Determinants of international migration to the Spanish Regions.

Author: Juan David Gil González

al364421@uji.es

EC1049 – Final Degree Project

Degree in Economics, course: 2020/2021

Dra. Inmaculada Martínez Zarzoso

Abstract:

This dissertation explains using a log-log panel data model, why a regular immigrant tends to choose a given Spanish region as destination. The major factors considered in host communities are related to policy, economic conditions, subsidies, language, and safety. I use a sample of 53 countries worldwide with data extracted from organizations and institutions such as The World Bank and the Statics National Institute from Spain. The study is relevant because in Spain there are many foreign-born citizens gathered in different areas of the country. I focus the study on annual movements of migrants in the two last decades, focused to explain flows through empirical analysis, using a panel data study using the software Stata. If the main results indicate that inflows were only explained by economic reasons, all immigrants would be in Madrid, Catalonia or the Bask country. The reality is nevertheless more complex, since: there are many Romanians working in Valencia, or many Portuguese in Galicia.

Keywords: income per capìta, panel data, international migration, hypothesis, European Union, Spanish regions.

Table of contents.

Graphs, tables, figures, and equations.

1 Introduction.

The study about foreign international migration is an important topic which affects several aspects of daily life for anyone, and surely touches aspects of the person that is presently reading this line. In the economic science, a single change with an immigration policy could affect any aspect or situation which was potentially economically studied. Thus, the GDP (Gross domestic product), population, welfare, employment rate, crime, or public spending are topics that could be affected in a positive or negative way. To clarify and adding value to the prior knowledge about immigration to Spain, the paper has the ambition to contribute the public administration, government, or competent institution to make better decisions and laws that affect immigration between autonomous communities. If politics elaborate laws given the empirical science existing about each topic, social welfare would be maximized for all the society.

First, to give the lector tools for understanding the topic this paper is going to talk about, it is needed to make a brief introduction about the subject, not only it is necessary acknowledge the previous authors which studied and wrote the Spanish patterns of immigration. This topic needs a background explanation of the available and more complex theory about general international migration. Therefore, firstly papers about causes of international migrations are going to be explained briefly, in addition to then putting the focus on Spain.

According to a reputed Australian university professor, international migration is one of the central parts that explain a bigger issue: globalisation. This phenomenon explains an enormous expansion of connection flows between multiple economic agents around the globe. These agents span every aspect of our life and cooperation in society and converge on a sort of ‘’*transnational network*’’ (Castles, 2000). The reputed sociologist Manuel Castells also writes about negative and positive aspects of globalisation. He approaches to this ‘’network society’’ a continuous struggle between different nation-state which only ends on exclusion, or social, technological, and economic inequality (Castells, March 2009). Equally, an international division of the labour whose achievements were merit of the States, finally are gone for the public sector, worsening the persistent inequality. Moreover, Manuel Castells describes that this quick process of globalisation is caused by six points process through a huge change of mentality and economical behaviour due to the blast of the financial crisis in 2008.

Another interesting, reputed, and empirical point of view about causes that explain international immigration is the one of Karen O’Reilly (2012), this author

approach starts talking about the existence of immigrants based on the existence of nations. Then, for the passage of 2 centuries and change of trends, crowds of available humans beaten for economic or social ravage are waiting the opportunity to cross borders (O'Reilly, 2012).

The main objective of this paper is to contribute to the economics science through the study of literature migration and confirm hypothesis modelling panel data with Stata program. This is going to be reached by confirming through statistical inference, two principal hypothesis (one economical and another social), each with two components, these are:

1. Well-off and poor immigrants tend to emigrate for similar reasons to the autonomous communities.
    i) Autonomous Communities with high rates of wealth growth tend to receive more people.
    ii) Rural and urban development also affects destination.

2. Refugees and immigrant fluency depends also on non-economic factors.
    i) Safety and wellness ratios by destination and origin influence the target autonomous community.
    ii) Technological development of origin country affects international arrivals to Spanish territory.

2 Development: International migration.

After the brief explanation made before on some approaches to the causes of migration, this section will explain what is migration, international migration, and Spanish migration, both inside and out. The ultimate interest (as indicated in the conclusion) in this paper is to determine which specific characteristics an immigrant who chooses the same communities shares. This phenomenon will be explained from the viewpoint of several authors and approaches of different models, like the gravitational or neoclassical model.

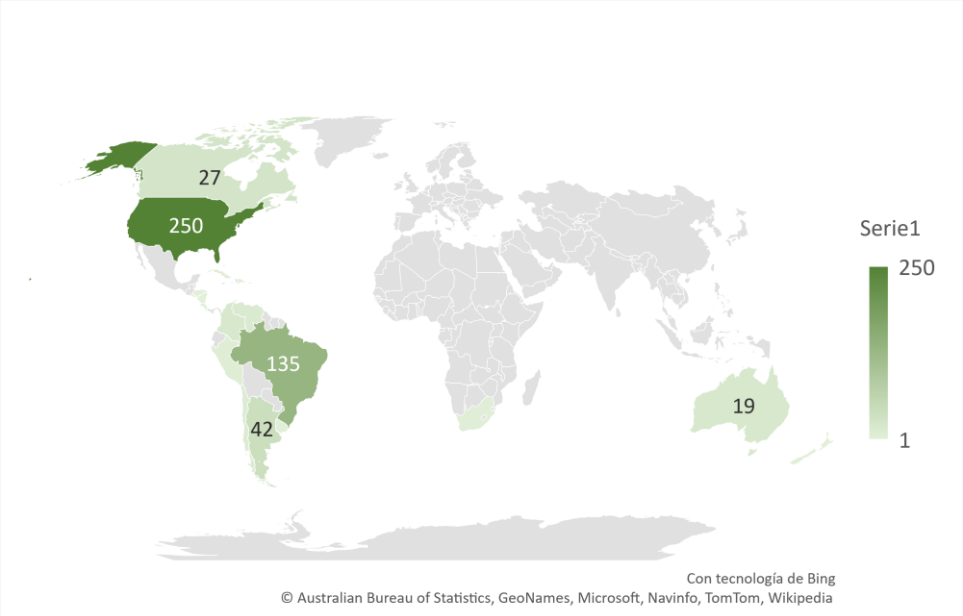
From a historical point of view, Roselló V.C.(2008) exposes the first massive migrations of people that were produced at the beginning of the XIX century, in a great migration that would move the current population of a country like Italy (60 million) from

Europe to America. In that case, it was mainly due to economic causes and to escape the poverty that in Europe would leave the two world wars. At that time, the United States have been already a benchmark for emigration, but not the only one, since countries like Argentina, Brazil, or Canada were very prosperous places full of wealth and resources. In Argentina, the foreign population reached 30% of the total population. It is noteworthy to know that the first constant migrations, although not massive, are due to the conquest of America by Spain in 1492 and the subsequent constant exchange of products and population. In this context, the cause of trade as a key factor in population exchange is significant (Roselló, 2008).

Later, the neoclassical school developed theories of general and partial equilibrium, and currently in all universities the international movement is studied thanks to the Heckscher-Ohlin-Samuelson model as a global equilibrium between productive factors (Oyarzun, 2008). The next concept map shows the number of populations in millions of European raids in the main host countries during the first large wave of migration. Because travel a century ago was very costly in proportion to people's wages, the difference is explained by prices and wages (Ravenstein, 1889). The old economic theory of prices and wages is outdated, but useful for understanding old perspectives.

Figure 1

First wave of migrants from Europe, 1820-1930



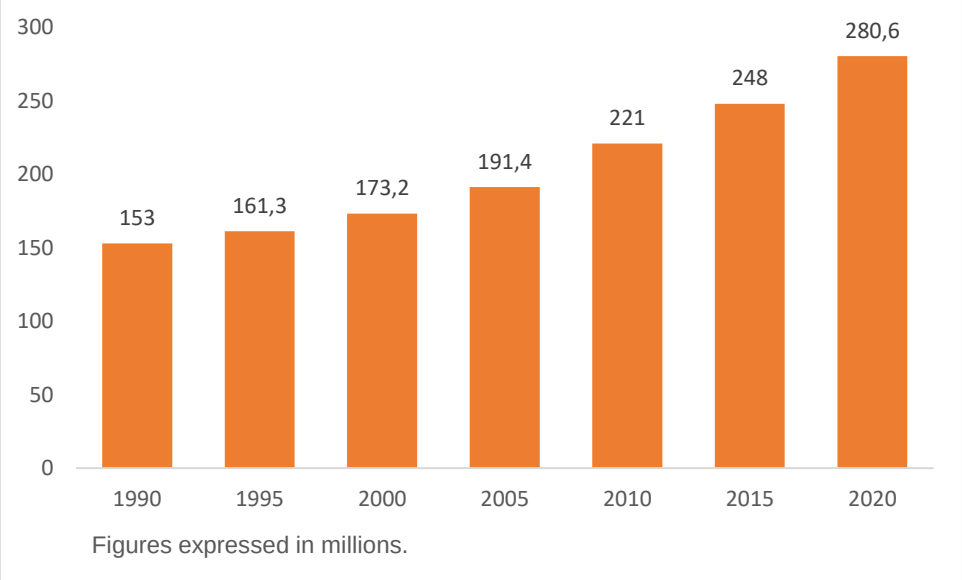
Source: Own elaboration with data collected from: https://migrationdataportal.org/

Immigration is a term and domain of the social sciences used to study a vital historical and human event since the first existence of homo sapiens. That is the study concept used when we want to understand why there are movements of individuals that are looking for a new place to live, different from the place they were born. Today, an immigrant is someone who moves to live or work (or both), in another political state, necessarily crossing a boundary configured by land, sea or air. The motivation to start studying about the flows of people coming to Spain is due to the importance they have on the economy and the configuration of society.

On the next graph, an evolution of world immigrant population is settled. This is the overall number of people living in a country where they were not born. Reasons of this are almost clear for the economic study, and it is interesting to acknowledge why this is happening with the whole world, because this work will aim to Spanish environment of migration from global to specific context.

