

Determinants of the demand for education in Spain and Catalonia

Bachelor's degree in Economics

Author: Agustin Fernández Felguera

Tutor: Esteve Sanromà

Academic year: 2019-20



UNIVERSITAT DE
BARCELONA

Facultat d'Economia
i Empresa

Abstract

The objectives of this work are to identify the determinants of demand for secondary and higher education in both Spain and Catalonia and to analyse whether there is a certain intergenerational determinism in education. The database analysed is the Labour Force Survey for the 2007-2019 period. A logistic regression model is used to determine the function of educational demand, while intergenerational educational mobility is estimated using mainly transition matrices.

The main results, referring to both the demand for post compulsory secondary and higher education, give a key and positive role to the parental education and to the gender of the individual, while the size of the household and the unemployed in the home have a negative influence. Moreover, differences between autonomous communities and between different years are found. On the other hand, and in relation to mobility, it can be seen for the Spanish case that children are more likely to achieve higher levels of education than their parents and that the fact that the parents have achieved higher education determines to a large extent that the children do too.

Keywords: Education demand modelling, post compulsory education, tertiary education, Labour Force Survey, intergenerational educational mobility, Spain and Catalonia.

Resum

Els objectius d'aquest treball són identificar els determinants de la demanda d'educació secundària post obligatòria i superior tant a Espanya com a Catalunya i analitzar si hi ha un cert determinisme intergeneracional en l'educació. La base de dades analitzada és l'Enquesta de Població Activa del període 2007-2019. S'utilitza un model de regressió logística per determinar la funció de la demanda educativa, mentre que la mobilitat educativa intergeneracional s'estima principalment mitjançant matrius de transició.

Els principals resultats, referits tant a la demanda d'educació secundària post obligatòria com a la superior, donen un paper clau i positiu a l'educació dels pares i al gènere de l'individu, mentre que la mida i l'atur de la llar tenen una influència negativa. A més, es troben diferències entre comunitats autònomes i entre diferents anys. D'altra banda, i en relació a la mobilitat, es pot veure pel cas espanyol que és més probable que els fills assoleixin nivells d'educació més elevats que els seus pares i que el fet que els pares hagin assolit una educació superior determina en gran mesura que els nens també ho facin.

Paraules clau: Modelització de la demanda d'educació, educació secundària obligatòria, educació superior, Enquesta de Població Activa, Espanya i Catalunya.

Index

1. Introduction	2
2. Education and economy in Spain and Catalonia	4
2.1. Spanish educational system	4
2.2. Transition to the labour market	5
2.3. Situation in Catalonia	7
2.4. Business cycle	8
3. Theoretical framework	9
3.1. Theoretical base.....	9
3.1.1. Human Capital Model	9
3.1.2. Effects of the business cycle.....	10
3.2. Research background	11
3.2.1. Common results	11
3.2.2. Studies on the effect of economic crisis	14
4. Methodological framework	15
4.1. Data	15
4.2. Analytical methods	17
5. Results	20
5.1. Spain	20
5.1.1. Post compulsory secondary education	20
5.1.2. Tertiary education.....	23
5.2. Catalonia.....	26
5.2.1. Post compulsory secondary education	27
5.2.2. Tertiary education	28
6. Intergenerational educational mobility in Spain	30
6.1. Reseach background.....	31
6.2. Methodology	31
6.3. Data	32
6.4. Results	33
7. Conclusions	37
8. Bibliography	39
9. Webgraphy	40
10. Appendix	41

1. Introduction

A good educational system is considered a key element for long-term economic growth and development. What it is more, competition in international markets brought about by globalization and the ever-evolving technology make increasingly important an agile adjustment of the qualifications of the population.

Moreover, both the Spanish and the Catalan education system have significant structural deficiencies that leave them in a disadvantageous position compared to other European regions. They are certainly peculiar cases within the European Union. In both Spain and Catalonia, the rates of school failure and early school leaving are very high and, on the other hand, the number of graduates with higher education is higher than what the labour market needs. These mismatches create a group of people with very low levels of education who will have trouble entering the labour market and another group form by people with high levels of education who will suffer from high rates of over-education. This causes great dissatisfaction among the young population and limits their possibilities for job development. At last, education not only provides individual benefits but also has important positive externalities that are manifested, for example, in higher quality political participation or better parenting.

It is therefore important to analyse how these distortions are caused in order to be able to implement appropriate public policies that can correct them.

Specifically, the hypothesis of this work is that both the socio-economic and geographical environment and the cyclical situation of the economy are determining factors in explaining the education that each individual decides to achieve.

The objectives of this work are first, to make a general analysis of the determining factors of the educational level that the population decides to reach in the long term. In addition, it is seen how the economic cycle, that is, the conjunctural situation, affects the demand for post-compulsory and higher education. It is expected that in an economically unfavorable situation, students will decide to extend their education. It will also be seen what impact has the economic structure of each autonomous community or each province on the educational level of its population. Finally, a more detailed analysis of intergenerational educational mobility in Spain is made.

In order to perform the analysis I use data from the Labour Force Survey from 2007 to 2019. This period of time is interesting because in 2007 the Catalan economy was at the peak of a long expansionary phase, later there is a crisis situation (2008-2013) and finally from 2013 to 2019 there is a strong economic recovery. Moreover, the Spanish area will be analysed first to later make comparisons with Catalonia.

Throughout the work, it is seen that parental education is an essential variable for understanding children's education but there are other relevant factors. In any case, the decisions that individuals make about their educational trajectories are complex and although it is known which factors may be the most influential, much remain to be explored.

This work consists mainly of four chapters:

In the first chapter, I describe the education system and its evolution throughout the last 13 years, emphasizing the Catalan case.

The next chapters describes the theoretical framework and the methodology that will be used throughout the work.

In chapter five, I analyse empirically the variables that influence the demand for post compulsory secondary education and higher education. This analysis is done at both the state and Catalan level.

Finally, an additional chapter is added to study the situation of intergenerational educational mobility in Spain.

2. Education and economy in Spain and Catalonia

2.1. Spanish educational system

The Organic Law of Education (LOE), approved in 2006, and the Organic Law for the Improvement of the Educational Quality (LOMCE) of 2013 that modifies it, are, at present, the basic norms that regulate the Spanish educational system and define its structure.

According to the constitutional system of distribution of powers, the powers of the Autonomous Communities (CCAA) in matters of education can only be those that the Constitution has not reserved to the State. That is, state powers act as a limit to the powers that the statutes of autonomy can attribute to the autonomous communities. Despite these limitations, the Catalan Education Law (LEC) was approved in 2009, which is still in force.

In 2019, the structure of the Education System is the following:

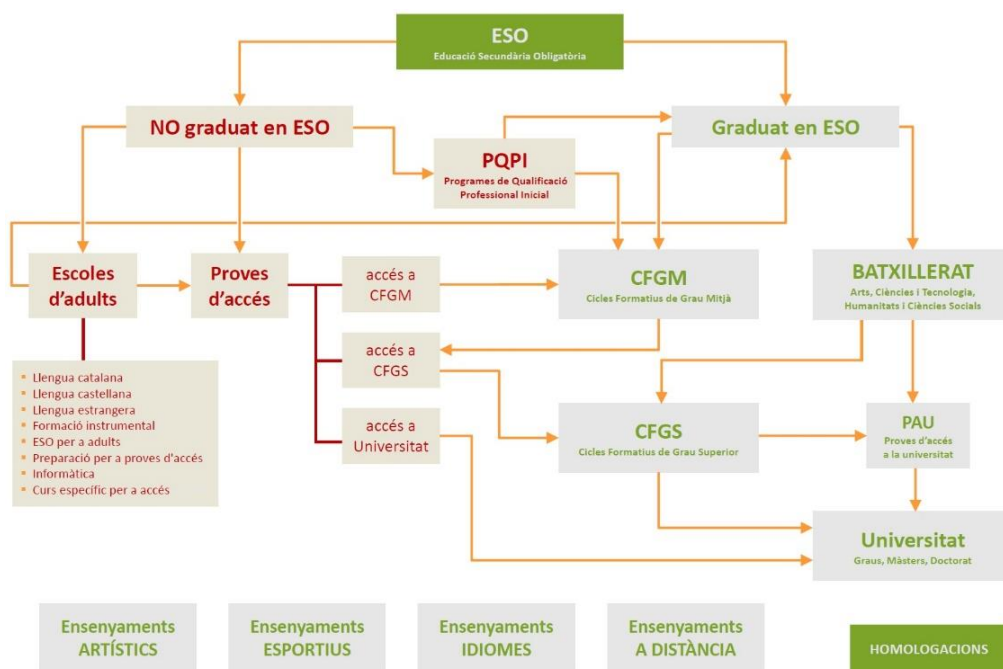


Figure 1. Source: <https://agora.xtec.cat/iecastell/el-sistema-educatiu-a-catalunya/>

The picture is valid for Spain as well as for Catalonia.

The levels offered by the education system are: Infant Education, Primary Education, Compulsory Secondary Education (ESO), Baccalaureate, Vocational or professional Training (FP), Adult Education and University education. On top of this, the Teachings of Languages, Artistic Teachings and Sport Teachings are offered, considered Teachings of Special Regime.

These levels can be grouped into three major blocks:

- **Primary or elementary education:** Is form by six academic years that are usually taken from 6 to 12 years.

- **Secondary education:** It comprises two distinct stages, one mandatory and the other voluntary:
 - The last compulsory stage of the education system, which is abbreviated as ESO in Catalan, is usually taken from 12 to 16 years old (although it can be extended to 18 years if two courses are repeated). It is organized in two cycles: the first is made up of 1st, 2nd and 3rd of ESO, and the second of 4th of ESO, in preparation for post-compulsory education.
 - The non-compulsory stage includes Baccalaureate, Basic Vocational Training and Intermediate Vocational Training.
- **Higher or tertiary education:** It includes both university education and higher vocational training. It usually starts at age 18. Unlike the previous stages it is not free for families.

Throughout the educational stage, a student has to decide on several occasions when he or she wants to stop studying to enter the job market: there is a succession of educational options and in addition, each one is conditioned to the previous. In general, the most important moments are the transition from compulsory to post-compulsory education (from ESO to Baccalaureate and CFPGM¹), and then from post-compulsory education to higher education (Baccalaureate and CFPGM towards university degrees or CFPGS²). However, in Spain it is also necessary to highlight the significant number of people who do not obtain the minimum degree required by the education system, that is, the graduate of compulsory secondary education (GESO which is the first stage of secondary education).

2.2. Transition to the labour market

This transition from the education system to the labour market is very complex and presents a large number of casuistry. That is, there is no single transition from school to the job market. Young people have several options to take:

- Working but not studying. It is the typical case of a worker who no longer studies.
- Not working nor studying (but looking for a job). When a person is in this situation he or she is called unemployed.
- Unemployed student. He or she is looking for a job.
- Working and studying. It is the most “desirable” situation.
- Studying outside the workforce. He or she is a full-time student, not actively looking for work.
- Not working nor studying (and not looking for a job).

Figure 2 shows these different dynamics and how they evolve as the population gets older³:

¹ CFPGM: It is the intermediate stage of professional training, in Catalan is called “Cicle formatiu professional de grau mitjà”.

² CFPGS: It is the highest stage of professional training, in Catalan is called “Cicle formatiu professional de grau superior”.

³ We have to be cautious when interpreting this figure because it is intended to provide longitudinal information but is based on cross-sectional data (LFS).

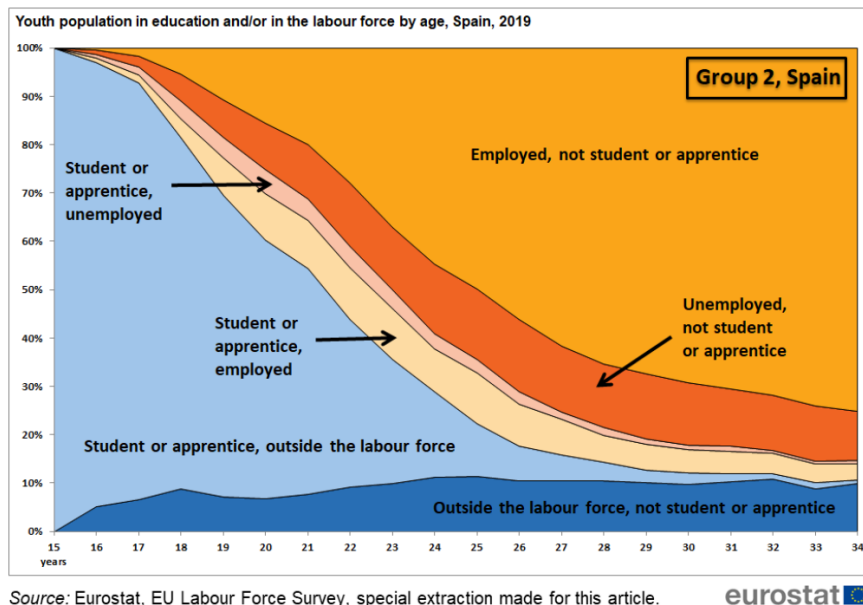


Figure 2. Youth population in formal education and/or in the labour force by age, Spain, 2019.⁴

At the beginning, students outside the labour force are declining rapidly and at the age of 25 they already account for only about 10%. These students typically enter the job market with a full-time job. However, three characteristic features of Spain with respect to other countries should be highlighted:

- From an early age there is a significant number of people who neither study nor look for work. It represents about 8% of the youth population.
- Not many students reconcile their studies with a job. This may be due to the fact that the vocational training system is not very widespread or that the youth unemployment rate is very high.
- There are many young people who are unemployed and not studying either. Probably, the main cause here is youth unemployment. This group grows when individuals are around 24 years old.

