestic&lt;/sub&gt;&lt;sup&gt;pt-1&lt;/sup&gt;</td>
<td>0.1223</td>
<td>0.1023</td>
<td>0.2395</td>
<td>0.1926</td>
</tr>
<tr>
<td></td>
<td>(0.0943)</td>
<td>(0.0983)</td>
<td>(0.1234)</td>
<td>(0.1179)</td>
</tr>
<tr>
<td>hiring&lt;sub&gt;natives&lt;/sub&gt;&lt;sup&gt;pt-1&lt;/sup&gt;</td>
<td>0.6124***</td>
<td>0.5301***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0807)</td>
<td>(0.0800)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hiring&lt;sub&gt;foreigners&lt;/sub&gt;&lt;sup&gt;pt-1&lt;/sup&gt;</td>
<td>0.3004***</td>
<td>0.1940**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0986)</td>
<td>(0.0867)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>human capital&lt;sub&gt;pt-1&lt;/sub&gt;</td>
<td>-12.4825</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(33.9091)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hiring&lt;sub&gt;highedu&lt;/sub&gt;&lt;sup&gt;pt-1&lt;/sup&gt;</td>
<td>-1.5509***</td>
<td>-1.3540***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.5172)</td>
<td>(0.4914)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hiring&lt;sub&gt;mededu&lt;/sub&gt;&lt;sup&gt;pt-1&lt;/sup&gt;</td>
<td>0.3712***</td>
<td>0.2575*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1420)</td>
<td>(0.1385)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hiring&lt;sub&gt;lowedu&lt;/sub&gt;&lt;sup&gt;pt-1&lt;/sup&gt;</td>
<td>0.2323***</td>
<td>0.1788***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0730)</td>
<td>(0.0616)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>trade openness&lt;sub&gt;pt-1&lt;/sub&gt;</td>
<td>26.2637</td>
<td>31.1720*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(17.3359)</td>
<td>(18.7066)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>firm closing&lt;sub&gt;pt-1&lt;/sub&gt;</td>
<td>-4.7434***</td>
<td>-6.4616***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.2884)</td>
<td>(1.5658)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>real GDPpc&lt;sub&gt;pt-1&lt;/sub&gt;</td>
<td>0.0085</td>
<td>0.0096</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0084)</td>
<td>(0.0087)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>%working-age&lt;sub&gt;natives&lt;/sub&gt;&lt;sup&gt;pt-1&lt;/sup&gt;</td>
<td>33.8553***</td>
<td>40.3538***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(12.0819)</td>
<td>(11.2585)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>%working-age&lt;sub&gt;foreigners&lt;/sub&gt;&lt;sup&gt;pt-1&lt;/sup&gt;</td>
<td>1.3790</td>
<td>5.1918</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.3222)</td>
<td>(6.2596)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant</td>
<td>-4.4230</td>
<td>-2364.9743***</td>
<td>40.1747</td>
<td>-3157.1420***</td>
</tr>
<tr>
<td></td>
<td>(17.2599)</td>
<td>(553.0124)</td>
<td>(32.9920)</td>
<td>(657.6317)</td>
</tr>
</tbody>
</table>

Fixed effects:
- origin-prov-sector: Yes, Yes, Yes, Yes
- year: Yes, Yes, Yes, Yes

Observations: 7,403
Number groups: 1,076
R-sq (within): 0.1199
R-sq (between): 0.7478
R-sq (overall): 0.5017
F: 51.5947
p-value: 0.0000
sigma_u: 204.1184
sigma_e: 315.3063
rho: 0.2953

Origing country-group (o), province (p), sector (s), year (t).
Standard errors clustered by province-sector are in parenthesis.
* p < 0.10, ** p < 0.05, *** p < 0.01
5 Conclusions

In this paper we investigate a selected FDI determinants in the Spanish provinces. In order to deal with this analysis we elaborate a novel database by merging information from two datasets: the SABI, which contains information about firms, and the MCVL, which contains information about workers.

We conduct a preliminary empirical analysis and we run a panel fixed effect model in order to detect the determinants of the new vacancies. We find that the foreign investors established in provinces and sectors more competitive and with higher human capital with respect to the domestic investors.

Furthermore, we detect that capital and immigration inflows in Spain followed different patterns. The FDI inflows came mainly from EU-15 and North America and was concentrated in services, while the immigrant inflows came mainly from Latin-America, Africa, and Rest Europe and was concentrated in low skill-sectors (construction and hotel).
References