Chart 1

International migrant population



Source: Own elaboration with data collected from: https://migrationdataportal.org/

As we have already seen, each year there is a steady increase in the number of international persons living in another country of origin. By putting this chart into perspective, now foreigners are not the majority, but grow with force each year. According to Castles (2010), economic reasons are significant for emigration as new sociological models. While there are many papers or studies to explain why people migrate, there is still a gap between reality and empirical evidence. What is true is that people's behavioural mobility is now far removed from the social structure of the Cold

War. Today, social transformation is more important than ever for such an event (Castles, 2010).

Below, it is shown a comparison between the importance of foreign population when the context is the whole world, or the framework is Spain. The difference here is because Spain for the past 20 years, has become in a host country, not a sending one. In 1970, 12.9% of people living in Spain coming from other countries, was only the 0,43%, namely thirty times less. That situation was not always like that, from 1885 to 1936 Spain was sending each year hundred thousand inhabitants abroad, mainly to Latin America (Montserrat Casado, 2003).

Chart 2

Proportion of world migrant population, 2020(%).

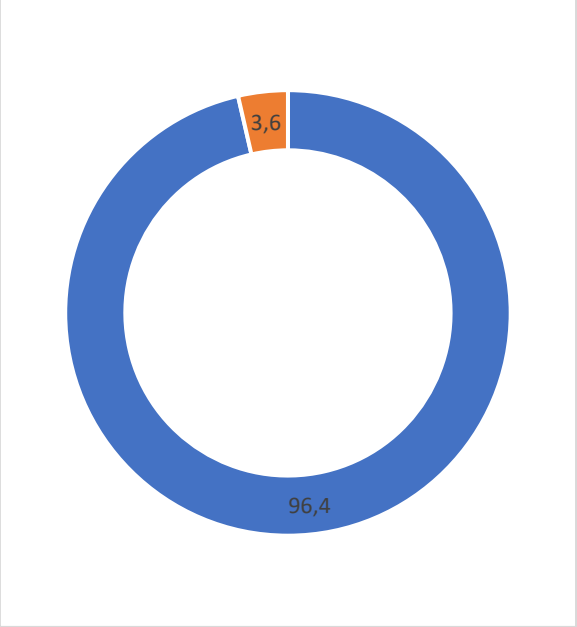
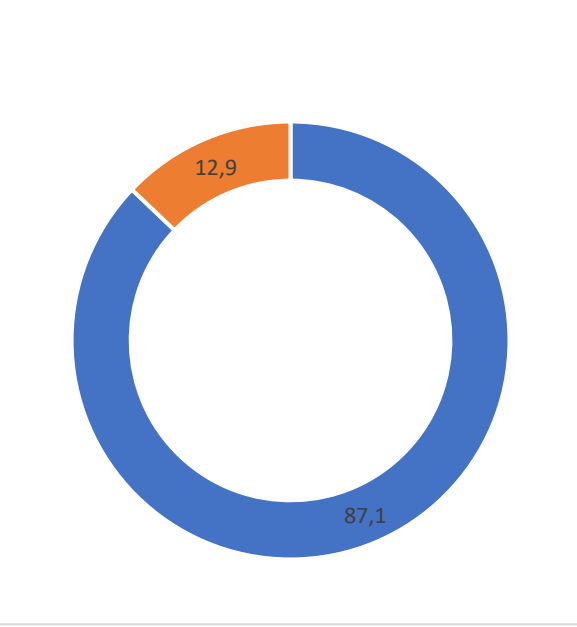


Chart 3

Proportion of foreign people in Spain, 2020.(%)



Source: Image 3 and 4 elaborated with data collected from: https://migrationdataportal.org

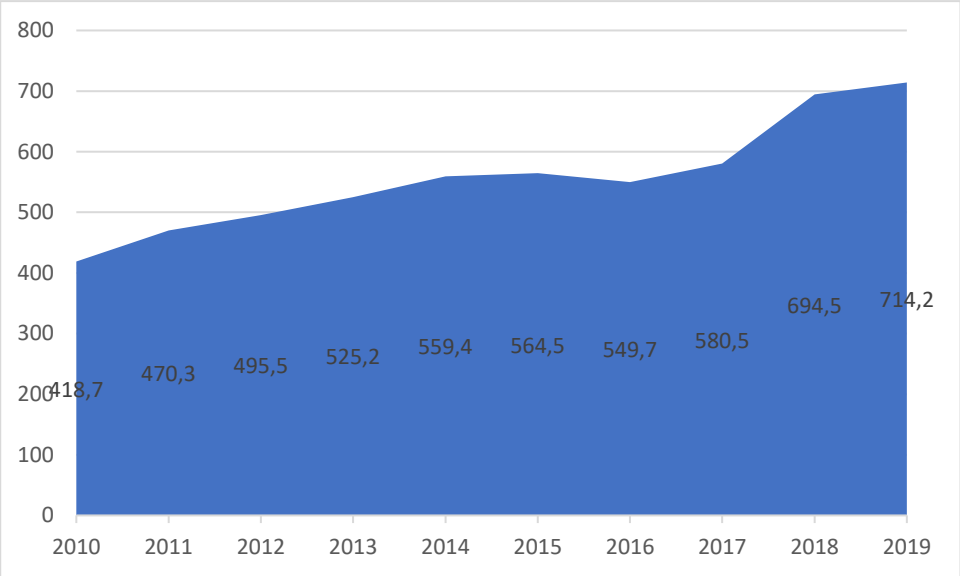
In this case, what is interesting for understand patterns of location in the Spanish case, is that there is an enormous importance of immigration, more than a standard country, and almost the seventy-five percent of they, are settled in Andalucia, Islas Canarias o Comunidad Valenciana, namely, very concentrated in three of the seventeen autonomous communities which shape Spain (Montserrat Casado, 2003). These characteristics will be explained later. Now, the analysis will follow arguing and approaching the causes through which there are massive immigration flows.

If lector look at another for causes that drive international migration, it is shown that international shipments to help the family, move so much money. If consignments

of 2019 are added up, would be as big an economy as Taiwan or Poland (Portal de datos mundiales sobre la migración, 2021). Then, the worldwide development of this metric:

Chart 4

International consignments received in billions, USD($)



Source: Own elaboration with data collected from: https://migrationdataportal.org

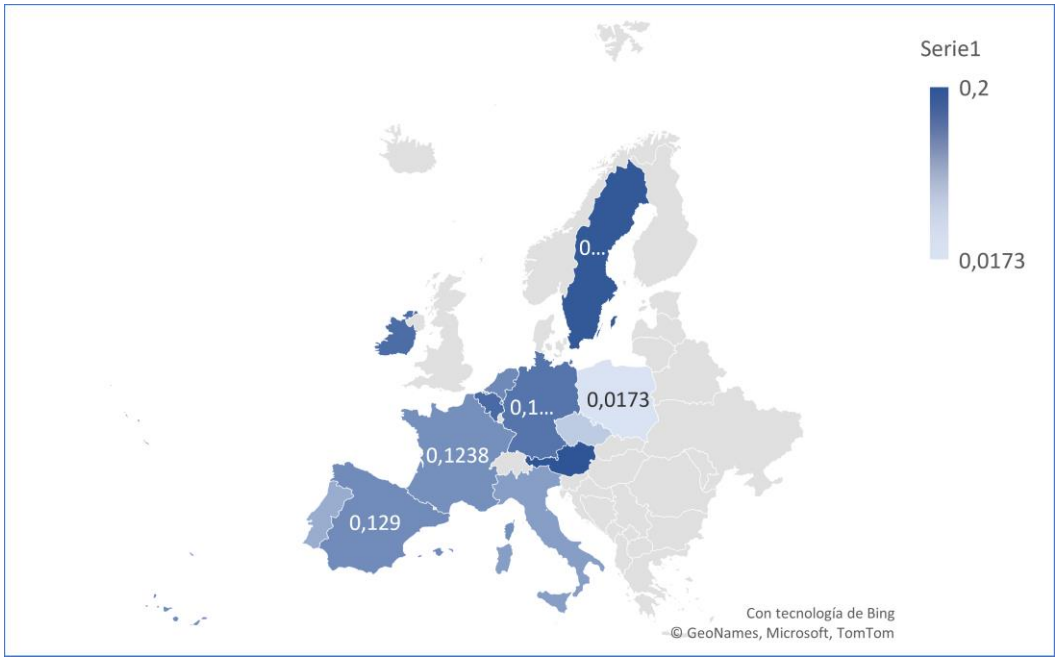
Above, Over the last year, the file of remittances sent to relatives or friends of migrants worldwide has been registered. The huge amount of money sent would express another important cause to explain the international movements of persons. Alma Rosa (2006) delves into the importance of remittances for development and, through an empirical study, argues that sending dollars manages to remit the deficit of nations, helps local economies, and finances the development of poorer countries. Specifically, each $ 1 would generate a wealth of $ 2.9 in the destination country, demonstrating that remittances have a high elasticity, and may partly explain the causes of international movements of people (Jumilla, 2006).

3 The exceptional European migration framework.

Before speaking more in depth about migration in Spain, the migration framework that governs the European Union will be briefly discussed, since to understand the large population flows that arrive in Spain, we will have to understand which borders operate in Europe and why it makes so unique in the world. This step is also intended to give the reader a vision progressively from the most general to the most concrete.

Figure 2

Percentage of immigrant population over the total with a color gradient.



Source: own elaboration with data collected from: www.datosmacro.com

Note: n= principal 15 european economies.