Another feature must be taken into account when analyzing this transition in Spain, which is that the Autonomous Communities have very different situations. Unemployment rates, early school leaving, school failure or education spending are very different among these territories. One might think that due to these differences there should be a significantly different educational demand in each autonomous community.

That is why in this work we will study the Spanish case differentiating the autonomous communities. In addition, the Catalan case will also be studied separately.

⁴ Source Eurostat: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Participation_of_young_people_in_education_and_the_labour_market#Participation_of_young_persons_in_formal_education_and_in_the_labour_market

2.3. Situation in Catalonia

Catalonia stands out for having a population with very high levels of education: In 2019 more than half of the people between 25 and 29 years old had reached the highest level of education. This percentage is much higher than in Spain and in the vast majority of European countries. This is positive because college graduates have higher salaries and are more likely to find work and be promoted. On the other hand, however, it creates underemployment and over education. In the first graph we can see how higher education has not stopped increasing, especially due to the sharp reduction in the number of people with primary education. This reduction in the early school leaving was accelerated as from 2012-2013.

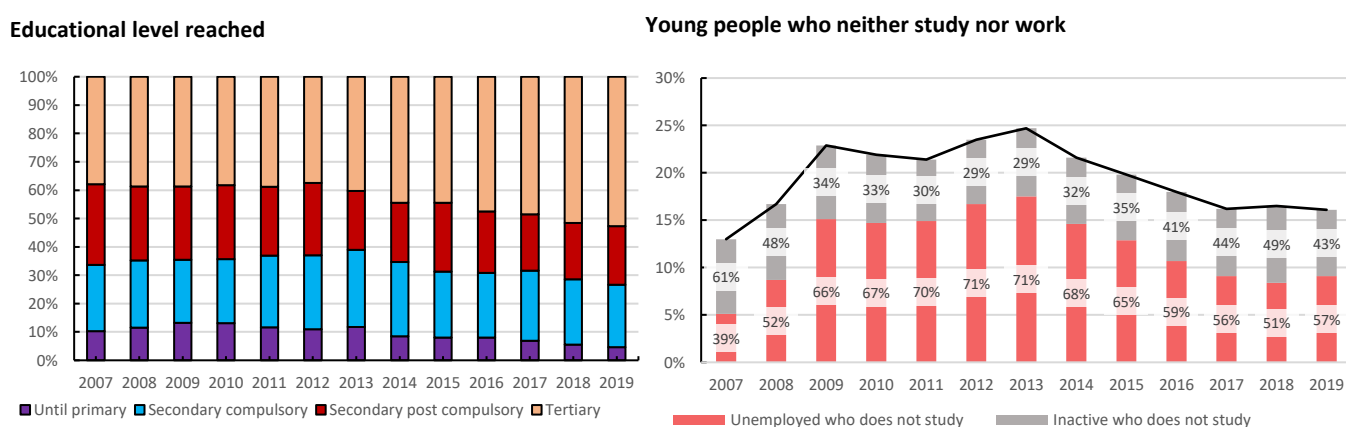


Figure 3.

Source: Own elaboration base on the labour force survey microdata. In the first figure it is considered the population between 25 and 29 years old. In the second, young people between the ages of 16 and 29 both in formal and non-formal studies.

In 2013 almost 40% of people between the ages of 25 and 29 had at most compulsory secondary education but by 2019 it was close to 25%. It should be noted, however, that these people probably made the decision to stop studying many years earlier, when they were younger.

On the other hand, if we look at people who do not study or work (also known as NINIs), we can see that their evolution has been very marked by the crisis. This rate increased rapidly with the onset of the crisis due to rising unemployment among the young population. With the recovery, the percentage of NINIs due to unemployment has decreased again, but still in 2019 it represented even more than half of the total NINIs.

In recent years, both the rate of early school leaving and NINIs have fallen sharply, while the number of students with tertiary education has increased. Mention should also be made of the increase in enrollment in vocational training courses - which has served to provide more applied education alternatives - and the decrease in repetition rates during primary and secondary schooling. This evolution has meant a convergence towards the most prosperous

European countries, that is, the crisis has europeanized us. The current model is more balanced and has a greater weight in the intermediate levels of education.

Moreover, even with the strong economic growth of the last few years, the levels of education have continue improving. Young people is not as attracted to enter the labour market as they were before. It seems that this period is different to the others in the sense that the demand for education has became less dependant on the business cycle.

2.4. Business cycle

During the period 2007-2019, three cycles can be seen in the Catalan economy. In 2007, the economy was at the peak of a long expansionary phase. As can be seen in the first serie, the young unemployment rate in Catalonia was at a record low. However, from this year onwards, unemployment starts to rise sharply until the end of 2012, when the worst moment of the crisis come. The unemployment rate more than tripled during these years from 13.4% to 50.4%. GDP per capita also falls sharply and in fact does not recover until the last quarter of 2015. Since 2012, though, these indicators have been improving. That year, a new expansionary phase begins, which still continues in 2019.

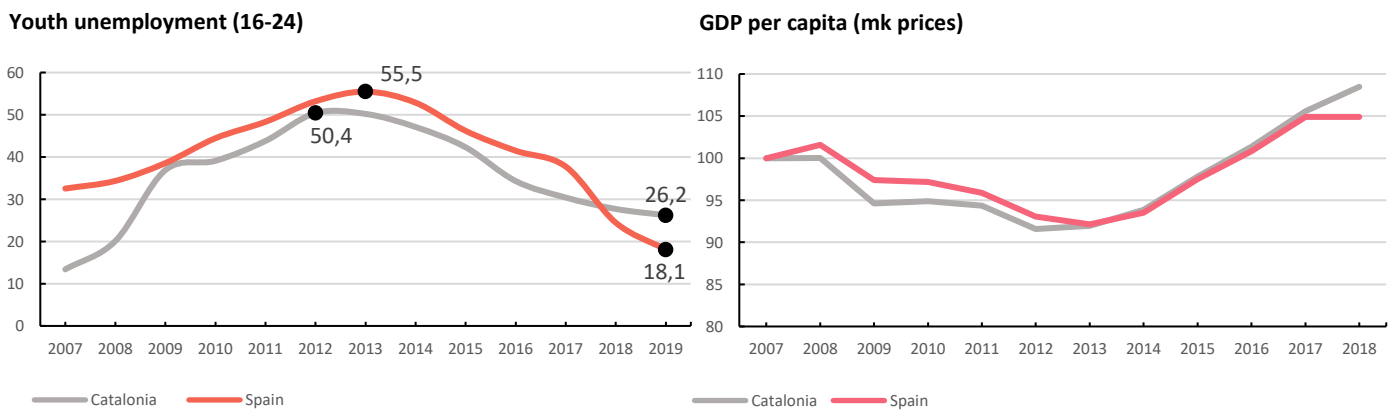


Figure 4. Source: INE, and INE regional system

Related to the economic cycle it is important to highlight spending on public education, which has been highly pro-cyclical due to economic policies pursued during the Great Recession. The policy of cut-backs started in 2012 affected the educational system in a number of forms. Some examples are the substantial increase in university tuition fees in several autonomous communities or the reduction in the amount of scholarships granted.

3. Theoretical framework

3.1. Theoretical base

3.1.1. Human Capital Model

The best-known theory of the decision to study-work is the model of human capital. This model holds that spending on education can be considered an investment in human capital that increases labor productivity (just like that of physical capital) and therefore has a rate of return.

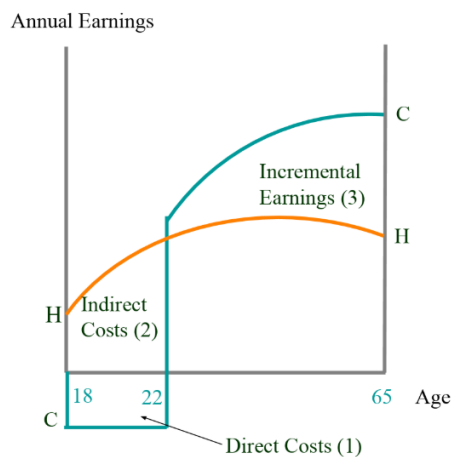
To decide whether to continue working or enter the labor market, from an economic point of view, individuals decide their optimal level of education by comparing the present value of expected earnings due to the additional years of education, with the costs associated with them.

- Costs are generally of two types:
 - The direct costs such as the tuition fee, the transport or the books.
 - Indirect or opportunity costs, which are earnings that you give up because you are studying and not working. These costs are generally higher the more education the individual has, as more training leads to better-paid jobs (increasing costs). Thus, the opportunity cost of post compulsory education will be lower than that of higher education.
- The main benefit of investing in education is the higher income you will receive over the professional career. This may be because there is less unemployment for highly qualified professionals or because their salary is higher. There are also other non-monetary benefits such as better job stability, less risky work and so on. In general, however, the benefit is decreasing with respect to the level of education because marginal productivity is lower.

Therefore, the marginal profitability of studying decreases because opportunity costs increase, often so do direct costs; and because marginal benefits decrease.

A third element must also be added because the costs and benefits of education occur at different times and are therefore not directly comparable. Costs occur in the short term, during the period in which you are studying, while the benefits extend over the entire working life. This is why the temporal preference of individuals must be taken into account: they prefer to have money today than to have it in the future.

With all these elements one could calculate the net present value to study, then the individual will decide to study if this value is greater than zero.



The following graph shows the evolution of income for a person who decides to study for a university degree at the age of 18, finish it at the age of 22 and finally decide to retire at the age of 65. This individual will decide to continue studying if the sum of Direct Costs (1) and Indirect Costs (2) updated to the present is smaller than the present value of the Incremental Earnings (3) that he can obtain throughout the professional career due to having studied a college degree.

Figure 5.

Apart from education as an investment good, in this sense, we must not forget that it can also be demanded as a consumer good, that is, in addition to future income, education also allows the achievement or reproduction of the social status, and it provides entertainment and intellectual enjoyment, among others.

There are other theories that seek to explain the increase in income caused by an increase in education. The screening hypothesis says that education increases future income not because it increases productivity but because it allows high-quality workers to be "signaled." That is, education allows companies to distinguish talented workers in a context where there is imperfect information in the job market.

3.1.2. Effects of the business cycle

Economic crisis force households to redefine economic strategies, and therefore investment in education. In this sense, the relationship between the economic cycle and the demand for education is not clear, insofar as causal relations of the opposite sign are involved, such as the income effect and the substitution effect, and therefore the net effect will depend on the intensity of each of these forces.

- **Income effect:** It occurs because, with crises, the purchasing power of families (household income) decreases, and therefore their ability to demand education. The greater the decline in family income, the less likely it is that there will be a chance of people studying in that family. This is caused by the pressure to incorporate workers who provide additional income to the meager income of the head of the family, but income necessary to prevent the household from falling into poverty. For young people this means interrupting their studies and entering a very competitive job market. On the other hand, the reduction of government spending (when austerity policies are implemented) should also be considered, as they increases direct costs.
- **Substitution effect:** Briefly, it consists of replacing employment with studies. That is, as employment is destroyed, the demand for education increases, as the opportunity cost to study decreases: the harder it is to find work or the lower wages are, the less you lose to continue studying. In this way, the labour situation becomes an important factor in understanding the variations in the demand for education.

3.2. Research background

3.2.1. Common results

This section reviews the state of the issue in the Spanish case. Several studies have been conducted, some of which date back to the 1980s, but the type of analysis has not changed much and generally similar methods are used and similar results are obtained. In any case, as far as i know, little has been studied at the regional level and there are few recent studies that have been done during the expansionary phase (2014-19). It has also been observed that there very few studies where data from a long series of years are used. Most of the studies work with one or two periods.

The most widely used data sources are the Income and Living Conditions Survey (SILC)⁵, the Labour Force Survey (LFS), the Family Budget Survey and the European Union Household Panel; all prepared by the INE or Eurostat. However, other sources of information such as census or surveys from the universities themselves have been used. At this point the main problem is that in most cases the data are cross-sectional and not longitudinal, thus reducing the possibilities of analysis.

Regarding the dependent variable, different levels of education have been analysed, although the one that arouses the most interest among researchers seems to be higher education. Some of them are:

- Administrative school failure, failure to achieve the minimum degree required by the education system (in Spain it is ESO).
- Staying enrolled between 16 and 17 years.
- Access to the University. University applicants are considered to be all those individuals who have either completed the university level or are currently pursuing it.
- Two levels of education can also be considered together. This type of analysis is interesting because the decisions that are made in the educational system are sequential. For example, the analysis of the demand for higher education is performed only for those individuals who have completed the previous educational level:
 - First stage: It is looked at whether the individual has completed compulsory secondary education.
 - Second stage: Higher education applicants are analysed.

On the other hand, not only the level of education must be taken into account, but also must be differentiated wheter the education is finished or not. Given the literature and the objectives of this work, it seems appropriate to consider three types of relationships:

- Achievement or completion of a certain educational level: Shows whether the individual has completed this level and has the corresponding degree.
- Demand for a certain level of education. It can have several interpretations:
 - a. The individual is studying this level of education.

⁵ For this topic, the SILC compared to the LFS has advantages and disadvantages. On the one hand, SILC database incorporates the variable Income of the household, while the LFS does not. As well the variable Age is not categorized. On the other hand, the SILC has a lower sampling rate and frequency. It is produced once a year and the observations only are renovated completely every six years.

- b. The individual is studying this level of education or he/she has already achieved it. It is a measure that encompasses more cases.

In addition, the results also depend a lot on the selection of the sample: Usually are studied the individuals with some age range inside the interval of 16-25 years.

Regarding the methods used to analyse the data, binary response models such as logit or probite stand out. Other more sophisticated tools have also been used, such as Heckman's double hurdle model. This method is useful to correct selection bias and to be able to study the demand for higher education given the level of education that must have been previously achieved.