## A Appendix

Table A.1: Economic activity classification

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>CNAE-93 Rev 1</th>
<th>CNAE 2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>Food, beverages, and tobacco</td>
<td>15 16</td>
<td>10 11 12</td>
</tr>
<tr>
<td>200</td>
<td>Textiles, leather, and wood</td>
<td>17 18 19 20 36</td>
<td>13 14 15 16 31 32</td>
</tr>
<tr>
<td>300</td>
<td>Paper and edition</td>
<td>21 22</td>
<td>17 18 58</td>
</tr>
<tr>
<td>400</td>
<td>Chemical, plastic, and petroleum transformation</td>
<td>23 24 25 26</td>
<td>19 20 21 22 23</td>
</tr>
<tr>
<td>500</td>
<td>Metallurgy and mechanical equipment manufacture</td>
<td>27 28 29</td>
<td>24 25 28</td>
</tr>
<tr>
<td>600</td>
<td>Electrical machinery, computer system, and medical instruments</td>
<td>30 31 32 33</td>
<td>26 27</td>
</tr>
<tr>
<td>700</td>
<td>Automotive</td>
<td>34 35</td>
<td>29 30</td>
</tr>
<tr>
<td>800</td>
<td>Energy</td>
<td>40 41</td>
<td>35 36 37 38 39</td>
</tr>
<tr>
<td>900</td>
<td>Construction</td>
<td>45</td>
<td>41 42 43</td>
</tr>
<tr>
<td>1000</td>
<td>Wholesale, retail sale, and vehicle motor repairs</td>
<td>50 51 52</td>
<td>45 46 47</td>
</tr>
<tr>
<td>1100</td>
<td>Hotel</td>
<td>55</td>
<td>55 56</td>
</tr>
<tr>
<td>1200</td>
<td>Transport</td>
<td>60 61 62 63</td>
<td>49 50 51 52</td>
</tr>
<tr>
<td>1300</td>
<td>Telecommunication</td>
<td>64</td>
<td>53 61</td>
</tr>
<tr>
<td>1400</td>
<td>Financial activities</td>
<td>65 66 67</td>
<td>64 65 66</td>
</tr>
<tr>
<td>1500</td>
<td>Real estate activities</td>
<td>70</td>
<td>68</td>
</tr>
<tr>
<td>1600</td>
<td>Renting</td>
<td>71</td>
<td>77</td>
</tr>
<tr>
<td>1700</td>
<td>IT and computer services</td>
<td>72</td>
<td>62 63 33</td>
</tr>
<tr>
<td>1800</td>
<td>R</td>
<td>73</td>
<td>72</td>
</tr>
<tr>
<td>1900</td>
<td>Administrative and support activities</td>
<td>74</td>
<td>69 70 71 74 78 82</td>
</tr>
<tr>
<td>2000</td>
<td>Public administration</td>
<td>75</td>
<td>84</td>
</tr>
<tr>
<td>2100</td>
<td>Education</td>
<td>80</td>
<td>85</td>
</tr>
<tr>
<td>2200</td>
<td>Services (health, leisure, sports, culture)</td>
<td>85 90 91 92 93</td>
<td>86 87 88 90 91 92 93 94 95 96</td>
</tr>
</tbody>
</table>

Codes from 100 to 800 refer to industry, code 900 to construction, and codes from 1000 to 2300 to services. The primary sector is excluded.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>New vacancies</td>
<td>Number of employees of new firms, province, and sector.</td>
<td>SABI, Bureau van Dijk</td>
</tr>
<tr>
<td>FDI inflows</td>
<td>Vacancies created by new firms that fulfill at least one of the following conditions: (i) belonging to a foreign enterprise with at least 10 percent of the capital and/or (ii) the parent company is located abroad. This variable varies by origin, province, and sector.</td>
<td>SABI, Bureau van Dijk</td>
</tr>
<tr>
<td>New hiring</td>
<td>People entering the Spanish job market. This variable varies by origin, province, and sector.</td>
<td>MCVL, Spanish Social Security</td>
</tr>
<tr>
<td>Immigrant inflows</td>
<td>New workers born in a foreign country. This variable varies by origin, province, and sector.</td>
<td>MCVL, Spanish Social Security</td>
</tr>
<tr>
<td>GDP / GDP per capita</td>
<td>The INE's database contains information about the GDP in current euros (Contabilidad regional de España). Using a price index provided by the Penn Word Table (pwt), nominal values are converted into real ones. We compute the GDP per capita using information about population, which is also available at the INE's database (Cifras de Población). This variable varies by province.</td>
<td>INEbase, INE; Penn Word Table (pwt)</td>
</tr>
<tr>
<td>Human capital</td>
<td>Average years of education of the working-age population. This variable varies by province.</td>
<td>Fundación Bancaja and ICE</td>
</tr>
<tr>
<td>Trade openness</td>
<td>DataComex provides information about trade flows (exports and imports) by different (foreign) origin country-groups. Using the GDP, we calculate an index of trade openness as trade flows over GDP. This variable varies by origin and sector.</td>
<td>DataComex; INEbase, INE</td>
</tr>
<tr>
<td>Risk</td>
<td>Number of firms that shut down. This variable varies by province.</td>
<td>SABI, Bureau van Dijk</td>
</tr>
<tr>
<td>Working age</td>
<td>Data about population are extracted from the INE's database Cifras de Población. We distinguish between working-age population as people aged between 15 and 64. We calculate the working age population as people aged between 15 and 64.</td>
<td>INEbase, INE</td>
</tr>
</tbody>
</table>