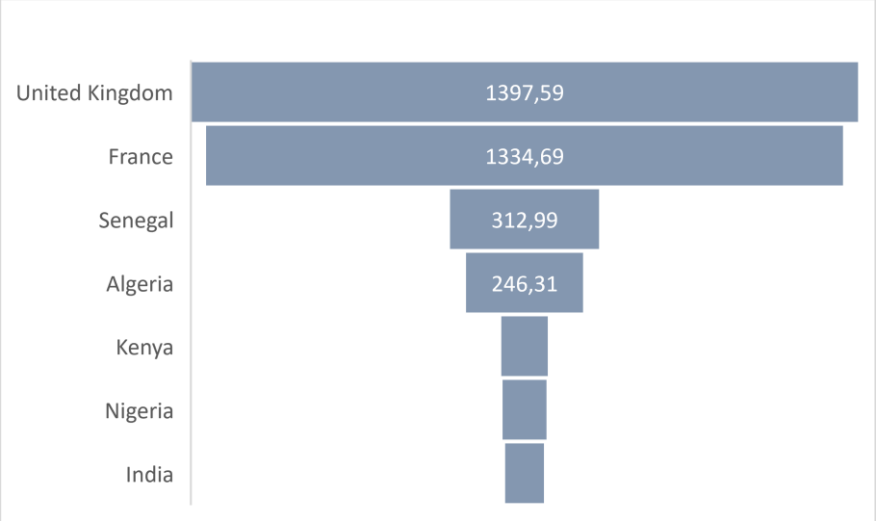
The European Union is an international union of a group of countries that share laws, regulatory frameworks, political institutions and even the monetary means of economic exchange, that is, the euro. To understand and study the entry of migration to Spain, it is essential to understand European laws, which are unique in the world, and allow freedom of movement and trade never seen before between different nation-states. Centuries and decades ago, the Roman Empire, Napoleon or Hitler, attempted to unite Europe under the pretext of military unions, based on identity or race. However, it was only in 1993, with the founding Treaty of the European Union, that this union took effect. Even though European rulers have always been very concerned about their national sovereignty, the search for a better future for all and, above all, the spectre of two world

wars that devastated Europe, fostered a union at all levels (Dinan D, 2004). Following the First World War in 1925, a French Prime Minister proposed a union at the European level after the First World War. Thus, in 1944 and already putting an end to the devastating Second World War, in the United Kingdom, the Federal Commission proposed a union in Europe to prevent another war from taking place again. Moreover, in 1948, the Congress of Europe met with hundreds of intellectuals from the principal countries of power.

In Spain there are many immigrants from Europe, partly explained by the proximity and by the same regulatory framework for the movement of people that exists between all the countries that make up the EU. Likewise, although it seems that free movement between people within Europe is out of date, according to Givens T., and Luedtke, A. (2005), this continues to be one of the main concerns of Europeans, and governments of both the left and the right, do not promote the improvement of the situation.

Chart 5

GDP per capita in 1960, USD ($).



Source: Own elaboration with data collected from Wikipedia.

Before the unification of frontiers, the major powers were historically restrictive on immigrants. After the independence of the French and British colonies, poor countries like India or Algeria sent a lot of workers to their former metropolises, or after 1991 with the fall of the Soviet Union, Germany faced the challenge of hosting hundreds of thousands of workers, from ex-Soviet countries, in collapse due to the economic crisis. In these cases, the observed causes of migratory movements between the former colonies and their respective invader countries are clearly favored by economic causes. These immigrants have the ease of knowing the language of the target country and are

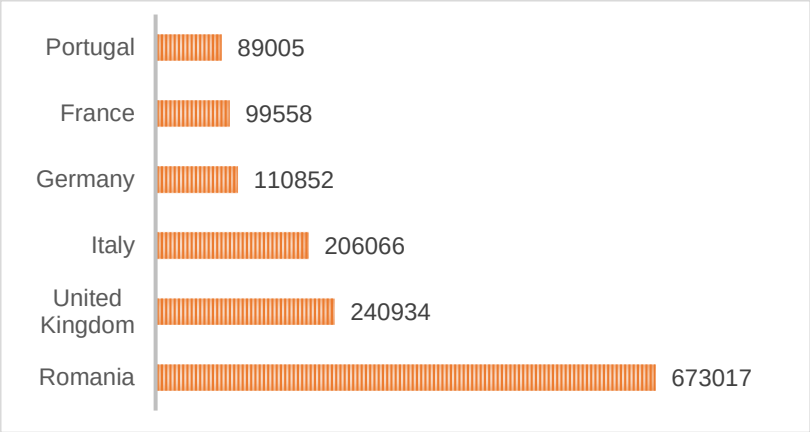
also motivated by the search for a better future, where wages are much higher. Although today the differences between the per capita wealth of the countries that made up the former colonial empires and the former colonies are still high, it is enough to see the disproportion that existed in the decade from 1960 to 1970, which is when most colonies achieve their independence, to understand people's motivation to emigrate for a better life. The chart below, with data obtained from the World Bank (https://www.worldbank.org/fr/home), graphically shows the differences in wealth generated per capita:

As it is shown, Although the French colonies appearing in the graph (Senegal and Algeria) enjoyed greater wealth than the English colonies and closer to the GDP of their metropolis, it is true that the disproportion continues to be 1 to 4, that is, at the time of independence France was 4 times richer. If we look at the countries invaded by the British Empire, the three were around 90 international dollars in annual output per capita. The fact that the United Kingdom was 15 times richer than its former colonies would explain the rise in immigration and would give way to a mentality less open to free borders that would later make negotiations difficult to open confines because of the founding Treaty of the European Union (Roselló, 2008).

We know that an institutional framework exists in the European Union which facilitates the exchange of population for reasons of study, work or tourism. Freedom of trade and freedom of movement established by European treaties are the main reason why millions of people on the continent continually cross borders (Oyarzun, 2008). There are clear causes of work, commerce and, above all: tourism. The data presented below show that a large proportion of Europeans who come to Spain do so due to tourism. To get to the part where we will expose the population movements in Spain, we will show the number of Europeans established in the country.

Chart 6

Top 6 European population living in Spain, 2020.



Source: Own elaboration with data from *INE.*

4 2015 refugee crisis

In this work, the refugee variable has been introduced, therefore in this section the importance this variable has from the year 2015, will be briefly contextualized. According to the UN, a refugee is a person forcibly displaced from their country of origin due to persecution and does not find protection in their country. Since the second world war until 2015, there had not been a refugee crisis so high in number of forcibly mobilized, and this has mostly affected Europe.

Europe was bordering on a diplomatic crisis that did not suffer since the twentieth century. Since the 2015 crisis, 45 years passed when free migration was closed, many controls were established because of the oil crisis. The Dublin Regulation was not prepared for this crisis, and alongside the Arab springs, Europe has received millions of refugees in just a few years (Enríquez, 2015). Finishing 2015, according to the European authorities, one million of refugees did get in the EU. The fundamental cause of this wave of migration has been preceded in previous years by the outbreak of the Arab springs in 2010 with greater or lesser intensity, in all the Arab countries that have access to the Mediterranean Sea. This revolutionary situation began to strain the borders of Europe, but the event that seriously challenged the EU was in 2011, when the revolution reached Syria and a civil war broke out.

Even though the migration crisis caused is highly focused from the perspective of European countries, Syria's neighbouring countries also suffered many economic and social consequences. For example, Lebanon is located at its closest point to 50 minutes

by car from the Syrian capital, Damascus. According to Benedetta Berti (2015), Lebanon's tourism and services sector was severely affected, causing massive job losses and widespread price increases, events that moved almost 200,000 Lebanese from the middle class to the lower / poor class.

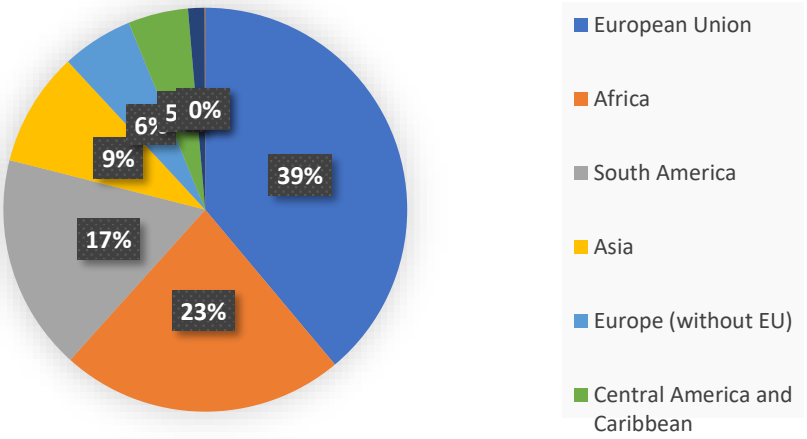
Although Spain is part of the European regulatory framework, it was not as affected by this crisis as other European countries. Germany was the country that received the most refugees, and according to Eurostat data, together with Sweden, France, and Italy they received about 70% of all asylum seekers. For this reason, it is considered interesting how the refugee variable has affected the study since it is analysed from the perspective of the country of origin, which does include countries such as Germany that are highly affected by this migratory wave. Next, after briefly explaining what the 2015 refugee crisis was, the theoretical analysis will be made based on the characteristics of immigration in Spain.

5 Immigration in Spain.

Immigration in Spain stands out for being almost 12% of the population. It is a country where thousands of people come to live for many years and millions to work.

Figure 3

Proportion of foreign population living in Spain by areas of the world (%).