Regarding the explanatory variables, in general all studies decide to include a number of similar variables. Some results are the following:

Gender: Studies have found gender differences because women are more likely to demand higher education, generally at all levels. Several studies have estimated models separating men from women to better appreciate these differences. One of the factors that can cause this effect is the fact that the gender gap (or wage gap) narrows as the educational level increases.

Birthplace and nationality: In terms of birthplace and nationality, negative effects are observed for foreigners. This may be partly due to the composition effect as many immigrants come from poor countries and come to Spain with few economic resources. It is worth noting the great wave of immigration that arrived in Spain before the Great Recession, many coming from Latin America or Morocco.

Parent education: It has often been shown that parental education is very important in determining the level of education that children will attain. It is even considered that this variable could be more relevant than household income or parental employment. In older studies it was not customary to include the studies of the mother, as only the studies of the head of the family were analyzed, but today they are both relevant. The question now is whether a parent's education can be more influential depending on the child's gender.

Household income: From an economic point of view, it is reasonable to think that family income has a positive effect on educational demand, because it shifts the marginal cost curve by lowering the economic effort to access education for higher-income families. In addition, it is plausible to think that the rate of intertemporal discount in families with higher incomes is lower, since its members will not be so impatient to enter the labour market, but will rather opt for investments in human capital that bring long-term returns. In any case, it is likely that this variable has been losing influence in recent decades due to the welfare state. In the 80's it was considered one of the most important variables but recent studies indicate that family income does not have a significant impact on the demand for education, probably because of public aid has generalized access to education.

Socio-economic level of the parents: The type of industry in which they work or the type of employment they have has often been studied. This variable is expected to be closely related to education and income.

Employment situation of the parents: It has been observed that the probability of applying for a university degree is higher among the children of professionals or managers than for the

children of unskilled workers. It is a variable that is related to parental studies and family income.

Size of the household (in terms of the number of members): In terms of family size and composition, until the 1980s it seems that the number of siblings did not have a significant effect on the demand for education, although from a certain family size a slightly negative effect was found. If this variable has an effect (presumably negative) then the order of the children could also be relevant. This variable, however, probably takes on more importance in underdeveloped countries. In addition, it is advisable to differentiate between those over and under 16 years of age, since the former can work while the latter cannot, and therefore they cannot provide income to the household.

Contextual variables: Some contextual variables that have been used, often at the level of the autonomous community or province (the spatial environment of the individual), are: the stock of aggregate human capital, the size of the individual's population, the degree of urbanization, or the economic level of the region. These factors have been measured with variables such as GDP per capita or the general and / or youth unemployment rate.

In any case, care must be taken because many of these variables are temporary and refer to the present moment, which does not necessarily coincide with the time when educational decisions were made. Therefore, using variables such as unemployment or size of the household which are generally transitory situations, it is not clear that the effect on educational demand decisions, which are generally taken with a wider time horizon, can be conclusively concluded. This is especially true when the response variable is the level of education attained and not the level the individual is currently studying.

Joint analysis of several educational levels

One issue that has not been studied much is the differences in the determinants of demand for education between different levels of education. In general, all studies focus on a single level of education, such as compulsory secondary education or higher education. It is surprising that such a limited analysis is made when with the same data and models the demand for different educational stretches can be studied.

The joint analysis can be interesting as the demand for basic education can be very different from that of higher education: For a young person aged 16-18 the decision to continue studying can depend a lot on the situation of the local labor market or more intrinsic factors such as gender, family expectations, etc. On the contrary, for higher education it should be kept in mind that direct costs are usually quite high and therefore family income and economic situation should carry more weight. In the case of university students as well, the spatial context should have less influence as mobility should be higher.

3.2.2. Studies on the effect of economic crisis

The underlying premise in studies that want to analyze the association between variations in the economic cycle and changes in the demand for education is that such an economic shock affects households differently according to their social and economic position. Particular concern has been raised about the educational decisions of low-income households.

A number of variables are generally used to distinguish between the income effect and the substitution effect:

The income effect usually shows how household income has changed. This can be seen both in the household income that can be found for example in the Income and Living Conditions Survey or in the employment status of household members, that is, whether they are unemployed or employed.

To measure the substitution effect, often researchers use the general unemployment rate or the youth unemployment rate of the individual's autonomous community as it indicates how hard it is to get a job.

The basic model could be summarised in the following equation:

$$P(\text{Study}) = a + b_1 \text{ year} + b_2 \text{ income effect} + b_3 \text{ direct costs} + b_4 \text{ opportunity cost}$$

In the practice,

$$P(\text{Study}) = a + b_1 \text{ year} + b_2 \text{ household income} + b_3 \text{ tuition fees} + b_4 \text{ unemployment rate}$$

With regard to the results obtained, it has been seen that in general the crisis in Spain did not cause structural changes in access to education according to socio-economic characteristics. Researchers don't appreciate short-term changes either, although this depends very much on the dependent variable used. On the other hand, it has been noted that in underdeveloped countries the substitution effect may be more relevant than in rich countries.

4. Methodological framework

4.1. Data

The Labor Force Survey (LFS) –also known as EPA in Spain- is a large household sample survey of quarterly frequency. It contains extensive information on issues of work and education of individuals and households in Spain. The National Institute of Statistics is the responsible for its production in all the territory. Idescat expands the sample to Catalonia, so this autonomous community is the region with the highest sampling rate.

In this work I use the microdata published on the INE website. Data from all over Spain are used, but in addition, data from Catalonia will also be used separately. Therefore, I work with two databases (the second a subset of the first). The use in a separate way of the sample of Catalonia allows to disaggregate by province⁶ and also cancels the effect that regional educational policies may have (i.e. differences in the cost of university tuition). In addition, since one of the hypotheses of the work is that there must be different behaviours between autonomous communities, it makes sense to create different models by region.

The data used correspond to the period 2007-2019; for each year the 4th quarter information is extracted and the rest of the quarters are ignored⁷. The 13 cross-sections are joined in the same data set by adding a dummy variable to differentiate each year. This results in a repeated or pooled cross-sectional design. A sample of 2.163 million individuals and 835 thousand households is obtained in Spain. For Catalonia, the sample consists of 220,000 individuals and 86,000 households. Using a pooled cross-sectional design instead of using a single period allows on the one hand to see the temporal effect on the variables of interest across the dummy *Years* but also allows for a larger sample and reduced thus the sampling errors (as long as the assumptions made are correct).

The choice of variables to be incorporated into the models is based, first, on the availability of the issues covered by the EPA and secondly on the review of the literature that has been done previously. With regard to the data, it is important to take into account two very important restrictions: First, the variable Age of the individual is categorized, the intervals are [16,19] years, [20,24], [25,29] and so on. Second, information is only available from the parents if the children have not left the family home. That is why young people living alone are not selected. In addition, the observations of Ceuta and Melilla are also ignored.

Regarding the dependent variables: Two different dichotomous variables are used as the demand for these two levels of education may depend on different factors.

- The first variable classifies individuals between the ages of 16 and 19 according to whether or not they have accessed post-compulsory secondary education⁸. This

⁶ In some autonomous communities the provinces cannot be used because there are too few observations and therefore the sampling errors of the estimates would be too high. The fact that in Catalonia the sample is expanded by Idescat allows a more detailed treatment.

⁷ Each year only one quarter is selected to reduce dependence between cross-sections. It is important to always choose the same quarter to avoid the effects of seasonality, but other than that there is no particular reason to choose the 4th. Section 4.2 explains the topic in more detail.

⁸ For survey design reasons, I had to use this age group, even if it has some problems : not all students can start studying post-compulsory studies at the age of 16, however, these represent a small part of the population aged 16-19.

includes all those individuals who have either completed baccalaureate or intermediate vocational training or are currently pursuing them. It is a concept similar to early school leaving.

- The second variable classifies individuals between the ages of 20 and 24 according to whether or not they have accessed tertiary education. This includes all those individuals who have completed either a bachelor's degree or the highest level of vocational training or are currently pursuing them.

Independent variables

<ul style="list-style-type: none"> • Individual:
<ul style="list-style-type: none"> ○ Gender: Dichotomous variable. ○ Nationality and birthplace: It is necessary to take into account both nationality and place of birth because they can provide different information. Therefore, I created a variable that combines the information from these two. The categories are: Spanish nationality and birth in Spain, Spanish nationality and birth abroad, and non-Spanish nationality.
<ul style="list-style-type: none"> • Household:
<ul style="list-style-type: none"> ○ Household's unemployment: Dichotomous variable. It indicates if there is some person working or all of the members are unemployed. ○ Father's education: It has five categories that correspond to the different educational levels, except the last one. Categories: Primary, compulsory secondary, post compulsory secondary, tertiary and no info. (when the individual doesn't have a parent or he/she doesn't live in the household). ○ Mother's education: Ídem. ○ Number of individuals living in the same household: Discrete variable. ○ Nivell socioeconòmica del cap de família: Dichotomous variable. In one category there are the individuals who are business owners with employees, directors or managers and skilled and specialized operators. In the other category there are grouped the rest of the cases.
<ul style="list-style-type: none"> • Context:
<ul style="list-style-type: none"> ○ Year: Dummy variable with 13 categories (from 2007 to 2019). ○ In the dataset of Spain: Autonomous community (CCAA), it is a dummy variable with 17 categories. For each CCAA it is computed: GDPpc, % graduates between 25 and 64 years old with higher education, general unemployment rate (population between 16-64 years old). All of them are numeric except for CCAA. ○ In the dataset of Catalonia: Province (4 categories) and unemployment rate of each province (16-64). The general unemployment rate is used and not the juvenile one because the two variables are very related and this one has more observations and therefore the estimation error is smaller. ○ → Year describes the temporal component and CCAA the geographical one. GDPpc and unemployment rate are used to measure the economic cycle while

the rate of population with higher education is related to the educational system.

Both the variables *Year* and *CCAA (or Province)* include all those characteristics of that period or region that are not included in the other variables such as unemployment rates or higher education. Some of these characteristics may be, for example, spending on education - especially relevant are the tuition fees-, the sectoral structure of the economy or the structure of the educational system (for example, the weight of private-state education or the implementation of vocational training), among many others.

Due to multicollinearity issues, I do not put all the contextual variables at once in a single model but I test several models separating the variables that can be more related and I choose the model with the best metrics. In the next section, these types of problems are discussed in more detail.

Overview of the datasets used during the analysis:

Data	Num. Individuals	Num. Households
Population between 16 and 19 years old in Spain	89696	47295
Population between 20 and 24 years old in Spain	110294	51532
Population between 16 and 19 years old in Catalonia	9088	5860
Population between 20 and 24 years old in Catalonia	10622	6409

4.2. Analytical methods

Steps followed during the analysis

1. The binomial family with a logit link will be used to model the binary response. The steps to be followed are as follows: First, a model will be created with all the covariates except the dummy *Year*. Then, to assess the impact of multicollinearity, different models will be created by dropping some of the variables that may be related to each other. Next, the variable *Year* will be added and the interactions with the rest of the covariate will be looked at. Those variables that do not provide information will be removed. Lastly, the Akaike information criterion (AIC) and the F test for nested models will be used to select the final model.
2. I will first look at the significance and sign of the coefficients. Some odds ratios will also be interpreted later.
3. To investigate whether the relations of interest change over time (structural change) I will check the significance of the dummy variable *Year*. If there are no changes and the

coefficients are constant then the data can be grouped, and thus achieve lower standard errors (due to the bigger sample size).

The use of LFS microdata presents a number of problems, mainly because they are observational data but also due to the fact that some variables have been discretized and others are not made public. A second source of problems is the use of statistical models, as like all models they have their limitations. Below is a brief description of some of these issues along with the actions taken to correct them (if any have been taken):

- The chosen dependent variables include two different casuistry: Students who are studying, and those who have already completed their studies. This usually leads to problems because effects can be confounded. In any case, I considered these variables to be the most interesting to analyze, in part due to the fact that the variable *age* is in intervals greatly restricts the possibilities to create other variables.
- Many of the contextual variables used refer to conjunctural situations of the economic cycle. The problem is that there is a gap between these variables and the moment in which the individual decides to start a new level of education. If you study the population between the ages of 20 and 24 who are accessing higher education, many of these individuals will have made the decision to enter college or vocational training (CFGS) four years earlier, perhaps even more. It is true that later each year the individual has to decide whether or not to continue studying. But the decision is no longer the same because first, the performance he has obtained that year affects a lot (you can think of the number of failed courses) and second, because the costs of finishing studies are lower (he or she has fewer years ahead).
- Most statistical analyzes assume that the data were obtained by simple random sampling (MAS). As most of the data obtained from surveys, LFS observations were not obtained with this type of design but used strata (socioeconomic factors) and conglomerates (census tracts and households). With the available microdata, only the household cluster can be taken into account as the variable indicated in the census section and the strata are not published. Because of this, standard parameter errors will be slightly underestimated.
- Since the LFS is a rotating panel, each quarter only 1/6 part of the total sample is renovated. Thus, if you use all the samples (4 for each year) there is a lot of covariance among them and it is better to use other panel models rather than the pooled model. To reduce this dependence between samples, I decided to use only one quarter for each year. Thus, in two consecutive cross sections (only) 2/6 of the individuals will be the same.
- There is a problem of sampling selection. Because educational decisions are sequential, only individuals who have completed the previous educational level can choose to pursue the next level. This is why some studies have used a double hurdle model (Heckman, 1990), although most researchers use the traditional logistics model.
- There is likely to be no structural stability, in fact it is one of the working hypotheses. As mentioned above, the *Year* variable will be used to detect this situation.
- Some variable can be related to each other, namely, there could be a problem of multicollinearity. Multicollinearity can cause unstable estimates and inflated variances which affects confidence intervals and hypothesis tests. This is especially important since the aim of the work is to make structural analysis. To do the multicollinearity

analysis, I distinguish two groups of variables. The parental studies and the contextual variables:

- Regarding the parent education variables, their correlation with Cramer's V statistic was calculated and values between 0.4 and 0.48 were found for the different data sets used (0 indicates no relation and 1 maximum relation). Some studies create separate models to differentiate between father's studies and mother's studies and avoid multicollinearity problems. Another reason why separate models are created is that the data have a considerable number of young people for whom there is no information from their parents. In any case, it has been decided that it is better to use a single model and create a category called "no info." for cases in which no information is available from any of the parents.
- Another important multicollinearity problem can emerge from the relationship between the dummy CCAA and the other contextual variables (GDPpc, unemployment rate and rate of higher education). That is why I created several models omitting some of these variables and then I evaluated the results. In all the cases, the coefficients for unemployment rate, GDPpc and rate of higher education of the CCAA were below 0.05 (though some of them were significant). In most of the final models I decided to keep the dummy CCAA and remove the other contextual variables.
- For tertiary education demand, the dependent variable is very balanced but not happens the same for the demand of post-compulsory secondary education. In any case, since the aim of the work is to analyze the relevance of the variables and not to make predictions, this is not a serious problem.
- Adding aggregate variables such as the region's unemployment rate to a microdata sample can skew standard errors. This problem has been notoriously discussed by Moulton (1990) and Wooldridge (2003). That being said, most of the studies do not correct this problem, and those who do, obtain similar results: One example can be found in Gil Izquierdo, María & Escobar, Laura & Torres, María. (2010).