Source: Own elaboration with data collected from *INE.*

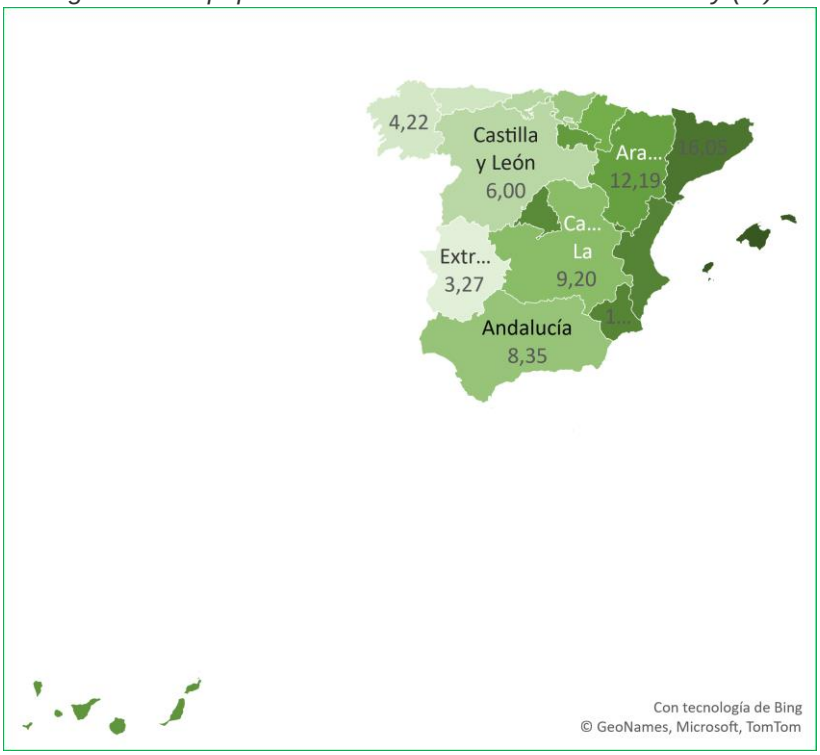
As can be seen in figure 3, in Spain the main origin of its foreign inhabitants is from the European Union. From that nearby area of the world, they are 39% of foreigners, clearly due to the internal freedom of movement that exists in the EU, thanks to the Schengen Treaty, signed in 1985 in Luxembourg. History of migration in Spain is broad,

recently closely linked to Europe by geographical proximity and legal freedom of movement but being very constant with America since the discovery in 1492. For example, in Spain in 1998 there were 130,203 people from Latin America with residence license. In 2008 the number reached more than one million: 1,333,886, that is, ten times more than ten years before (García, Jiménez and Redondo, 2009).

It should be noted that although the causes of external migration to Spain are not fully delimited, there are studies such as Silvestre (2005) that show clear causes of internal movements. These causes are mainly economic and due to differences in wages between regions, with a large increase at the beginning of the 20th century due to the strong growth that was registered in the Spanish economy was driven by sectors of the economy with low intellectual qualification, such the industry. Next, it will also be demonstrated if the causes between external and internal migrations converge.

Figure 4

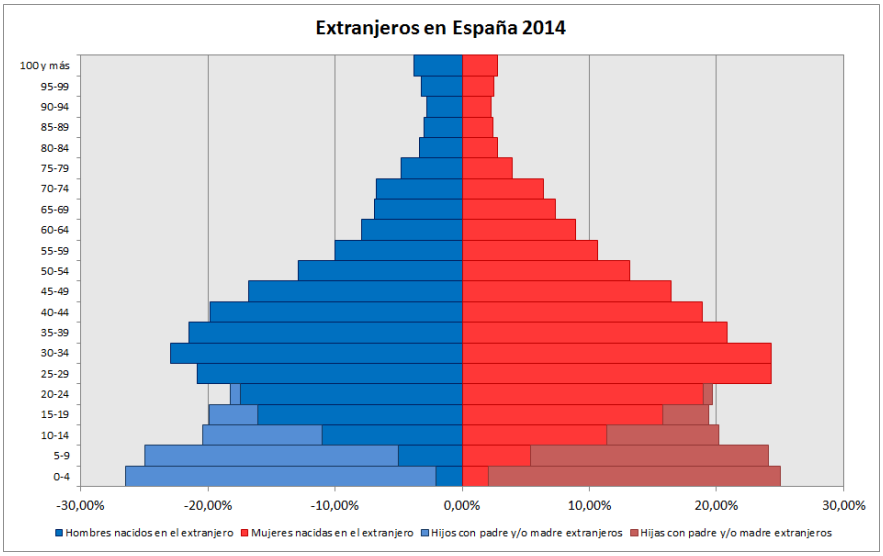
Foreign resident population for each autonomous community (%).



Source: Own elaboration with data collected from *INE.*

Figure 5

Population pyramid of foreign inhabitants in Spain, 2014.



Source: Wikipedia (free license)

In the figure 4 above it is observed the importance of the foreign inhabitants by region. In the case of Madrid or Catalonia, the mobility hypothesis for economic reasons would be confirmed, since they are two rich regions with a high percentage of migrants. However, other regions like Valencia, Murcia, or the Balearic Islands, also have many migrants, but they are economically less developed regions. Figure 5 shows the youth of the foreign population in Spain, in contrast to the local population, with an average population of 44.9 years in 2020. Between the decade of the 90s and the beginning of the 21st century, the positive contribution of immigration contributed to more than half of the newly created jobs were held by foreigners (Moreno, 2012).

In fact, the hypothesis that the economy is not a decisive factor for emigrating to Spain, is stated by Karen O'Reilly (2000) in her paper with subjective assertions of British emigrants living here. The interviewees describe Spain as a country with many leisure opportunities, where “you are not cast aside when you are old”, while “England is so depressing these days. It’s cold and grey, and everyone is miserable” (O’Reilly, 2000, p.6). However, this work is 21 years old, therefore in the following pages it will be shown if the reality remains the same, and if it is just as likely to find English people in Spain nowadays who affirm the same.

6 Econometric model.

6.1 Variables and data.

In this section, a panel data model with three dimensions will be developed to confirm or reject the hypotheses raised at the beginning and through which it is intended to establish the differences between migrants who choose one autonomous community over another. The final model to develop the model and extract causes and conclusions to migratory flows will be log-log and throughout the following pages the reason for the decision and the advantage of working with this model for these data specifically over others will be explained. The dependent and main variable of the model is the logarithm of immigration, but the control variables can independently express characteristics of countries of origin, or autonomous communities of arrival.

The data is composed of 3 dimensions: autonomous communities of Spain, country of origin of immigrants and years. In the first dimension, the 17 official Spanish autonomous communities and the 2 autonomous cities are represented. The second dimension includes the 53 main countries worldwide which send immigration flows to Spain. To sum up, the third dimension represents the temporal data whose scope is from 2008 to 2019. Total number of observations are 12,085, although abated in some variables because of missing values.

The data that express the reality of the autonomous communities it's extracted from the National Institute of Statistics from Spain (INE), and the data that correspond to the context of each of the 53 collected countries annexed to the timeline that has been indicated, they have been extracted from the World Bank database. There is also a multitude of data and information used to create graphs and figures, taken from the Global Migration Data Portal (https://migrationdataportal.org/).

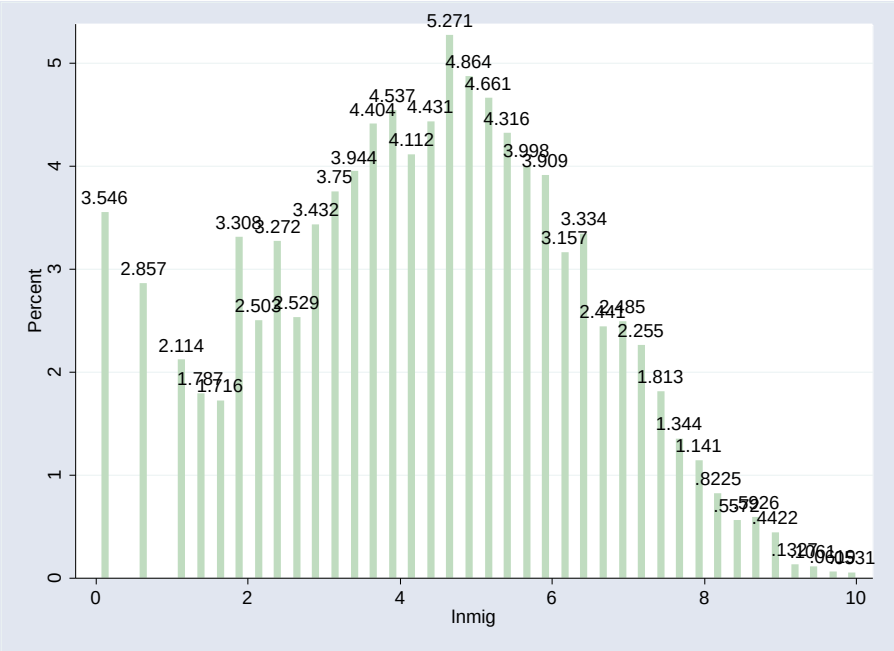
6.1.1 Dependent variable.

The dependent variable and that we want to explain with the model, are international arrivals by country, segregated by each autonomous community that constitutes Spain. This variable is expressed in units, worked in panel data. It is modulated by autonomous communities (19), countries of sending migrants (53) and the years in which the events analysed occur, from 2008 to 2019. Because the model has as an explained variable the annual arrival of immigrants segregated by autonomous communities, international arrivals in the final model will be explained as logarithms, that

is, as their percentage evolution due to the advance of a certain period, or account of the interaction than the other independent variables included in the model.

Chart 7

Relative distribution of the annual logarithm of migrations.



Source: Own sample of the model

Chart 7 could show a first impression, in which it is confirmed at first glance that the dependent variable treated follows a normal distribution. Visually, the sample distribution is a Gaussian bell, a property needed to do statistical work with a high number of N population. According to probability theory, most of the data sets associated with morphological, psychological, or economic phenomena should collect most of the data in the central part of the distribution, with a low deviation towards the extremes (Pértegas and Fernández, 2001). This would indicate that the variables of a normal distribution are grouped around the mean, and that rare cases are found at the extremes where occur with less likelihood.

Equally, chart 7 expresses on the Y axis the relative importance (in percentages) of the values seen on the X axis. For example, in the middle of the distribution, between what would be a percentage increase between 4 and 6 per year in immigrants by autonomous communities, there is a 5.27% of the sample. Moving away from this value, forward or backward, the importance gradually decreases, although it declines faster when we look for ever higher percentages of increased immigration flows.

To sum up, there are slightly more values of increase that move away from the mean towards smaller values than towards higher values. In fact, immediately after a 0%

increase in international annual entries, there is 3.5% of the sample, which decreases, and when it reaches the value 2 (+ 2%), the values increase until reaching the sample mean. We could conclude with the information provided by the graph, that there are more very small annual increases in migrant flows (close to + 0%) than very large annual increases (close to 10%). This is because the decision of large population groups to move or not is not made just by willingness but is caused in part by the ease or difficulty they must undertake the trip, based on their previous material conditions (Castles, 2010).

6.1.2 Independent variables.

The control and independent variables that shape the model are made up of 5 variables expressed in logarithms, and another 3 linear ones. The variables and the reason for choosing them will be explained below, followed by the analysis using tables created in Stata. In table 1 of correlations shown below, the degree from 0 to 1 (positive or negative) of a priori relationship between all the variables that make up the model is expressed. Firstly, the meaning of each variable is explained:

1. *lnmig*: Dependent variable of the model and already explained, which expresses the logarithm in annual terms of immigrants for each autonomous community.

2. lngdpco: Logarithm of the per capita GDP of the migrants' countries of origin. It has been chosen since it is a very important factor in economic science to express effects on the real well-being of citizens of one policy over another. Mathematically it expresses the annual production in dollars of goods and services, divided by the population of a country.

3. *lngdpCCAA*: Logarithm of the GDP per capita of the Spanish autonomous communities. It is a variable that can be key to modelling whether the factor that migrants choose to reach one territory over the other is mainly economic. Data is extracted from the National Institute of Statistics (INE), and it is important to know because could be a influencing factor of destiny for immigration.

4. *lnunemd :* Logarithm of unemployment rate of each autonomous community. From the point of view of a migrant, this variable is also based on economic reasons for moving away.

5. *lngdpgrowth:* This is a factor that modulates the destination of the migration site. It is the logarithm of the evolution of the per capita GDP of the autonomous communities. Faced with an evolution of + 1% in this factor, international entries would evolve by X% because of the model. If this factor is ultimately decisive, it would

be interpreted that the autonomous communities to emigrate are chosen in part because of their economic situation.

6. *lnrefugees:* The logarithm of refugee arrivals is important for the paper and serves to solve one of the hypotheses of the model. Although it should be noted that this variable includes refugee asylums in the 53 countries of origin of the migrants heading to Spain. Therefore, it is significant to study the effect of the saturation of refugees from the countries of origin when they are preparing the arrival to Spain.

7. *lnrandd*: This variable expresses the annual expenditure per country on research and development as a percentage of GDP. It is also expressed in logarithms and helps to determine or approximate the level of technological development of a country, in this case the country of origin. For now, one might think that the higher spending on research and development in a country, the less incentive to emigrate. This hypothesis would be confirmed if it is assumed that technological advance is accompanied by social / economic advance.

8. *ruralt*: This variable corresponds to the logarithm of the annual percentage of the urban population by country of origin. It is an interesting variable to study since it conditions the economic and social characteristics of potential emigrants. The rural areas of the countries tend to be underdeveloped compared to the urban centers, and even, belonging to an urban, a person may be more willing to emigrate abroad due to multiple social factors that are impossible to fully analyze in their entirety. Later, using the model, it will be concluded if in this case, a higher proportion of the urban (the opposite will be rural) population favors emigration, and if it is a significant cause for an immigrant to move on.

9. *ev:* This variable is expressed in absolute numbers and indicates the average life expectancy of the total inhabitants of a country of origin. Among the countries with the highest life expectancies are, for example, Spain, Japan or Italy. We include it in the model because the expectations of quality of life in country X are likely to be important when deciding whether an emigrant will settle in one autonomous community or another. If life expectancy is significant for the model, it will be understood that a person who chooses one community over another to befriend, will do so in part or totally as the objective of improving their general quality of life.

10. *ch*: Finally, to analyze causality between the state of the social system and well-being, a variable that analyzes an aspect of health in the countries of origin has been

included: the number of hospital beds per thousand inhabitants. To analyze causes of emigration due to general aspects of well-being towards the Spanish autonomous communities, we have considered that it is interesting to include a factor that is a symptom of whether a health system works well or not. According to international organizations and studies carried out each year, Spain has one of the best health systems in the world. In case of confirming the positive significance of this variable, we would affirm that the good state of the Spanish health system has an attractive effect.

6.2 Descriptive analysis.

In this section, a descriptive analysis of the variables in question will be carried out, also mentioning other interesting variables for the model but which have not been included for various reasons. A descriptive analysis is a detailed exposition of the sample that will be used to work with the model, as well as the explanation of interesting aspects that can help the reader to understand about the final conclusions.

Table 1

Correlation matrix of the model.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) lnmig	1.000									
(2) lngdpco	-0.134	1.000								
(3) lngdpCCAA	0.206	0.020	1.000							
(4) lnunemd	-0.051	0.009	-0.722	1.000						
(5) lngdpgrowth	0.018	-0.458	0.005	-0.008	1.000					
(6) lnrefugees	-0.045	0.343	0.010	-0.016	-0.142	1.000				
(7) lnrandd	-0.119	0.734	0.021	-0.029	-0.347	0.655	1.000			
(8) ruralt	-0.065	0.688	0.010	0.026	-0.388	-0.018	0.324	1.000		
(9) ev	-0.038	0.872	0.021	0.036	-0.418	0.216	0.566	0.656	1.000	
(10) ch	0.007	0.277	0.005	-0.041	-0.249	0.066	0.377	0.146	0.122	1.000

Source: Made with Stata from the model.

As can be seen in Table 1, all the variables of the model have been crossed between them to firstly observe the existing correlations. The values of the correlations are expressed on a scale from 0 to 1, both positively and negatively. A correlation is considered significant when it exceeds the value of 0.5 and more intense the closer it is to 1. This process is very interesting to carry out when working with empirical samples,

but it is not constitutive of drawing definitive conclusions. It helps us better understand the interaction between variables, although it is not fully reliable. It could be the case that two variables presented collinearity bias, or that they gave negative correlation but then in the execution of the econometric model they expressed different conclusions.

So, if the table 1 is looked, the main correlations are found in the following three cases:

- Lngdpco - ev (0.872): There is a very strong correlation between life expectancy and per capita GDP in the migrants' countries of origin. The evidence about this data is clear: the higher the income level, the population is expected to enjoy better living conditions, and therefore a longer life.
- Lnrandd – lngdpco (0.734): In this case, a strong correlation is also observed between two variables that in principle are assumed to be positively related. When the percentage of spending on research and development over GDP increases in the countries of origin, the per capita GDP increases in the same way and strongly. Subsequently, it will be determined to what extent these variables condition migration, although right now it could be said that each euro of annual increase in spending on research and development has a positive collateral effect of 0.75 euros.
- Lnunemd – lngdpCCAA (-0.722): The variables of this correlation correspond to the destination communities and not to the countries of origin. In this case, expected results are also obtained for the economic evidence: when the local unemployment rate increases, a strongly decrease in GDP per inhabitant in each autonomous community is expect.

Based on the results presented, it could be said that for the moment the model and its variables are faithful to the empirical and economic reality, since they yield rational results. Next, the variables will be exposed from a univariate analysis perspective, with the aim of providing the reader with a greater understanding of the data at a mathematical level.

Table 1.1

Other interesting correlations of the model.

Variables	(1)	(2)	(3)	(4)	(5)
(1) lnmig	1.000				
(2) lndp	0.639	1.000			
(3) homd	-0.122	-0.196	1.000		
(4) pg5	0.054	0.008	-0.037	1.000	
(5) lnruralp	0.098	0.017	-0.021	0.460	1.000

Source: Made with Stata from the model.

In Table 1.1 we can see other variables that may have been included in the model, but this has not been the case for various reasons. Firstly, we will see what relationship exists between the variables and the main variable of the model: the logarithm of international arrivals (lnmig).

- (1) lnmig – (2) lndp: 0.639. The destination population of the autonomous communities has a significant correlation between the destination population of the autonomous communities of 63%. Although, it is a mathematically interesting correlation, it has been preferred not to include it in the model since with regard to the conclusions that this variable can yield, they are not interesting for the paper. It is very easily deductible that the larger the destination population, the greater the infrastructure and capacity that territory must host international migrants.
- (1) lnmig – (3) homd: -0.122. The relationship between international migrants and the homicide rate in the autonomous communities is not high, but the interesting thing is to know that it is negative. The proof that this variable is correct, and objective is that the results are logical: if there is a greater number of homicides in an autonomous community x, international migrants will decrease by 12% that year. In short, it is an interesting variable, but the model already contains more fruitful variables that modulate characteristics of the autonomous communities and we have preferred not to include too many control variables so as not to saturate it.
- (1) lnmig – (4) pg5: 0.054. This variable is interesting to modulate migrations between autonomous communities. This is the percentage of the total population by country of origin that lives on less than $ 5.5 international per day. Although it is an interesting variable, other economic variables have already been included, so it has been decided to include more important economic variables such as GDP per inhabitant. There is also no significant correlation that is motivating for their inclusion, it is weak, 0.05.
- (1) lnmig – (5) lnruralp: 0.098. This correlation reflects the interaction between the logarithm of migrations and the logarithm of the percentage of rural population in the country of origin over the total population. A variable has already been included in the model that studies the proportion of urban population, therefore including a variable of rural population would be the same concept but in reverse. In this case, empirical evidence of other interesting events would not be provided, and it could cause collinearity problems. The relationship is weak, but positive.

To sum up, in the last case, it could be deduced that the larger the rural population, the higher the possibility of emigrating and of choosing between different autonomous communities based on economic aspects. These economic aspects can occur because if a country is more rural and therefore less economically developed, it would tend to discriminate between richer or poorer rich autonomous communities. All these statements will be clarified later with the execution of the model.