While these issues may affect the reliability of the results, I consider that the overall conclusions remain valid, i.e., these problems are not determinant to the analysis.

5. Results

The results consist of four sections. The first two analyse the demand for post compulsory and tertiary education in Spain and the last two do the same for Catalonia.

In each section I only include a table with the most important variables, which are statistically significant, of the final model. More information can be found in the appendix (including the complete output and graphics).

5.1. Spain

5.1.1. Demand for post compulsory education in Spain

The next table provides a descriptive summary of the variables described in the previous section, differentiating between two groups: individuals who have accessed post-compulsory secondary education (are studying or have already reached the level) and individuals who have not. Some variables show notable differences.

In relation to individual variables, it can be seen that the frequency of women seeking tertiary education is much higher than that of men: 87% of women aged 16-19 have access to post-compulsory secondary education. By nationality and birthplace there are also differences, especially for those who don't have the Spanish nationality. With regard to the education of the parents, there are very remarkable differences. If the parents have achieved tertiary education most surely their children will reached that level too, especially in the case of the father. Depending on the family's unemployment situation, it can be seen that the frequency of access is not the same. The same goes for the autonomous communities, as there are some such as the Basque Country or the Floral Community of Navarre where the detachment of young people from the education system at this age is minimal.

Summary table - Year 2019 - Spain

Variable	Categories	Secondary post compulsory education	
		No access	Access
Total		16,8%	83,2%
Gender	Male	20,3%	79,7%
	Female	13,1%	86,9%
Nationality and place of birth	Nationality and birth place - Spain	14,5%	85,5%
	Birth place - other	22,6%	77,4%
	Nationality and birth place - other	32,0%	68,0%
Father's educational attainment	Primary	37,7%	62,3%
	Secondary compulsory	20,8%	79,2%
	Secondary post compulsory	11,4%	88,6%
	Tertiary	5,3%	94,7%
	No info.	22,5%	77,5%
Mother's educational attainment	Primary	42,2%	57,8%
	Secondary compulsory	23,2%	76,8%
	Secondary post compulsory	13,7%	86,3%
	Tertiary	6,0%	94,0%
	No info.	25,9%	74,1%
Socio-economic status	Others	17,1%	82,9%
	High	15,8%	84,2%
Unemployment household members	All members are unemployed	27,8%	72,2%
	Not all	16,1%	83,9%
CCAA	Catalunya	15,6%	84,4%
	Andalusia	21,4%	78,6%
	Aragó	18,5%	81,5%
	Astúries	14,6%	85,4%
	Canàries	24,1%	75,9%
	Cantàbria	16,2%	83,8%
	Castella-La Manxa	18,7%	81,3%
	Castella i Lleó	14,9%	85,1%
	Comunitat de Madrid	13,8%	86,2%
	Comunitat Foral de Navarra	7,6%	92,4%
	Comunitat Valenciana	16,4%	83,6%
	Euskadi	5,7%	94,3%
	Extremadura	17,9%	82,1%
	Galícia	12,9%	87,1%
	Illes Balears	22,7%	77,3%
	La Rioja	6,4%	93,6%
	Regió de Múrcia	19,6%	80,4%
	Number people household	3,98	3,91
Regional	GDPpc	24179	25105
	Unemployment rate (16-64)	11,04	10,41
	Tertiary studies rate (16-64)	36,48	37,99

Note: The dependent variable is Access to post compulsory secondary education. For the categorical variables the relative frequency is shown. For the numerical ones, the mean is calculated. The calculations are done only using the data from the year 2019 since the period can influence the values.

Figure 6. Source: Own elaboration, using the LFS data for the period 2019.

As already mentioned in the methodology, logistic regression models are estimated and the variables are interpreted independently, looking at the sign and the significance.

Below is the model that best fits the data. Therefore, those variables that did not provide information were ignored.⁹

Logit model of demand for Post compulsory secondary education in Spain					
Independent variables	Categories	Estimate	Std. Error	t value	Pr(> t)
Gender	Female	0,5108	0,021	23,92	9,6E-126
Origin	Spanish nationality and Birthplace Abroad	-0,4486	0,058	-7,72	1,22E-14
	No spanish nationality	-0,8997	0,039	-22,93	9,8E-116
Father's education	Compulsory secondary	0,1924	0,034	5,73	9,93E-09
	Post compulsory secondary	0,7067	0,042	16,86	1,41E-63
	Tertiary	1,0628	0,046	23,11	1,5E-117
Mother's education	Compulsory secondary	0,3223	0,032	9,98	1,96E-23
	Post compulsory secondary	0,8659	0,038	22,60	1,7E-112
	Tertiary	1,4139	0,045	31,74	9,6E-219
Household	Not all member are unemployed	0,1645	0,040	4,08	4,57E-05
	Number individuals	-0,1266	0,011	-11,46	2,35E-30

Category of reference: Male, with Spanish nationality, who lives in a household where both parents have primary education, all the members of the household are unemployed, the period is 2019, the socio-economic level of the head of the family is not high and the autonomous community of residence is Catalonia.

Figure 7. Source: Own elaboration from the LFS data, 2007-2019 period.

According to the results obtained, the variables that appear to be the most significant when it comes to explain why an individual has demanded post-compulsory secondary education, refer to being a woman or to the highest level of education of the mother or of the father. Conversely, the fact that all members of the household are unemployed or the fact that there are many people living in the household have the opposite effect.

Socio-economic variables

If there were causality between economic variables and low academic achievement, this would be a major shortcoming in the principle of equality of opportunities, in the absence of a government policy aimed at correcting it. In this first model, it can be seen that when there is at least one member of the household working (and therefore the household has an inflow of income) the likelihood of access to post-compulsory education increases. Therefore, a significant income effect can be seen. On the contrary, the fact that there are many individuals living in the household, which means more expenses, decreases this probability. However, the socio-economic status of the head of the family turn out to be non-significant (the fact that the head of the family is an entrepreneur or manager does not influence the educational achievement of the children).

⁹ It is recommended to take a look at the appendix to see the complete output. Two important variables, Year and autonomous communities, are not shown in the tables.

Concerning the education of the parents, it is clear that is a very relevant factor. The fact that parental education seems to be the main determinant of children's educational level is worrying because it produces intergenerational persistence in the stock of human capital. This is even worse when the parents' educations are also closely related, i.e. individuals marry people of a similar education (there is selective pairing).

Another important result is that non-Spanish people have much more problems to remain in the educational system than Spaniards.

Contextual variables

Only the autonomous communities of Cantabria, Navarre and the Basque Country increase the probability of accessing the post-compulsory education level. Therefore, Catalonia is in a good position overall. In addition, the dummy of the Autonomous Communities shows, using the F-test, that its interaction with the variable *birthplace and nationality* is informative. Therefore, immigrants demand different educational levels depending on the autonomous community in which they live.

The unemployment rate in the autonomous community has certain influence (but not enough to be statistically significant) on the level of education. As might be expected, when unemployment rises, the likelihood of the individual continuing to study is higher (substitution effect).

Temporal variable

With regard to the time period, it is surprising that no interaction of the variable Year with the other covariates has been significant. In addition, of the contextual variables that evolve with the economic cycle, only the unemployment rate of the autonomous communities has emerged slightly significant. The results are shown in the appendix.

It can be seen, however, that every year except 2016 and 2018 increases the probability of dropping out of school without having reached the post-compulsory secondary level (with respect to the year 2019). This is in line with the fact presented at the introduction that school dropout has decreased considerably in the last decade. However, this model cannot provide much more insights about this trend.

5.1.2. Demand for tertiary education in Spain

As in the previous case, the descriptive table and the econometric model with the coefficients of the independent variables are shown below.

Summary - Year 2019 -Spain

Variable	Categories	Tertiary education	
		No access	Access
Total		41,3%	58,7%
Gender	Male	48,4%	51,6%
	Female	33,9%	66,1%
Nationality and place of birth	Nationality and birth place - Spain	35,3%	64,7%
	Birth place - other	50,6%	49,4%
	Nationality and birth place - other	75,9%	24,1%
Father's educational attainment	Primary	68,4%	31,6%
	Secondary compulsory	48,0%	52,0%
	Secondary post compulsory	32,8%	67,2%
	Tertiary	17,0%	83,0%
	No info.	51,4%	48,6%
Mother's educational attainment	Primary	70,1%	29,9%
	Secondary compulsory	51,9%	48,1%
	Secondary post compulsory	35,7%	64,3%
	Tertiary	17,7%	82,3%
	No info.	60,7%	39,3%
Socio-economic status	Others	41,2%	58,8%
	High	41,7%	58,3%
Unemployment household members	All members are unemployed	51,6%	48,4%
	Not all	40,6%	59,4%
CCAA	Catalunya	42,2%	57,8%
	Andalusia	49,4%	50,6%
	Aragó	37,0%	63,0%
	Astúries	38,2%	61,8%
	Canàries	49,2%	50,8%
	Cantàbria	30,8%	69,2%
	Castella-La Manxa	45,7%	54,3%
	Castella i Lleó	38,2%	61,8%
	Comunitat de Madrid	30,6%	69,4%
	Comunitat Foral de Navarra	35,1%	64,9%
	Comunitat Valenciana	40,0%	60,0%
	Euskadi	27,3%	72,7%
	Extremadura	46,5%	53,5%
	Galícia	36,4%	63,6%
	Illes Balears	56,7%	43,3%
	La Rioja	34,1%	65,9%
Regió de Múrcia	46,8%	53,2%	
	Number people household	3,73	3,71
Regional	GDPpc	24312	25226
	Unemployment rate (16-64)	10,96	10,32
	Tertiary studies rate (16-64)	36,67	38,14

Note: The dependent variable is Access to tertiary education. For the categorical variables the relative frequency is shown. For the numerical ones, the mean is calculated. The calculations are done only using the data from the year 2019 since the period can influence the values.

Figure 8. Source: Own elaboration, using the LFS data for the period 2019.

The results are similar to the same table in the previous section, but some differences can be highlighted. First, the percentage of people between the ages of 20 and 24 who access higher education is fairly balanced (although there are 8% more in the higher education category). Regarding the covariates, the most notable difference is found in the studies of the parents: The frequency of individuals with parents with primary education who access the higher education level is very low (around 30%). Therefore, at first glance, it seems that in higher education, the education of the parents is even more important than in the previous levels.

The most relevant results of the econometric model are the following:

Logit model of demand for tertiary education in Spain					
Independent variables	Categories	Estimate	Std. Error	t value	Pr(> t)
Gender	Female	0,7146	0,017	41,36	< 0
Origin	Spanish nationality and Birthplace Abroad	-0,7843	0,054	-14,64	1,82E-48
	No spanish nationality	-1,5342	0,042	-36,68	8,8E-291
Father's education	Compulsory secondary	0,1596	0,028	5,77	7,84E-09
	Post compulsory secondary	0,7059	0,032	21,97	1,6E-106
	Tertiary	1,3234	0,035	37,95	< 0
Mother's education	Compulsory secondary	0,2903	0,027	10,89	1,45E-27
	Post compulsory secondary	0,8530	0,030	28,41	3,7E-176
	Tertiary	1,7117	0,035	49,42	< 0
Household	Number individuals	-0,1392	0,009	-14,69	8,95E-49
	High socio-economic status	-0,0423	0,023	-1,85	0,064281

Category of reference: Male, with Spanish nationality, who lives in a household where both parents have primary education, all the members of the household are unemployed, the period is 2019, the socio-economic level of the head of the family is not high and the autonomous community of residence is Catalonia.

Figure 9. Source: Own elaboration from the LFS data, 2007-2019 period.

In this econometric model there are four features that I would like to highlight:

As mentioned, here parental education is even more important than in the previous case and is certainly the variable that most clearly influences the dependant variable. Thus when the mother has higher education instead of primary education, the odds of the individual also studying this level of education (probability of accessing the higher level over that of not accessing) increase 5.5 times; as long as the other variables remain constant. Another interesting fact is that in the both models seen the OR of the category *Tertiary education of the mother* is roughly 1.5 points higher than the OR of *Tertiary education of the father*.