6.3 Univariate analysis of the sample.

In the following table 2 we can see the main univariate statistical indicators (number of observations, mean, standard deviation, minimums, and maximums) of the estimated model. The data collected by our model on the 53 countries of origin and 19 autonomous communities of destination, give us interesting statistics to consider:

Table 2

Summary statistics.

Variable	Obs	Mean	Std. Dev.	Min	Max
lnmig	11307	4.299	2.063	0	10.071
lngdpco	12084	9.066	1.388	6.409	11.542
lngdpCCAA	12084	10.014	.198	9.648	10.489
lnunemd	12084	2.885	.376	1.887	3.611
lngdpgrowth	10355	1.06	.83	-2.924	3.225
lnrefugees	11818	9.212	2.605	2.398	14.458
lnrandd	7581	-.196	1.062	-4.255	1.321
ruralt	12084	67.524	17.175	28.968	98.041
ev	11077	73.766	7.512	49.913	83.754
ch	8398	3.471	2.369	.1	9.85

Source: Made with Stata from the model.

As we can see in Table 2, most of our variables have a very high sample population, which is an indicator of high statistical quality of the sample. There are also many missing values due to gaps in the sample for various reasons of data collection, or lack of transparency in some countries in providing data. The variable with which contain the least amount of data is spending on research and development as a percentage of GDP. On the other hand, there are four variables that contain all the data of the sample in a complete way, among which are the GDP per capita of the

countries of origin, or the GDP per capita of the Spanish autonomous communities.

- The largest standard deviation of the sample is found in ruralt, that is, the total rural population by country of origin. This is because the data for the main 53 countries of origin of migrants cover countries that are structurally as different in ways of living as Senegal to the United Kingdom. For this reason, the evidence indicates that a rural country like Senegal, in this respect deviates a lot from a country as urban as the United Kingdom.
- The lowest standard deviation of the sample is found in the GDP per inhabitant of the Spanish autonomous communities: lngdpCCAA. This data is because between the wealth per inhabitant of the richest community in Spain with respect to the poorest, there is not as much disparity as there is in the data collected in any other variable of the model. For example, in 2018, the GDP per inhabitant of the richest community (Madrid) compared to the poorest (Melilla), was almost the double. It is a difference of 89%, which socially implies great social differences, but not statistically and relatively.

6.4 Methodology and Model results.

In this section, the main existing methodologies and models that may be useful to the work of our sample will be discussed. Econometric statistical inference was carried out using Stata and following the three-dimensional panel data structure.

To carry out an empirical analysis of the phenomenon that concerns us, we have worked with data from secondary sources collected through observation, and not through direct experimentation with those data. At an advanced level of scientific analysis in the field of economics, there are 2 basic established methodologies to work through statistical programs: pooled cross section or panel data (Wooldridge, 2010).

In this paper, panel data was selected because it is better adapted to the characteristics of our dataset. They are observed chronologically over time. On the other hand, in pooled cross section, information is collected from random agents at different moments of time, and it is not useful to analyse, as in our case, a macro panel (few individuals with high temporal observations). The general panel data template follows a structure such as this:

Equation 1: *Basic structure from a fixed effects model.*

$$Y_{it} = \beta_0 + \delta_0 d2_t + \delta_1 d3_t \ldots + \delta_{n-1} dn\text{-}1_t + \beta_1 X_{it} + \alpha_i + u_{it}$$

- Y_{it} = dependent variable of the model, in our case lnmig.

- X_{it} = independent variable of the model, usually there are more than only one.

- β = model slope parameter.

- $d2_t$ = binary variable that collects the time.

- α_i = variable that collects the error that can be influenced by Y, but that is not altered over time.

- u_{it} = disturbance term that cannot be observed.

According to Wooldridge (2010), this type of estimation has several advantages that will be briefly named below:

- It allows to research on a very large population sample.
- It allows studying unobservable heterogeneity and acting accordingly.
- It is efficient to avoid biased results by including individuals in the samples who are not equal between them.
- It is efficient studying adjustment and interaction processes between individuals.

Previously, different forms of the general panel data model will be analysed to process the sample with the most convenient model. Although we are going to show different variables of the model, what is clear is that the basic model with which we are going to work is a Fixed Effects model.

Finally, before advancing on the approach to make different statistical contrast will be explained. To contrast if a result model is suitable for validating a theory or variables effect on the individual or common way, there are a few performances. The one following in this paper is through the p-value. Following a structure based on alternative or null hypothesis, if p-value issued by State it is a number below 0.05, is said this specific variable is statistically significant with a confidence of 95%. There are principles of this system that prove its validity, and which can be consulted in Wooldridge (2010).

$H_0 : \beta_n = 0$

$H_1 : \beta_n \neq 0$

H_0 is the null hypothesis and H_1 is the alternative. When the contrast is carried out, if the p-value is less than 0.05, we affirm that the null hypothesis is rejected, and the alternative hypothesis is validated. The fact of confirming H_1 tells us that β_n is different from 0 and is relevant, also it has a modulable effect and that affects the dependent variable. Therefore, this approach will be followed in the following pages when explaining the individual or joint meanings of the model.

6.4.1 Fixed effects model.

The importance of the fixed effects model is that it collects the individual effects of the sample of the 53 countries that change over time and for each annual period. In the case of the model in question, must be accomplished by fixed effects which changes over time and is not captured by any variable, for example: the motivation of the inhabitants of a country to emigrate for personal or cultural reasons that are impossible to collect and analyse totally. This variable is the one mentioned before (α_i). Therefore, with the robustness command, we are going to make sure through Stata that the model does not present heteroscedasticity.

The finished and complete model will be as the following representation:

Equation 2: *Fixed effects model with our variables included.*

$$lnmig_{it} = \beta_0 - \beta_1 lngdpco_{it} + \beta_2 lngdpCCAA_{it} + \beta_3 lnunemd_{it} + \beta_4 lngdpgrowth_{it} + \beta_5 lnrefugees_{it} + \beta_6 lnrandd_{it} + \beta_7 ruralt_{it} + \beta_8 ev_{it} + \beta_9 ch_{it} + \alpha_i + u_{it}$$

The following pages will show several model types in panel data using fixed effects. They are all fixed effects, built with different approaches. These different approaches are grouped in the following formatted Table 3, which includes a different model for each of the columns. Therefore, to understand models each column, the description is below:

- (1) PairFE: lin-log model, a semi-log regression.

- (2) PairYearFE: lin-log model with years. It is the same semi-log regression but affected for the dummy years variables.
- (3) PairYearFE: log-log model with years. It is the chosen model for the sample, with the reasons explained later.

Table 3

Different interpretations of the model

VARIABLES	(1) PairFE	(2) PairYearFE	(3) PairYearFE
lngdpco	-352.1***	-632.1***	-0.540***
	[116.9]	[161.2]	[0.0915]
lngdpCCAA	731.1	357.7	0.822
	[475.5]	[941.9]	[0.631]
lnunemd	-319.7***	-366.8***	-0.634***
	[94.15]	[94.93]	[0.114]
lngdpgrowth	24.55	29.07	-0.00864
	[18.73]	[20.72]	[0.00996]
lnrefugees	6.046	3.388	-0.0456***
	[23.83]	[24.72]	[0.0150]
lnrandd	-165.2*	-143.2	-0.290***
	[86.15]	[88.25]	[0.0584]
ruralt	31.33***	21.41*	0.0261**
	[11.66]	[12.84]	[0.0129]
ev	25.23	-15.50	-0.0938***
	[27.61]	[22.90]	[0.0200]
ch	1.047	68.38	-0.0569*
	[48.86]	[60.39]	[0.0311]
2009.year		-69.21	0.0314
		[73.80]	[0.0756]
2010.year		-54.80	0.0376
		[73.65]	[0.0758]
2011.year		125.3	0.235***
		[101.1]	[0.0889]
2012.year		139.0	0.213*
		[145.8]	[0.114]
2013.year		168.3	0.196
		[156.1]	[0.120]
2014.year		135.6	0.205*
		[136.8]	[0.110]
2015.year		54.82	0.111
		[102.0]	[0.0945]

2016.year		79.65	0.178**
		[85.61]	[0.0831]
2017.year		185.3**	0.311***
		[78.98]	[0.0774]
2018.year		312.7***	0.322***
		[98.24]	[0.0884]
Constant	-6,891	3,199	8.891
	[5,256]	[9,412]	[6.572]
Observations	5,485	5,485	5,398
R-squared	0.069	0.087	0.254
Number of id	844	844	837
rmse	450.4	446.4	0.358
ll	-41291	-41238	-2105

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Source: Made with Stata from the model.

6.4.2 Lin-log model.

This model approach determines a change in the dependent variable Y due to the interaction of a percentage change in an independent variable X. This approach is usually used to model in situations or studies that need to know the results in absolute numbers. For example, if we need to build an expense model in a house renovation based on percentage increases in surface area in square meters, percentage increases in the value of materials, or the total cost of living in that area.

In the Table 3 column 1 (PairFE), it is observed how the model would be following this typology. As it is seen, in this case it is observed that only three of the nine independent variables would be significant at an acceptable level of 95% confidence. Also is it interesting to present the results in absolute data, since they are more difficult to analyse, process and draw precise and optimal econometric conclusions.

Although this interpretation is not the definitive or the chosen model, several interesting data will be interpreted to understand the data processing in this model:

- With 99% confidence, a 1% increase in per capita GDP in the countries of origin causes a drop of 352 international migrants.
- With a confidence level of less than 90%, an increase of 1% in the expenditure on research and development of the countries of origin over the GDP, reduces the number of migrants to the autonomous communities by 165. This is because

there is a strong relationship between technological development and positive economic developments.