The socio-economic level is slightly significant (p-value of 0.064). However, the sign of the coefficient does not have a clear economic interpretation, as it is the opposite of what was expected. It is confusing that individuals living in households where the head of the family is an entrepreneur or manager are less likely to demand higher education. In any case, a possible interpretation is that these individuals live in a very comfortable situation and have no incentives to strive.

The variable *Year* is no longer significant; the coefficients but remain negative with respect to the year 2019. In this case though, its interaction with the unemployment of the household is significant: The effect of this variable on the probability of studying higher education depends on the period considered.

Apart from the dummy *Year* and *Autonomous Community*, none of the others contextual variables is significant. However, although the coefficients are not statistically different from zero, the variables do provide information according to test F. In any case, this could be because individuals have made educational decisions much earlier in time.

With regard to the dummy of the autonomous communities, the F test again shows that its interaction with the variable associated to the place of birth and nationality is informative.

5.2.Catalonia

As we have seen, there are strong differences between autonomous communities that are manifested in the significance of most of the categories of the dummy variable, and in its interactions with the other covariates. The source of these differences is not clear though, it could be the uneven costs of education among communities or the reputation of the universities, to name a few. In any case, it seems that the unemployment rate or the number of people with higher education in the autonomous community are not very relevant.

In this section I will do an analysis similar to what has been done for the Spanish territory. Special emphasis will be placed on the differences that can be seen between Catalonia and the rest of Spain, bearing in mind that care must be taken when statistically interpreting results from two different population groups.

In order to be as short as possible, the descriptive tables have been placed in the appendix.

5.2.1. Demand for post compulsory secondary education in Catalonia

Logit model of demand for Post compulsory secondary education in Catalonia					
Independent variables	Categories	Estimate	Std. Error	t value	Pr(> t)
Gender	Female	0,5608	0,063	8,85	1,1E-18
Origin	Spanish nationality and Birthplace Abroad	-0,2253	0,153	-1,47	0,141866
	No spanish nationality	-0,8532	0,092	-9,32	1,55E-20
Father's education	Compulsory secondary	0,1223	0,107	1,14	0,254726
	Post compulsory secondary	0,7189	0,124	5,81	6,4E-09
	Tertiary	1,0656	0,136	7,82	6,13E-15
Mother's education	Compulsory secondary	0,0528	0,105	0,50	0,616833
	Post compulsory secondary	0,6112	0,112	5,45	5,33E-08
	Tertiary	0,9242	0,126	7,31	2,96E-13
Household	Number individuals	-0,1040	0,029	-3,60	0,000318
Province	Tarragona	-0,2046	0,079	-2,59	0,009667
	Girona	-0,1077	0,078	-1,38	0,167448
	Lleida	-0,0370	0,108	-0,34	0,731996

Category of reference: Male, with Spanish nationality, who lives in a household where both parents have primary education, all the members of the household are unemployed, the period is 2019, the socio-economic level of the head of the family is not high and the province of residence is Barcelona.

Figure 10. Source: Own elaboration from the LFS data, 2007-2019 period.

The relevant variables for the Spanish case continue to be so for Catalonia: gender, birthplace and nationality, studies of the parents and the number of individuals in the household. Instead of autonomous communities, now I consider provinces. Thus, it can be seen how not living in Barcelona reduces the chances of studying post-compulsory education, though the difference is only significant for Tarragona. What is not clear are the factors that cause it: It could be due to issues of mobility (if for example it is more difficult to move to the center of studies), productive structure (you can work in the tourist industry with very little training) or others.

From 2007 to 2013 the temporal variable is significant and has negative coefficients. Moreover, the Year has more influence on the demand for secondary education than for tertiary education, both in Spain and Catalonia. The odds ratios are shown in the next figure. Until 2014 the odds ratios are less than 1, so these year are associated with lower odds of accessing this level of education. This evolution is consistent with what was expected: Nowadays, the probability of studying (any educational level) is clearly higher than back in 2007. However, the causes are not clear. It could be due to the changes that both the economy and society have experienced due to the economic crisis, reduced spending on education or technological change, among others.

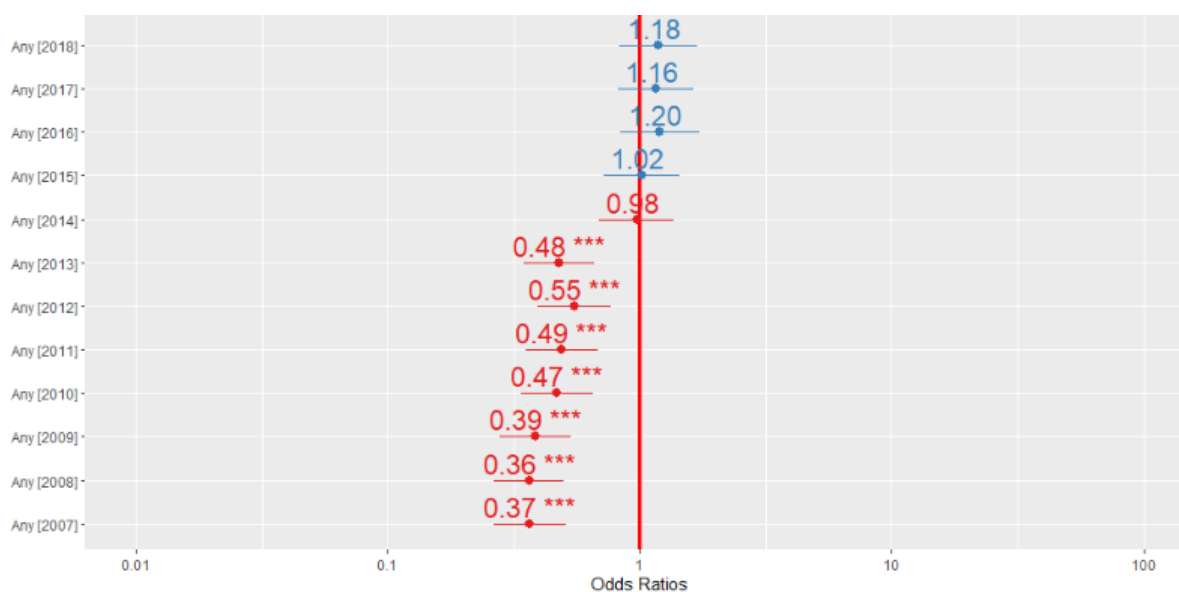


Figure 11. Odds ratio summary of the final model. The confidence intervals of the OR are calculated assuming normality. OR are ordered by year.

5.2.2. Demand for tertiary education in Catalonia

Logit model of demand for tertiary education in Catalonia					
Independent variables	Categories	Estimate	Std. Error	t value	Pr(> t)
Gender	Female	0,6535	0,052	12,65	3,05E-36
	Male				
Origin	Spanish nationality and Birthplace Abroad	-0,9236	0,129	-7,13	1,08E-12
	No spanish nationality	-1,3029	0,087	-14,94	1,32E-49
Father's education	Compulsory secondary	0,2052	0,088	2,33	0,019783
	Post compulsory secondary	0,7976	0,098	8,17	3,65E-16
	Tertiary	1,3104	0,101	12,94	7,57E-38
Mother's education	Compulsory secondary	0,0762	0,085	0,89	0,371147
	Post compulsory secondary	0,5757	0,089	6,50	8,68E-11
	Tertiary	1,2838	0,099	12,98	4,58E-38
Household	Number individuals	-0,1389	0,026	-5,28	1,36E-07
	High socio-economic status	-0,1029	0,064	-1,61	0,107285
Province	Tarragona	-0,2889	0,061	-4,76	1,99E-06
	Girona	-0,2047	0,063	-3,25	0,001154
	Lleida	0,0272	0,091	0,30	0,765489

Category of reference: Male, with Spanish nationality, who lives in a household where both parents have primary education, all the members of the household are unemployed, the period is 2019, the socio-economic level of the head of the family is not high and the province of residence is Barcelona.

Figure 12. Source: Own elaboration from the LFS data, 2007-2019 period.

The results obtained with this model are very similar to the others. In general though, in Catalonia the odds ratio are smaller, less significant than in the Spanish case, though the descriptive tables are very similar.

Overall, it can be seen that in the models dealing with access to higher education the socio-economic variables are very important: The coefficients of the variables parental education, birthplace and nationality of the individual and the number of household members; are very significant. Thus, there is much difference in the probability of studying higher education between a Spanish female, with parents with higher education, and a male immigrant with parents with little education and a large family.

6. Intergenerational educational mobility in Spain

As seen in the previous results, the level of education of the parents is one of the most important factors in determining the level of education that the child will achieve. The influence of this factor has been contrasted in numerous studies over time and can only be compared to the effect that household income had in the 80s. That is why it has been considered appropriate to proceed with a more detailed analysis of this factor, especially with regard to the intergenerational mobility. In this context, intergenerational mobility means the association between the educational attainment of parents and that of their children.

Parents can influence the educational potential of their children in a very diverse way: Providing a stable and suitable environment for learning, facilitating the access to the consumption of cultural goods such as reading or art, through the formation of aspirations and preferences of educational achievement, through investment and financial support of educational activities etc.

In addition, in Spain the influence of parents is probably greater than in other countries due to the strong family relationships. According to Eurostat¹⁰, in Spain children leave the family home on average at the age of 29, three years longer than the European average and almost ten years older than the young Swedes do.

The role of public policies must also be considered. These can significantly affect educational intergenerational mobility, counteracting inequalities caused by the family environment. By public policies one could think of the subsidies of the educational costs (in the form of scholarships), the facilitation to the access of cultural goods, a good orientation in the educational trajectory of the students, quality public educational centers; among many others.

In particular, in this work we want to delve into intergenerational educational mobility, in order to see to what extent there is a certain willingness to perpetuate the achievement of educational levels of origin. It is important to analyze this issue in depth because one of the main arguments used to defend public intervention in education is equal opportunities. Society considers desirable a high degree of intergenerational mobility. The lack of intergenerational educational mobility would highlight a lack of effectiveness of the educational system.

¹⁰ Source: https://ec.europa.eu/eurostat/statistics-explained/index.php/Being_young_in_Europe_today_-_family_and_society. The data can be found in https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=yth_demo_030&lang=en

6.1. Research background

Mobility studies try to determine the processes that condition the intergenerational changes of variables such as wages, income, social status or education. It is about capturing the transformations that take place in society between two consecutive generations.

Below I will mention some of the results that have been found in the topic of the intergenerational educational mobility:

Education systems that are more understanding of students are more conducive to intergenerational mobility than those that are stricter and that encourage students to decide very early which path want to follow.

Educational systems in which children start school at a very early age promote mobility as the influence of the family is reduced. The same can be said of those systems where young people are less dependent on family support networks.

In societies in which individuals are paired with individuals with similar socio-economic characteristics (selective pairing), mobility will be reduced largely.

Another source of the persistence of parents' educational levels on their children is territorial segregation. This can happen for example when the best schools are usually located in affluent neighborhoods as is usually the case in the US. Since only high-income families have access to these neighborhoods, we see how economic inequality indirectly influences education.

Finally, it is necessary to take into account the possible differences in intergenerational educational mobility due to gender. In general, the studies make the differentiating analysis between the father-son, father-daughter, mother-son and mother-daughter relationship. Another factor to consider is the evolution and historical change in these relationships.

6.2. Methodology

Three different methods have been used widely in the literature to study this topic: Correlation analysis, regression analysis and transition matrices. Below is a brief description of each method and how it has been used in the work.

a) Correlation

It serves to measure the relationship between the educational levels of parents and their children.

The correlation between Spearman and Kendall will be used to see the association between numerical variables. These tests are not as popular as the Pearson correlation but have the advantage that they do not assume normality (they are range-based tests).

The Gamma correlation (λ) will be used to see the association between ordinal categorical variables. This measure has two versions: One is symmetrical and the other is not, and therefore the explanatory variable must be distinguished. Symmetrical measurement will be used in this work. The chi-square test is much more popular but has the problem that the value of the statistic is not delimited. Other options would be the correlation of Thiel's U or that of Cramer V.

All statistics that will be used to see the correlation between variables have values between -1 and 1 where 1 indicates maximum association (one variable can be determined from another).

b) Regression analysis

Various regression models can be useful to study the factors influencing mobility. In the literature, the most commonly option is regression to the mean using a Markov model. This regression consists in relating the dependent variable *education achieved by the children* to the explanatory variable *education of the parents*. However, other options are also valid, such as the use of multinomial regressions to model the likelihood that an individual will increase, decrease, or maintain the level of education their parents have. Although these regression models are not used in this section (it may be an interesting proposal for future projects), binary response regressions have already been used in the previous chapter.

c) Transition matrices

In general, transition matrices are used to indicate the probabilities of movement through the different groups of the studied variable, between two successive generations. In particular, the educational transition matrix is a square table of conditioned frequencies. Each element (ij) represents the probability that an individual will access an educational level j if his or her parent belongs to an educational level i . Within this matrix, three areas must be distinguished:

- Immobility: It is the main diagonal of the matrix. It shows the probability that an individual will get the same level of education as their parent.
- Upward mobility: Above the diagonal are the individuals who have exceeded the educational level of their parents.
- Downward mobility: On the contrary, below the diagonal are the individuals who have worsened the educational level of their parents.

6.3. Data

With the LFS data, a whole series of information on the composition of households can be obtained. Thus, one can know for example the number of members of each household, the role of each individual (father, mother, son); the educational level, the age, among others. Two types of variables are used in this section:

- Educational level of son, father and mother. It is a categorical ordinal variable.

- Age at the time the maximum level of education was reached. It is a numerical variable.

Two notes: Parental information is not always available because not everyone may live in the same household or may not be alive at the time of the survey. By education I means formal (official) education.