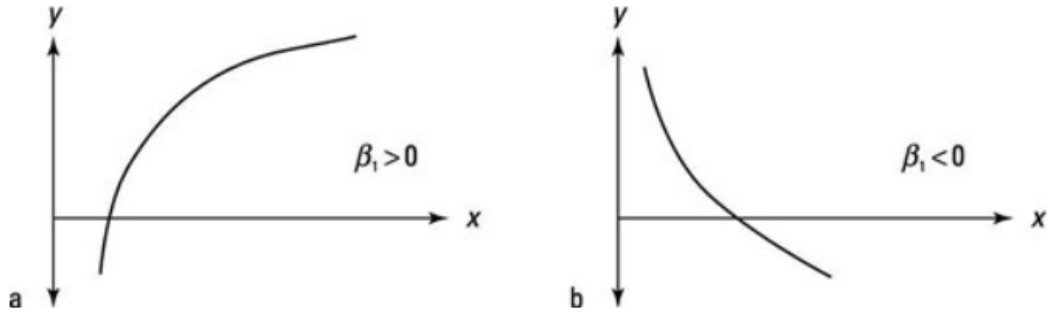
Typically, the lin-log model approximates the interpretation of a quadratic graph, such as the ones shown below (see on Figure 6):

$\beta_1 > 0$ = It is shown on the left, and represents a positive effect of the independent variables on the dependent variable.

$\beta_1 < 0$ = It is shown on the left, and represents a negative effect of the independent variables on the dependent variable.

Figure 6

Quadratic prediction from a lin-log model.



Source: economipedia.com.

6.4.3 Lin-log model with years.

This model also follows the same typology but introducing the time dummy variables. Following table 3 second column, the reasons for not choosing this model will be explained in this case.

The model in table four gives more negative results despite including time variables. In this case, the only variable that improves the model is the main control variable, lngdpco. In this case, for each percentage increase of 1% in per capita GDP in the countries of origin, international entries in Spain decrease by 632 people. On the other hand, observing another important variable for the model (lnrefugees), the inclusion of annual dummy variables causes a lower increase in migrants than the previous model (by half). Likewise, in these two cases the variable is not statistically significant.

Regarding the dummy variables of years, they are also not significant for the most part. Only the years 2017 and 2018 of the sample are significant with 95% confidence. The most probable here is when adding more variables (the years), on a previous model that was not correct, the errors are fed back.

6.4.4 Log-log model with years.

After having studied several models and their validity for the sample, in this section it is chosen the model that finally will perform que data, and with which the relevant contrasts will be made to confirm or reject the statements made at the beginning, and to extract empiric and useful theories for economic science.

The model shown in the column three ((3) PairYearFE) table 3, is the final model and the one used to interpret the sample. This log-log approach is the one that best fits to the data, and due to its configuration, is the one that validates the previous arguments to a greater extent, and that yields the statistical significance of a greater number of variables. This model is interpreted as the elasticity of Y over β_1. An increase in an independent variable X_1 of 1% causes a change in the dependent variable of β_1%.Next, all the econometric and economic conclusions of the model will be presented:

- Ceteris paribus, with a high confidence of 99%, jointly our model is significant to explain the logarithm of international migrations.
- Ceteris paribus, with a confidence level of 90%, our model is individually significant for all variables, except with lngdpCCAA and lngdpgrowth. For the time dummy variables, the years 2009, 2010, 2013 and 2015 are also not individually significant for a moderate confidence of 90%.

Figure 7

Contrast hypothesis.

1. Well-off and poor immigrants tend to emigrate for similar reasons to the autonomous communities.
   i) Autonomous Communities with high rates of wealth growth tend to receive more people.
   ii) Rural and urban development also affects destination.

2. Refugees and immigrant fluency depends also on non-economic factors.
   i) Safety and wellness ratios by destination and origin influence the target autonomous community.
   ii) Technological development of origin country affects international arrivals to Spanish territory.

Source: Own Stata processing.

As it is said before, and remembered in the Figure 7 but settled in the paper's introduction, are here the economic conclusions linked with initial approaches:

- According to the model, and a confidence of 99%, migrants who choose between different autonomous communities, behave differently when the economic evolution of their countries of origin changes. For example, if in a country x between the 53 chosen, and GDP in the country of origin evolves by 3%, international migrants to the Autonomous Communities will decrease by -1.62%. Given a greater increase in per capita GDP, a greater decrease in international inflows.
- Due to the results produced by the variable lngdpgrowth, we must discard the hypothesis that communities with high growth rates gather migrants in part, for that reason. With a p-value of 0.386, we do not reject the null hypothesis that tells the slope of the parameter is equal to 0. This would indicate a very important statement that the economic growth factor in the autonomous communities is not significant to explain whether a migrant chooses between one community or another to live. The annual evolution of the GDP per inhabitant of the autonomous communities is not significant either, supporting this hypothesis that the economic situation of the autonomous communities is secondary when a person decides to settle in Spain.

- At this point, an answer is given as to whether there is a starting bias related to the environment where they live among people who emigrate to the autonomous communities. With a 95% confidence, higher percentage of urban population in origin, causes increases in migrants in a positive way. Although it is a weak relationship: every + 10% of the urban population percentage of the total, migration increases by + 0.2%. This is a controversial aspect since it would be reasonable to think that more rural countries are poorer, they tend to look for opportunities abroad. In this case, the connections caused by an urbanized country with infrastructure that ease international transport, outweigh the poverty feature of a rural country. Consequently, the thesis of the *Network Society* promulgated by Castells (2009) is validated.

Until here, the economic / social hypotheses of the model raised at the beginning have already been analysed. Next, there is a review of the results of the variables necessary to confirm or deny the other 3 hypotheses that have been raised at the beginning, related to: refugees, security, and technology.

- The refugee variable is significant with 99% confidence. An increase in the arrival of refugees of + 10% would cause a decrease in the arrival of international migrants by -0.4%. In this case, it can be affirmed that refugees increase international arrivals legally, or that there is a small substitution effect. It surely seems insignificant, but if it is presented in absolute numbers, the impact of refugees is better understood. For every 5,000 refugees who enter, 230 legal immigrants stop coming. The possible connection will be discussed below in the conclusions.
- In this case, safety and wellness ratios are represented by three important variables: lnunemd, ev and ch. The logarithm of the unemployment rate of the autonomous communities is not significant. This output confirms the evidence that economic welfare factors are not the backbone of the destination decision to emigrate to Spain. Likewise, the other two variables are significant with a confidence level of 90% and 99%. Every 2 years of increase in the average life expectancy in the countries of origin (ev) causes a decrease in migrants of -0.188%. In the variable of beds per 1,000 inhabitants, a decrease in international arrivals of -0.57% is recorded for every 10 extra beds that are added to the system per 1,000 inhabitants. In a country like Spain, it would be adding 460,000 beds to the entire health system. Although it seems that these variables affect

migration very little, it must be considered that a phenomenon like immigration is conditioned by many different variables, which are difficult to fully capture in an econometric model.

- Finally, the variable of technological development in the countries of origin also yields the expected result. With 99% confidence, the logarithm of R&D over GDP in the countries of origin (lnrandd), causes a decrease in international arrivals to Spain. With higher technological spending on research and development, the economic situation of the country of origin improves and consequently its inhabitants are not so motivated to emigrate. Ceteris Paribus, each increase of + 10% in the percentage of expenditure over GDP in R&D, reduces international arrivals by -2.9%.

7 Verification and robustness.

Before proceeding to the conclusions of the model, a comparative table will be presented that shows the influence of working with robust or non-robust models. The objective of knowing the robust errors is to satisfy the assumption of homoscedasticity, by which the errors are distributed in a normal way. Otherwise, heteroscedasticity is obtained, implying that the variance of the errors changes for each value of X_n. For the model, the second column of fixed effects with robust errors has been used, although it is necessary to understand what differences would imply not controlling for robust errors in the results:

Table 4

Comparison table of OLD, FE, and RE models.

	(1)	(3)	(4)
VARIABLES	OLS	FE	RE
lngdpco	-0.913***	-0.540***	-0.403***
	[0.0648]	[0.0915]	[0.0724]
lngdpCCAA	6.499***	0.822	1.552***
	[0.231]	[0.631]	[0.365]
lnunemd	3.723***	-0.634***	-0.496***
	[0.160]	[0.114]	[0.113]
lngdpgrowth	-0.0906**	-0.00864	-0.00570
	[0.0369]	[0.00996]	[0.00988]
lnrefugees	0.0856***	-0.0456***	-0.0252*
	[0.0146]	[0.0150]	[0.0142]
lnrandd	-0.245***	-0.290***	-0.154***

	[0.0519]	[0.0584]	[0.0494]
ruralt	0.00347	0.0261**	0.0161***
	[0.00253]	[0.0129]	[0.00579]
ev	0.178***	-0.0938***	0.0210
	[0.0123]	[0.0200]	[0.0157]
ch	0.0983***	-0.0569*	-0.0263
	[0.0125]	[0.0311]	[0.0220]
2009.year	-2.198***	0.0314	-0.0374
	[0.190]	[0.0756]	[0.0711]
2010.year	-2.192***	0.0376	-0.0836
	[0.144]	[0.0758]	[0.0692]
2011.year	-2.506***	0.235***	0.0662
	[0.149]	[0.0889]	[0.0787]
2012.year	-3.030***	0.213*	0.0339
	[0.168]	[0.114]	[0.0961]
2013.year	-3.428***	0.196	-0.0235
	[0.170]	[0.120]	[0.101]
2014.year	-3.026***	0.205*	-0.0512
	[0.164]	[0.110]	[0.0923]
2015.year	-3.047***	0.111	-0.145*
	[0.157]	[0.0945]	[0.0836]
2016.year	-2.617***	0.178**	-0.106
	[0.151]	[0.0831]	[0.0741]
2017.year	-2.129***	0.311***	-0.0123
	[0.142]	[0.0774]	[0.0650]
2018.year	-1.842***	0.322***	-0.0144
	[0.155]	[0.0884]	[0.0688]
Constant	-75.49***	8.891	-8.545**
	[2.722]	[6.572]	[3.882]
Observations	5,398	5,398	5,398
R-squared	0.189	0.254	
rmse	1.882	0.358	0.400
ll	-11064	-2105	.
Number of id		837	837

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Source: Made with Stata from the model.

Above, on Table 4 there are three different interpretations of the same final model used to perform the sample:

- (1) OLS: It is the ordinary least squares model, the best known and most used model, but not suitable for the sample. If this model is used, the heteroscedasticity derived from robust errors could not be controlled. Bias and inconsistency can be generated.

- (3) FE: In fixed effects, dichotomous variables are incorporated that examine the effects not collected by the variables and that do not change over time. In this case, each year incorporates the annual disturbance observed in the international arrivals of each country to the autonomous communities. Robust deviations are also interpreted.
- (4) RE: The random effects estimation is also interpreted controlled for robust errors. This approach works assessing whether the unobservable effects are not correlated with any explanatory variable.

Moreover, RE and FE equations are both achieved considering robust mistakes. Robust regressions are set up not to be overly influenced by assumption breaches due to the underlying data generation process. It is advisable to use robust models most of the time, but it is more important when heteroscedasticity (atypical error variance) is present or as a suspicion (Labra and Torrecillas, 2014).

As can be seen in table 4, differences are significative applying the model when it is executed by fixed effects (*FE*), or by random effects (*RE*). However, when looking at the first column (*OLS*), the results are very disparate and contrary to the evidence. This model is not capable of working with the complexity required by the sample, and yields results that are far from reality. An example of the inconsistency in the *OLS* model without robustness and with dummy variables of years, executed with a method that does not fit are some facts that I will show below:

- According to OLS, with 99% confidence, a 1% increase in per capita GDP of the regions, causes a 6.5% increase in migrants. It has been previously shown that this fact does not correspond to reality, better interpreted with fixed or random effects.
- According to OLS, with 99% confidence, the increase in life expectancy and the number of beds in hospitals in the countries of origin has a positive effect on migration. This is contrary to all evidence and has previously been demonstrate the opposite.

Finally, the model affirms that the annual dummies, keeping constant all the other factors, have a negative effect on migrants, namely, when years elapse, migration tends to decrease. This has also been shown contrary to the evidence, organizations such as the *INE* or the *World Bank* collect every year sustained increases in international flows of migrants.

8 Conclusions.

Final model will be the following:

Equation 3: *Final log-log model.*

$$lnmig_{it} = 8{,}891 - 0{,}54\ lngdpco_{it} + 0{,}822\ lngdpCCAA_{it} - 0{,}634\ lnunemd_{it} - 0{,}009\ lngdpgrowth_{it} - 0{,}046\ lnrefugees_{it} - 0{,}29\ lnrandd_{it} + 0{,}026\ ruralt_{it} - 0{,}094\ ev_{it} - 0{,}057\ ch_{it} + \alpha_i + u_{it}$$

In this conclusion I have wanted to analyse what types of migrants arrive in the Spanish autonomous communities, with several objectives of impact and social and economic improvement. Conclusions could be useful for various reasons related with business or public administration:

- For businesses: Depending on the migration of a type of person to one community or another, companies can take this into account when distributing their products more efficiently throughout the territory or establishing new companies in places where migrants arrive for work reasons.
- For the public administration: For the state, or local or regional administrations, this dissertation can clarify aspects in more efficient allocation of resources. If the empirical evidence shows there are migrants who come for reasons unrelated to the economy, a priori they will not need public allocation of resources to social protection, as much as a migrant would need it for economic reasons.

The first conclusion I can affirm is that there are many different reasons for emigration to Spain. There is no major reason to be established in the countryside. Also, depending on the model, the economic aspect is the least important. There are no significant positive links when macroeconomic well-being indicators increase in response to the increase in migrants. In this case, I offer the reader two assumptions on this aspect, through the model, the readers will be able to derive its best performance:

1- The improvement of the economy is not decisive to attract migrants to Spain.
2- The model does not include this relationship because the migrant would not decide to come to Spain the same year that economic improvements take place, but later, when the potential migrant has already assimilated that Spain is a prosperous country once those improvements have reached the productive economy and several periods of growth have been chained.

After confirming or accepting the different hypotheses raised at the beginning of the dissertation, now I am going to explain what I think is the average migrant who comes to Spain based on the data of the model and the existing theory. Next, to add wealth to work and give it a current context, through the econometric model, I will describe what would be an optimal international and local situation in Spain to receive many international migrants. I think that with a future perspective, it is interesting to know the context that would favor receiving many migrants, since, as we have seen previously, the population pyramid of foreigners is much younger, they can relatively contribute more than local workers to the challenges that Spain will face in the future: pensions, aging...

Based on the data, I think that the average migrant to Spain is cosmopolitan and educated. It is not extremely poor, and they generally seek a balance between quality of life and standard of living. These statements are because the variables for the evolution of GDP in the regions of Spain are not significant, whereas the variables that reflect the evolution of the economy and technology spending as the source are sensitive. Furthermore, migrants are not sensitive to the economic evolution in Spain, but they are especially sensitive to the evolution of employment, it could be deduced that although the economy does not evolve significantly if they have a guaranteed job, they remain in Spain. Afterward the assertions, I conclude a foreign immigrant marginally renounces to stay in his country, even if the economy evolves slightly better than in Spain, but if his country of origin improves to more than two digits, entries in Spain are greatly affected.

I say they are cosmopolitan and educated because the inputs get better when the percentage of the urban population gets better in their home country. They have increased access to education; universities are in cities, and there is also fast and cheap internet. To the statement that I have made previously, I add that the migrant seeks a balance between level and quality of life because he flees his country significantly and arrives to Spain when life expectancy and the quality of health at origin decrease. The interesting and fundamental fact is that they keep coming even if the Spanish economy deteriorates, but only if they have a guaranteed job. Finally, it is also relevant to know that refugees influence the inflow of legal migrants, but few: for every 100 refugees who have entered Spain, legal immigrants have declined by four. Although it is significant, in the dummy variable for the year following the refugee crisis (2016), it is observed that for every 1% influence of the external factors collected in the dummy, migrants increase by 0.17%. In other words, the positive effects on non-inputs to the model variables are four times greater than the negative effects on the influx of refugees. Additional external conclusion that we could draw for the knowledge of the rulers or poli ticians opposed to hosting refugees, is that humanitarian aid hosting refugees does not

harm the country's economy, nor the general situation, nor the future entry of legal inte rnational migrants.

9 References.

Castells, M. (March 2009). *The Rise of the Network Society.* Santa Monica: Wiley-Blackwell.

Castles, S. (2000). International migration at the beginning of the twenty-first century: Global trends and issues. . *International Social Science Journal*, 269-281.

Castles, S. (2010). Understanding Global Migration: A Social Transformation perspective. *Journal of Ethnic and Migration Studies* , 1565-1586.

Castles, S. (2010). Understanding Global Migration: A Social Transformation Perspective . *Journal of Ethnic and Migration Studies*, 1565-1586.

Dinan, D. (2004). Europe recast: a history of European Union (Vol. 373). Basingstoke: Palgrave Macmillan.

Enríquez, C. G. (2015). La crisis de los refugiados y la respuesta europea. *Análisis del Real Instituto Elcano (ARI)*, (67), 1.

García Ballesteros, A., Jiménez Basco, B., & Redondo González, Á. (2009). La inmigración latinoamericana en España en el siglo XXI. *Investigaciones geográficas*, (70), 55-70.

Givens, T., & Luedtke, A. (2005). European immigration policies in comparative perspective: Issue salience, partisanship and immigrant rights. Comparative European Politics, 3(1), 1-22.

Jumilla, A. R. (2006). Remesas familiares y su impacto en el crecimiento económico 1950-2004. *Revista análisis económico*(46), 23-57.

Labra R., and Torrecillas C. (2014). Guía CERO para datos de panel. Un enfoque práctico. *Cátedra UAM-Accenture en Economía y Gestión de la Innovación, 16, 1-60.*

Macmillan, P. (2011). International migration and social theory. En K. O'Reilly, *International migration and social theory* (pág. 130). Loughborough : Palgrave Macmillan.

Montserrat Casado, L. M. (2003). *El análisis económico de los movimientos migratorios internacionales: ampliación y nuevas propuestas de regulación.*

Moreno, L. (2012). Inmigración y Estado de bienestar en España.

OReilly, K. (2000). The new Europe/old boundaries: British migrants in Spain.

O'Reilly, K. (2012). *International migration and social theory.* Basingstoke: Palgrave Macmillan.

Oyarzun, J. (2008). Causas y efectos económicos de la inmigración. Un análisis teórico-empírico. *Papeles del este, 17*, 4-20.

Portal de datos mundiales sobre la migración. (29 de Mayo de 2021). *IOM's GMDAC*. Obtained from https://migrationdataportal.org/es?i=stock_abs_&t=2020

Ravenstein, E. G. (1889). The Laws of Migration. *Journal of the Royal Statistical Society, 52*(2), 241-305. Obtenido de https://www.jstor.org/stable/2979333

Roselló, V. C. (2008). Las migraciones desde una perspectiva histórica. *Revista de treball, Economia i societat, 49*, 9-13.

Pértegas Díaz, S., Pita Fernández, S. (2001). Unidad de Epidemiología Clínica y Bioestadística. Complexo Hospitalario Universitario de A Coruña (España) Cad Aten Primaria; 8: 268-274.

Silvestre, J. (2005). Internal migrations in Spain, 1877–1930. *European Review of Economic History*, *9*(2), 233-265.

Wooldridge, J. (2019). Introductory Econometrics: A Modern Approach. *South-Western College Publishing*; 007 edition.

Berti, B. (2015). The Syrian refugee crisis: Regional and human security implications. *Strategic Assessment*, *17*(4), 41-53.