The LFS could also be useful in analysing variables such as type of industry or type of employment of the parents but also of the children. One might expect parents to influence their children in the type of work they want to do: If the parents are scientists or doctors it seems more likely that the children want to be that too. The same can happen if the parents are entrepreneurs and have a family business, the children will probably collaborate in this business. The downside is that when children start working and stop studying, they usually leave home and then the LFS can't keep track of them. In any case, this work does not follow this line but only focuses on the educational issue.

Some important clarifications need to be made regarding the population under analysis. The LFS only registers information of the individuals who live in that moment in the household. Therefore, it is important to select an age range in which most young people are still living with their parents in order to obtain information from both generations. If for example young people emancipate or parents get divorced, the relationship is lost. At the same time, however, young people must be old enough to have already completed their studies. That is why the most suitable band is 25 to 29 years old. In this band, however, only 80.1% of young people live with one of their parents and therefore a selection of the sample must be made. This selection can create a certain bias if the distribution of young people who emancipate is different from that of young people who continue to live with their parents. In the 20 to 24 age group, 93.9% of young people live with one of their parents but there are many individuals who are still studying.

During the analysis, I will take into account possible differences in mobility due to gender – for both parents and children. Moreover, as the objective of this section is to study the relationship between parents and children, data from the whole period 2007-2019 will be used (regardless of the year). However, to see the educational level of the population, it is better to use only the data of the year 2019.

6.4. Results

First, a basic description of the level of education achieved has been made, similar to that already made in the initial descriptive part of the work. The next table describes the level of education achieved by the population aged between 25 and 29 in 2019, differentiating between parents and children. To have a more complete picture it has been decided to disaggregate the educational categories used so far and add the categories of illiterate, incomplete primary and in terms of post-compulsory secondary is distinguished between Baccaalaureate, which focuses on university access, and professional training, which is education focused on the professional world.

Level of education	Educational attainment				
	Children	Son	Daughter	Father	Mother
Illiterate	0,48%	0,56%	0,39%	0,73%	1,50%
Incomplete primary	0,45%	0,53%	0,36%	3,12%	3,73%
Primary	4,15%	4,72%	3,55%	13,13%	13,00%
ESO	20,66%	24,94%	16,27%	38,34%	35,85%
Baccalaureate	11,46%	12,41%	10,48%	12,86%	11,46%
CFPGM	12,54%	13,23%	11,84%	7,88%	11,08%
Tertiary	50,27%	43,60%	57,11%	23,95%	23,37%
Total	n = 6728 100%	n = 3408 100%	n = 3320 100%	n = 4137 100%	n = 5007 100%

Figure 13.

Consistent with the results obtained so far, in today's youth women demand more education than men do. In contrast, in the previous generation men achieved higher levels of education. It should also be noted that the education of parents is considerably lower than that of their children. Thus, in the LFS sample, more than half of the parents have only reached a basic educational level (up to ESO). Especially large is the intergenerational difference in primary and secondary education.

To see the degree of intergenerational mobility transition matrices are especially useful. Two matrices have been calculated, one relating the father's education to that of his children and the same to the mother's education. The data used are from the period 2007-2019. The results are shown in terms of conditioned frequencies.

Father	Educational attainment children				Total
	Primary	Sec.-comp.	Sec.-non comp.	Tertiary	
Primary	17,4%	31,4%	20,6%	30,6%	100% n = 20166
Secondary-compulsory	4,4%	39,4%	21,8%	34,3%	100% n = 23854
Secondary-non compulsory	3,0%	14,9%	34,2%	47,8%	100% n = 13201
Tertiary	1,3%	7,4%	18,1%	73,2%	100% n = 14148

Mother	Educational attainment children				Total
	Primary	Sec.-comp.	Sec.-non comp.	Tertiary	
Primary	18,7%	31,7%	19,6%	29,9%	100% n = 23177
Secondary-compulsory	3,9%	40,6%	20,6%	34,9%	100% n = 10158
Secondary-non compulsory	2,4%	13,8%	38,4%	45,5%	100% n = 7660
Tertiary	1,2%	6,7%	16,8%	75,3%	100% n = 10554

Figure 14.

The two matrices show very similar results. It can be seen how the most likely case is that the children of parents with higher education have also completed this level of education (73.2% for the case of the father, 75.3% for that of the mother). This point is very important, especially when it comes to implementing public policies. It would be desirable that the probability of

achieving higher education be the same for all individuals, regardless of their family background. Currently, children whose parents have higher education are more than twice as likely to reach this level of education as children whose parents only completed primary education.

For the other categories, there is as well a high probability of an intergenerational increase in the level of education (upward mobility). Of course, the category where children are most likely to improve their parents' level of education is primary education (more than 80% of children will improve this level).

In addition, the Gamma correlation has been calculated and the values obtained are 0.42 for the case of the transition matrix of the fathers and 0.43 for that of the mothers. In both cases, the statistic is highly significant. This indicates two things: First, the relationship between the education of parents and that of children is very significant, that is, not everybody has the same opportunities. Second, there is no difference between the educational influence of the father and the mother. In general, these results are similar to other studies done.

Continuing with the analysis, a summary can be made of the types of mobility that exist: Ascending, descending and unchanged (or immobility).

With respect to the	Ascending mobility	Immobility	Descending mobility
Father	36358 50,9%	27794 38,9%	7217 10,1%
Mother	42649 51,3%	33157 39,9%	7309 8,8%

Figure 15.

There is a great upward mobility regardless of the gender of the parents. The level of education of the second generation is much higher¹¹. With regard to downward mobility, care must be taken when interpreting. Part of this mobility can be explained by early school leaving, especially of children with parents with post-compulsory secondary education, but another part can be explained by the preferences of children of a lower level of education than college. For example, some people may prefer to do a CFPGS that is shorter than a college degree and often has good career opportunities.

Briefly, to conclude the analysis of educational intergenerational mobility, the correlations of the variable "Age at which the highest level of education was reached" were calculated. As in the previous cases, the sample was divided into four categories.

¹¹ Seeing these results it is not surprising that the phrase "This generation is the most educated in history" has become so popular in Spain.

**Correlation coefficients between age at which they reached
the highest educational level**

	2007-2019		2019	
	Kendall	Spearman	Kendall	Spearman
Father-son	0,26	0,34	0,21	0,28
Father-daughter	0,16	0,22	0,16	0,21
Mother-son	0,20	0,27	0,19	0,25
Mother-daughter	0,22	0,28	0,20	0,25

Figure 16.

As a continuous variable, two different types of correlations have been applied, Kendall and Spearman. All values have come out significant. Here you can see how the father has more influence on the son than on the daughter and the mother in reverse. It is not clear that this difference is statistically significant but in any case, it seems plausible, as there are works that are still associated with a particular genre. One can think of mothers who are nurses, primary school teachers or who do household chores, or men who are firefighters or police officers.

7. Conclusions

The present work analyses which variables influence the educational level demanded by young people between 16 and 24 years old. The educational levels covered are post compulsory secondary education and tertiary education and the period analysed goes from 2007 to 2019. The empirical analysis has been done with data from the Labour Force Survey, differentiating the Catalan territory from the Spanish one.

It has been seen that the factors affecting the demand for post-compulsory and higher education are similar to those obtained in the literature on this subject, with some notable nuances:

- Many of these factors are relevant to both access to secondary and higher education. However, the socio-economic variables such as education of the parents or size of the household seem to have even more influence on the demand for tertiary education.

- In Spain, women demand more education than men do, but this is not true in many countries (especially in the underdeveloped ones).

- Nationality and place of birth are also determining factors: foreigners significantly demand less education than nationals do. The integration of this group of the population has to improve yet (although the wave of immigration was earlier than the period analyzed). Furthermore, these factors do not have the same influence in all the autonomous communities.

- Contrary to other studies, the number of individuals living in the household has proved to be relevant in all the cases. As the size of the household increases, the probability of the individual entering secondary and tertiary education decreases. The results suggest that policies address to large families should be taken more seriously.

- Other factors that seemed interesting a priori turned out not to be so, especially with regard to the contextual variables GDPpc, unemployment rate and higher education rate. However, although these variables have not appeared significant in the models, it has been seen that the autonomous community of residence does influence the probability of studying. Not only that, it also modifies the effect of other variables such as the nationality and place of birth.

- With regard to the temporal component of the analysis, some results are consistent with what was expected but others are not. On the one hand, nowadays the probability of studying (any educational level) is clearly higher than back in 2007. On the other hand, it has not been seen that the time period has a significant effect on other explanatory factors.

Since the variables related to the economic cycle are not significant, but the *CCAA* and *Year* are, it seems that there are relevant factors that have not been included in the models. Some of them, related to the autonomous communities, could be spending on education or geographic mobility. Others, related to the time, could be technological change, or changes in the public opinion.

Among all the variables used in this work, however, the education of the parents is especially noteworthy. Therefore, it was decided to make additionally an introductory analysis of

educational intergenerational mobility. The conclusions I take from this analysis are the following:

Human capital has increased very rapidly in the last decade, that is, there has been a very high level of intergenerational upward mobility. The demand for education has continued growing in the last years without being influenced by the economic cycle. However, inequality of opportunities has not changed much: It has been seen that there is a persistent intergenerational immobility in the educational levels achieved, which is further aggravated by the selective pairing of the population. The main problem seen in this work is that the education of the parents is very influential on the level of education that will achieve their children. This should not happen in a society that considers itself meritocratic. The causes, however, are complex and the data do not allow a much more in-depth analysis.

This work contains some relevant contributions, including the use of current data of the LFS, the fact of having used several periods of time and not only a cross-sectional sample, the territorial division by autonomous communities and provinces or the separate analysis of two levels of education.

On the other hand, several extensions could be made. First, it would be interesting to compare the results obtained in this work with those that could be obtained from the Income and Living Conditions Survey (SILC) -considering the same time period-. Additionally, I propose other more complex issues that could also be worth to explore:

- Intergenerational mobility is a broad topic that deals not only with education but also with many other aspects. For example, the LFS is especially adequate to analyse the intergenerational mobility in the labour market.

- It could be interesting to see the effect that the various educational laws introduced over the years have had. In particular, you could think of the change from the LGE to the LOGSE. In this sense, it would also be interesting to include information that has not been added in this work, such as the differences between autonomous communities in educational expenditure.

8. Bibliography

Albert, C. (2000). "Higher Education Demand in Spain: the Influence of Labour Market Signals and Family Background". *Higher Education*, 40(2), 147-162.

Bank of Spain (2019). "Cambios en la decisión de participación laboral de los jóvenes tras la crisis". Informe trimestral de la economía española.

Bank of Spain (2019). "Relación entre el crecimiento del PIB y el del empleo en la economía española". Informe trimestral de la economía española.

Carmen Pérez Esparrells, María Ramos, Marta Adiego, Leonel Cerno (2013). "Principales determinantes del acceso a la educación superior en España: ¿existen diferencias antes y después de la crisis?". XXII Jornadas de la Asociación de Economía de la Educación.

Gil, M., de Pablos, L., y Martínez, M. (2010). "Los determinantes socioeconómicos de la demanda de educación superior en España y la movilidad educativa intergeneracional". *Hacienda Pública Española*, 193(2), 75-108.

Heckman J., James (1979). "Sample Selection Bias as a Specification Error". *Econometrica*, 47(1), 153-161.

Moulton, B.R. (1990). "An illustration of a pitfall in estimating the effects of aggregate variables on micro units". *Review of Economics and Statistics*, 72(2), 334-338.

Rodríguez, C. (1992). "La adquisición de capital humano: un modelo teórico y su contraste. *Investigaciones Económicas*". XVI(2), 305-316.

R., M. C. (2007). "ECONOMÍA LABORAL". Madrid, España: McGraw-Hill.

Salas, M., y Martín-Cobos, M. (2006). "La demanda de educación superior: un análisis microeconómico con datos de corte transversal". *Revista de Educación*, 339, 637-660.

9. Webgraphy

European Commission (2020). “Estructura y Organización del Sistema Educativo”. Retrieved from:

https://eacea.ec.europa.eu/national-policies/eurydice/content/organisation-education-system-and-its-structure-79_es

European Statistical Office (2020). “Participation of young people in education and the labour market”. Retrieved from:

https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Participation_of_young_people_in_education_and_the_labour_market#Participation_of_young_persons_in_formal_education_and_in_the_labour_market

European Statistical Office (2020). “When are they ready to leave the nest?”. Retrieved from:

<https://ec.europa.eu/eurostat/en/web/products-eurostat-news/-/EDN-20190514-1?inheritRedirect=true&redirect=%2Feurostat%2Fen>

Statistical Institute of Spain (2019). “Encuesta de población activa”. Retrieved from:

https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176918&menu=resultados&secc=1254736030639&idp=1254735976595#!tabs-1254736030639

Statistical Institute of Spain (2019). “Contabilidad regional de España”. Retrieved from:

https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736167628&menu=resultados&idp=1254735576581#!tabs-1254736158133

Hoang Anh (2019). “Correlation between discrete (categorical) variables”. Retrieved from:

<https://rpubs.com/hoanganhngo610/558925>

10. Appendix

Demand for post-compulsory secondary education in Spain

Output of the logit models

Initial or Full model:

Final model

Survey design:
svydesign(ids = ~NVIVI, weights = ~FACTOREL, data = eso)

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	6.892e-01	9.754e-01	0.707	0.47984
GenereDona	5.109e-01	2.135e-02	23.930	< 2e-16 ***
NaixNacioNacio Espanya naix estranger	-4.488e-01	5.816e-02	-7.716	1.23e-14 ***
NaixNacioNo nacionalitat	-9.005e-01	3.923e-02	-22.957	< 2e-16 ***
estudis_pareSecundària Obligatòria	1.917e-01	3.359e-02	5.707	1.16e-08 ***
estudis_pareSecundària Postobligatòria	7.067e-01	4.190e-02	16.867	< 2e-16 ***
estudis_pareSuperior	1.064e+00	4.602e-02	23.127	< 2e-16 ***
estudis_pareno info.	-1.907e-01	3.972e-02	-4.801	1.58e-06 ***
estudis_mareSecundària Obligatòria	3.213e-01	3.232e-02	9.940	< 2e-16 ***
estudis_mareSecundària Postobligatòria	8.654e-01	3.833e-02	22.578	< 2e-16 ***
estudis_mareSuperior	1.413e+00	4.453e-02	31.740	< 2e-16 ***
estudis_mareno info.	-1.576e-01	6.051e-02	-2.605	0.00919 ***
atur_llarAlgun ocupat	1.625e-01	4.049e-02	4.014	5.98e-05 ***
num_indiv	-1.263e-01	1.106e-02	-11.416	< 2e-16 ***
socio_eco_hfAlt	2.351e-02	2.762e-02	0.851	0.39463
taxa_atur	1.783e-02	1.008e-02	1.769	0.07686 .
taxa_estudis_sup	1.590e-02	1.229e-02	1.293	0.19585
PIBpc	-1.157e-05	3.208e-05	-0.361	0.71841
CCAAAndalusia	-5.163e-01	3.073e-01	-1.680	0.09300 .
CCAAAragó	-2.707e-01	9.452e-02	-2.864	0.00418 **
CCAAAstúries	-2.097e-01	2.276e-01	-0.921	0.35684
CCAACanàries	-2.626e-01	2.418e-01	-1.086	0.27744 .
CCAACantàbria	1.380e-01	2.129e-01	0.648	0.51685
CCAACastella-La Manxa	-4.046e-01	2.782e-01	-1.454	0.14583
CCAACastella i Lleó	-2.574e-01	1.886e-01	-1.365	0.17229 .
CCAAComunitat de Madrid	-2.858e-01	1.542e-01	-1.853	0.06383 .
CCAAComunitat Foral de Navarra	1.258e-01	1.050e-01	1.198	0.23083
CCAAComunitat Valenciana	-3.243e-01	2.197e-01	-1.476	0.13994 .
CCAAEuskadi	1.787e-01	1.463e-01	1.221	0.22206
CCAAExtremadura	-3.915e-01	3.405e-01	-1.150	0.25027 .
CCAAGalícia	-9.232e-02	2.201e-01	-0.419	0.67487
CCAAIlles Balears	-3.118e-01	1.328e-01	-2.347	0.01891 *
CCAALa Rioja	-2.624e-01	1.171e-01	-2.241	0.02501 *
CCAARegió de Múrcia	-2.409e-01	2.550e-01	-0.945	0.34482
Any2007	-7.921e-01	1.431e-01	-5.534	3.14e-08 ***
Any2008	-8.704e-01	1.203e-01	-7.235	4.72e-13 ***
Any2009	-8.762e-01	1.197e-01	-7.321	2.49e-13 ***
Any2010	-8.010e-01	1.160e-01	-6.904	5.13e-12 ***
Any2011	-7.420e-01	1.245e-01	-5.962	2.51e-09 ***
Any2012	-8.198e-01	1.474e-01	-5.562	2.68e-08 ***
Any2013	-8.004e-01	1.497e-01	-5.347	8.96e-08 ***
Any2014	-2.501e-01	1.415e-01	-1.768	0.07706 .
Any2015	-1.919e-01	1.120e-01	-1.713	0.08670 .
Any2016	-2.562e-02	9.316e-02	-0.275	0.78331
Any2017	-8.359e-02	7.295e-02	-1.146	0.25186
Any2018	1.901e-02	6.601e-02	0.288	0.77334

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Survey design:
svydesign(ids = ~NVIVI, weights = ~FACTOREL, data = eso)

Coefficients:

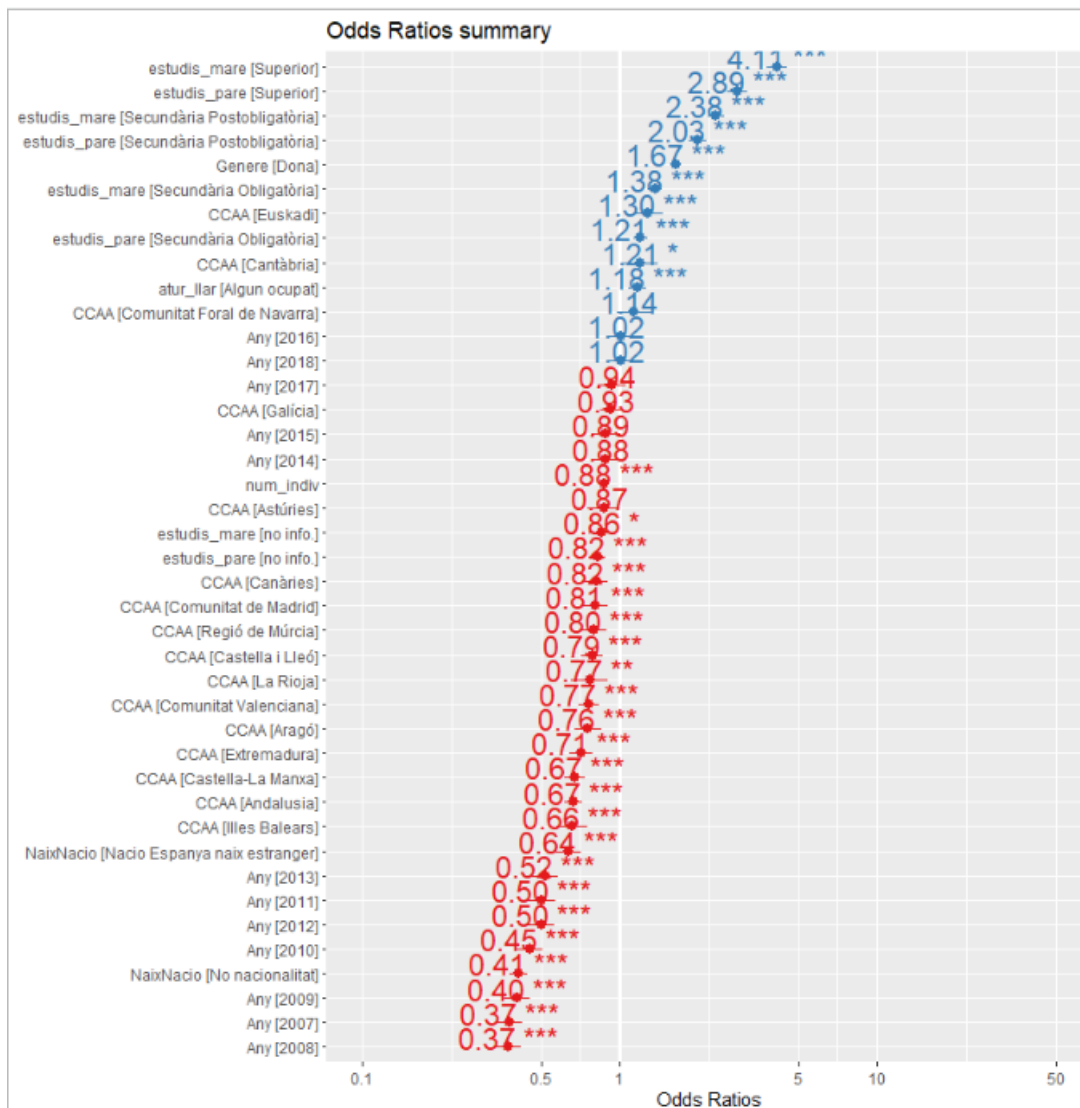
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.14382	0.08500	13.457	< 2e-16 ***
GenereDona	0.51081	0.02135	23.924	< 2e-16 ***
NaixNacioNacio Espanya naix estranger	-0.44857	0.05813	-7.717	1.22e-14 ***
NaixNacioNo nacionalitat	-0.89974	0.03924	-22.931	< 2e-16 ***
estudis_pareSecundària Obligatòria	0.19241	0.03356	5.733	9.93e-09 ***
estudis_pareSecundària Postobligatòria	0.70674	0.04192	16.858	< 2e-16 ***
estudis_pareSuperior	1.06283	0.04598	23.114	< 2e-16 ***
estudis_pareno info.	-0.19450	0.03941	-4.935	8.03e-07 ***
estudis_mareSecundària Obligatòria	0.32227	0.03229	9.980	< 2e-16 ***
estudis_mareSecundària Postobligatòria	0.86587	0.03831	22.601	< 2e-16 ***
estudis_mareSuperior	1.41386	0.04455	31.737	< 2e-16 ***
estudis_mareno info.	-0.15585	0.06051	-2.576	0.010010 *
atur_llarAlgun ocupat	0.16451	0.04035	4.077	4.57e-05 ***
num_indiv	-0.12661	0.01105	-11.458	< 2e-16 ***
CCAAAndalusia	-0.40391	0.04018	-10.052	< 2e-16 ***
CCAAAragó	-0.27577	0.06132	-4.497	6.91e-06 ***
CCAAAstúries	-0.13709	0.07151	-1.917	0.055260 .
CCAACanàries	-0.19947	0.05462	-3.652	0.000261 ***
CCAACantàbria	0.18886	0.08337	2.265	0.023505 *
CCAACastella-La Manxa	-0.39544	0.04663	-8.481	< 2e-16 ***
CCAACastella i Lleó	-0.23849	0.04846	-4.921	8.64e-07 ***
CCAAComunitat de Madrid	-0.21205	0.05560	-3.814	0.000137 ***
CCAAComunitat Foral de Navarra	0.13452	0.08045	1.672	0.094509 .
CCAAComunitat Valenciana	-0.26753	0.04850	-5.516	3.49e-08 ***
CCAAEuskadi	0.26144	0.06827	3.829	0.000129 ***
CCAAExtremadura	-0.33846	0.05519	-6.132	8.74e-10 ***
CCAAGalícia	-0.07738	0.04662	-1.660	0.096996 .
CCAAIlles Balears	-0.41654	0.06948	-5.995	2.04e-09 ***
CCAALa Rioja	-0.26232	0.08340	-3.145	0.001661 ***
CCAARegió de Múrcia	-0.21952	0.05808	-3.780	0.000157 ***
Any2007	-0.98472	0.05808	-16.955	< 2e-16 ***
Any2008	-0.99116	0.05804	-17.078	< 2e-16 ***
Any2009	-0.91003	0.05815	-15.650	< 2e-16 ***
Any2010	-0.79614	0.05889	-13.519	< 2e-16 ***
Any2011	-0.68928	0.05998	-11.491	< 2e-16 ***
Any2012	-0.69765	0.05995	-11.636	< 2e-16 ***
Any2013	-0.66172	0.06038	-10.959	< 2e-16 ***
Any2014	-0.12500	0.06456	-1.936	0.052858 .
Any2015	-0.11512	0.06456	-1.783	0.074571 .
Any2016	0.02039	0.06491	0.314	0.753379 .
Any2017	-0.06606	0.06309	-1.047	0.295073
Any2018	0.01578	0.06420	0.246	0.805839

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Category of reference: Male, with Spanish nationality, who lives in a household where both parents have primary education, all the members of the household are unemployed, the period is 2019, the socio-economic level of the head of the family is not high and the autonomous community of residence is Catalonia.

Source: Own elaboration from the LFS data, 2007-2019 period.

Odds ratio summary of the final model including the dummy variables. The confidence intervals of the OR are calculated assuming normality. OR are ordered decreasingly.



Results of the Test F for the variables of the Final model:

```
Anova table: (Rao-Scott LRT)
svyglm(formula = ABAN ~ Genere, design = my_design2, family = quasibinomial())
      stats      DEff      df      ddf      p
Genere    777.6069    1.6142    1.0000  47293 <2e-16 ***
NaixNacio 2839.0133    2.2615    2.0000  47291 <2e-16 ***
estudis_pare 6766.0408    1.8049    4.0000  47287 <2e-16 ***
estudis_mare 3556.9789    1.7879    4.0000  47282 <2e-16 ***
atur_llar      2.2582    1.6184    1.0000  47281  0.2398
num_indiv    295.2601    2.1110    1.0000  47280 <2e-16 ***
CCAA        397.3309    1.2720   16.0000  47264 <2e-16 ***
Any        2249.8179    1.7449   12.0000  47252 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Demand for tertiary education in Spain

Output of the logit models

Full model:

Final model:

Survey design:
svydesign(ids = ~NVIVI, weights = ~FACTOREL, data = superior)

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	7.207e-01	7.648e-01	0.942	0.346025
GeneroDona	7.143e-01	1.728e-02	41.348	< 2e-16 ***
NaixNacioNacio Espanya naix estranger	-7.819e-01	5.356e-02	-14.597	< 2e-16 ***
NaixNacioNo nacionalitat	-1.533e+00	4.194e-02	-36.562	< 2e-16 ***
estudis_pareSecundària Obligatòria	1.601e-01	2.765e-02	5.790	7.08e-09 ***
estudis_pareSecundària Postobligatòria	7.061e-01	3.216e-02	21.956	< 2e-16 ***
estudis_pareSuperior	1.324e+00	3.487e-02	37.968	< 2e-16 ***
estudis_pareno info.	-8.652e-02	3.188e-02	-2.714	0.006653 **
estudis_mareSecundària Obligatòria	2.894e-01	2.667e-02	10.852	< 2e-16 ***
estudis_mareSecundària Postobligatòria	8.533e-01	3.001e-02	28.433	< 2e-16 ***
estudis_mareSuperior	1.714e+00	3.463e-02	49.486	< 2e-16 ***
estudis_mareno info.	4.770e-03	4.154e-02	0.115	0.908580
atur_llarAlgun ocupat	-1.052e-01	3.274e-02	-3.214	0.001308 **
num_indiv	-1.398e-01	9.461e-03	-14.777	< 2e-16 ***
socio_eco_hfAlt	-4.359e-02	2.288e-02	-1.906	0.056710 .
taxa_atur	-1.319e-02	8.061e-03	-1.636	0.101762 .
taxa_estudis_sup	-1.193e-02	9.968e-03	-1.197	0.231451 .
PIBpc	-7.021e-06	2.490e-05	-0.282	0.777999
CCAAAndalusia	-4.176e-01	2.423e-01	-1.723	0.084853 .
CCAAAragó	7.217e-03	7.479e-02	0.097	0.923123
CCAAAstúries	-1.753e-03	1.780e-01	-0.010	0.992142
CCAACanàries	-3.631e-01	1.944e-01	-1.867	0.061844 .
CCAACantàbria	-4.596e-02	1.637e-01	-0.281	0.778902
CCAACastella-La Manxa	-2.871e-01	2.194e-01	-1.309	0.190699
CCAAComunitat Foral de Navarra	-1.037e-01	1.462e-01	-0.709	0.478284
CCAAComunitat de Madrid	1.060e-01	1.236e-01	0.857	0.391186
CCAAComunitat Valenciana	3.643e-01	8.274e-02	4.403	1.07e-05 ***
CCAAEuskadi	-1.629e-01	1.723e-01	-0.946	0.344330
CCAAExtremadura	4.594e-01	1.162e-01	3.953	7.71e-05 ***
CCAAGalícia	-2.759e-01	2.665e-01	-1.035	0.300643
CCAAIlles Balears	-6.263e-02	1.704e-01	-0.368	0.713168
CCAAIlla de Mallorca	-5.796e-01	1.085e-01	-5.342	9.24e-08 ***
CCAAIlla de Menorca	-9.869e-02	9.361e-02	-1.054	0.291751 .
CCAARegió de Múrcia	-4.053e-01	2.009e-01	-2.017	0.043652 *
Any2007	-5.115e-01	1.140e-01	-4.486	7.26e-06 ***
Any2008	-4.773e-01	9.705e-02	-4.918	8.75e-07 ***
Any2009	-4.176e-01	9.679e-02	-4.315	1.60e-05 ***
Any2010	-3.243e-01	9.391e-02	-3.453	0.000555 ***
Any2011	-2.818e-01	1.005e-01	-2.805	0.005038 **
Any2012	-2.014e-01	1.187e-01	-1.697	0.089611 .
Any2013	-1.556e-01	1.203e-01	-1.293	0.196122
Any2014	-1.639e-01	1.114e-01	-1.471	0.141288
Any2015	-1.526e-01	8.780e-02	-1.738	0.082296 .
Any2016	-5.539e-02	7.260e-02	-0.763	0.445523
Any2017	-4.250e-02	5.783e-02	-0.735	0.462439
Any2018	-7.978e-03	5.107e-02	-0.156	0.875863

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

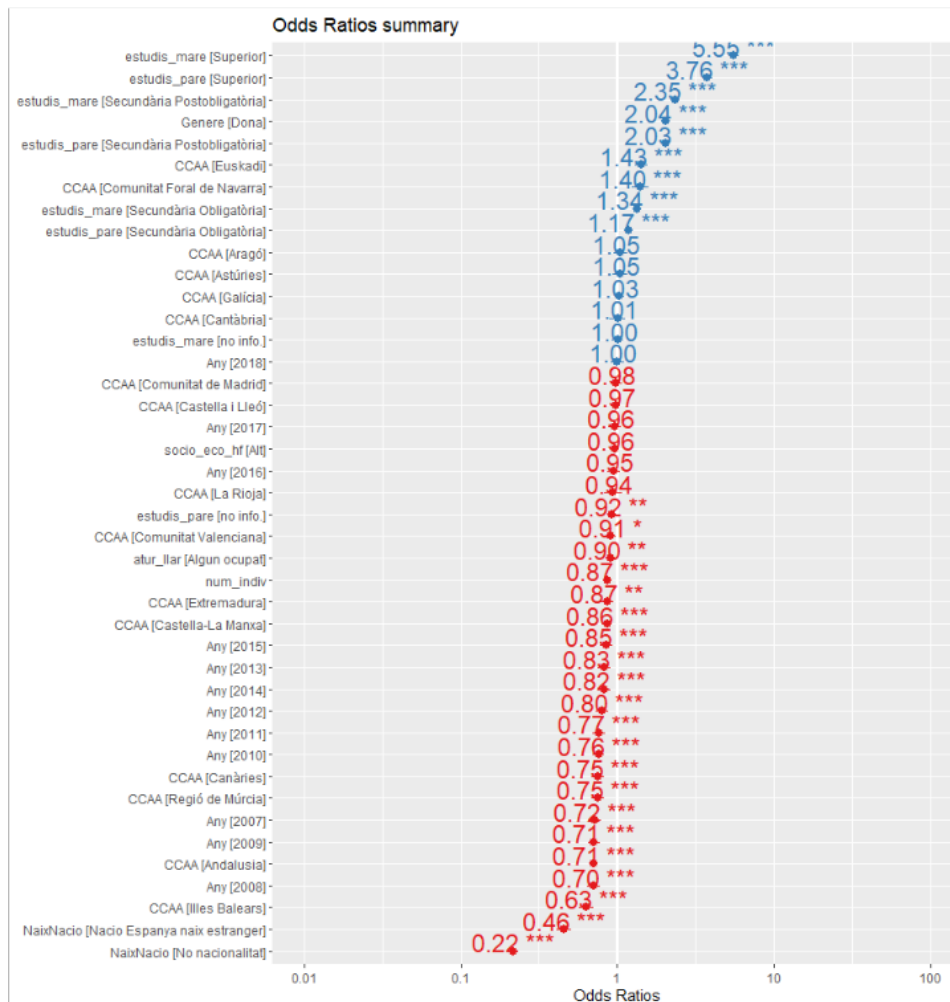
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.221314	0.129183	-1.713	0.086685 .
GeneroDona	0.714572	0.017278	41.357	< 2e-16 ***
NaixNacioNacio Espanya naix estranger	-0.784261	0.053553	-14.645	< 2e-16 ***
NaixNacioNo nacionalitat	-1.534233	0.041828	-36.680	< 2e-16 ***
estudis_pareSecundària Obligatòria	0.159588	0.027645	5.773	7.84e-09 ***
estudis_pareSecundària Postobligatòria	0.705868	0.032124	21.973	< 2e-16 ***
estudis_pareSuperior	1.323445	0.034871	37.952	< 2e-16 ***
estudis_pareno info.	-0.086302	0.031893	-2.706	0.006813 **
estudis_mareSecundària Obligatòria	0.290286	0.026668	10.885	< 2e-16 ***
estudis_mareSecundària Postobligatòria	0.853024	0.030027	28.408	< 2e-16 ***
estudis_mareSuperior	1.711747	0.034638	49.417	< 2e-16 ***
estudis_mareno info.	0.004869	0.041519	0.117	0.906638
atur_llarAlgun ocupat	0.030857	0.125038	0.247	0.805080
Any2007	0.365486	0.211418	1.729	0.083862 .
Any2008	0.019629	0.168326	0.117	0.907169
Any2009	-0.161705	0.158535	-1.020	0.307736
Any2010	0.050306	0.152412	-0.330	0.741350
Any2011	0.052198	0.158937	0.328	0.742597
Any2012	-0.108146	0.146449	-0.738	0.460244
Any2013	-0.168084	0.149187	-1.107	0.281099
Any2014	-0.206865	0.159828	-1.294	0.195570
Any2015	-0.263366	0.152727	-1.724	0.084637 .
Any2016	0.061901	0.159763	0.387	0.698423
Any2017	-0.059541	0.170969	-0.348	0.727647
Any2018	0.066306	0.170205	0.390	0.696860
num_indiv	-0.139186	0.009473	-14.693	< 2e-16 ***
socio_eco_hfAlt	-0.042338	0.022882	-1.850	0.064281 .
CCAAAndalusia	-0.342505	0.032996	-10.380	< 2e-16 ***
CCAAAragó	0.044844	0.050084	0.895	0.370592
CCAAAstúries	0.042098	0.056669	0.743	0.457561
CCAACanàries	-0.286815	0.044795	-6.403	1.54e-10 ***
CCAACantàbria	0.009146	0.057766	0.158	0.874201
CCAACastella-La Manxa	-0.150979	0.038835	-3.888	0.000101 ***
CCAACastella i Lleó	-0.027961	0.038388	-0.728	0.466389
CCAAComunitat de Madrid	-0.017713	0.045184	-0.392	0.695043
CCAAComunitat Foral de Navarra	0.338834	0.062323	5.437	5.45e-08 ***
CCAAComunitat Valenciana	-0.093548	0.039937	-2.342	0.019163 *
CCAAEuskadi	0.356568	0.049742	7.168	7.69e-13 ***
CCAAExtremadura	-0.142685	0.045764	-3.118	0.001823 **
CCAAGalícia	0.029067	0.037103	0.783	0.433401
CCAAIlles Balears	-0.462178	0.056769	-8.141	3.99e-16 ***
CCAAIlla de Mallorca	-0.058571	0.068507	-0.855	0.392572
CCAAIlla de Menorca	-0.290883	0.048183	-6.037	1.58e-09 ***
atur_llarAlgun ocupat:Any2007	0.729928	0.216482	3.372	0.000747 ***
atur_llarAlgun ocupat:Any2008	-0.395760	0.175113	-2.260	0.023824 *
atur_llarAlgun ocupat:Any2009	-0.193048	0.165754	-1.165	0.244160
atur_llarAlgun ocupat:Any2010	-0.244403	0.159696	-1.530	0.125917
atur_llarAlgun ocupat:Any2011	-0.347944	0.166456	-2.090	0.036595 *
atur_llarAlgun ocupat:Any2012	-0.126307	0.154712	-0.816	0.414274
atur_llarAlgun ocupat:Any2013	-0.025895	0.157660	-0.164	0.869536
atur_llarAlgun ocupat:Any2014	0.020872	0.167581	0.125	0.900879
atur_llarAlgun ocupat:Any2015	0.115556	0.160842	0.718	0.472488
atur_llarAlgun ocupat:Any2016	-0.124105	0.167617	-0.740	0.459055
atur_llarAlgun ocupat:Any2017	0.028526	0.178144	0.160	0.872778
atur_llarAlgun ocupat:Any2018	-0.067954	0.177490	-0.383	0.701826

Category of reference: Male, with Spanish nationality, who lives in a household where both parents have primary education, all the members of the household are unemployed, the period is 2019, the socio-economic level of the head of the family is not high and the autonomous community of residence is Catalonia.

Source: Own elaboration from the LFS data, 2007-2019 period.

Odds ratio summary of the final model (without interactions) including the dummy variables. The confidence intervals of the OR are calculated assuming normality. OR are ordered decreasingly.



Results of the Test F for the variables of the Final model:

```
Anova table: (Rao-Scott LRT)
svyglm(formula = SUP ~ Genere, design = my_design1, family = quasibinomial())
      stats      DEff      df      ddf      p
Genere 1803.8515  1.5428  1.0000 51530 < 2.2e-16 ***
NaixNacio 7239.2390  2.4288  2.0000 51528 < 2.2e-16 ***
estudis_pare 12548.9783  1.6681  4.0000 51524 < 2.2e-16 ***
estudis_mare 6467.7183  1.7020  4.0000 51520 < 2.2e-16 ***
atur_llar 38.1782  1.5819  1.0000 51519 9.842e-07 ***
Any 305.4660  1.6650 12.0000 51507 < 2.2e-16 ***
num_indiv 457.2312  1.7816  1.0000 51506 < 2.2e-16 ***
socio_eco_hf 3.6312  1.6738  1.0000 51505 0.143417
CCAA 678.6287  1.1927 16.0000 51489 < 2.2e-16 ***
atur_llar:Any 59.4498  1.5810 12.0000 51477 0.000252 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Demand for post compulsory secondary education in Catalonia

Descriptive table of relative frequencies

Relative frequencies table - Year 2019

Variable	Categories	Secondary post compulsory education	
		No access	Access
Total		15,6%	84,4%
Gender	Male	19,9%	80,1%
	Female	11,0%	89,0%
Nacionality and place of birth	Nationality and birth place - Spain	12,9%	87,1%
	Birth place - other	22,6%	77,4%
	Nationality and birth place - other	29,4%	70,6%
Father's educational attainment	Primary	38,0%	62,0%
	Secondary compulsory	18,9%	81,1%
	Secondary post compulsory	11,3%	88,7%
	Tertiary	4,1%	95,9%
	No info.	20,6%	79,4%
Mother's educational attainment	Primary	26,7%	73,3%
	Secondary compulsory	24,3%	75,7%
	Secondary post compulsory	16,3%	83,7%
	Tertiary	5,1%	94,9%
	No info.	25,5%	74,5%
Socio-economic status	Others	15,7%	84,3%
	High	15,3%	84,7%
Unemployment household members	All members are unemployed	25,9%	74,1%
	Not all	15,2%	84,8%
Province	Barcelona	15,0%	85,0%
	Tarragona	16,3%	83,7%
	Girona	19,1%	80,9%
	Lleida	15,1%	84,9%

Output of the logit models

Full model:

Coefficients:	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.297170	0.381432	3.401	0.000676 ***
GenereDona	0.559381	0.063387	8.825	< 2e-16 ***
NaixNacioNacio Espanya naix estranger	-0.226479	0.153256	-1.478	0.139521
NaixNacioNo nacionalitat	-0.856832	0.091548	-9.359	< 2e-16 ***
estudis_pareSecundària Obligatòria	0.113832	0.107617	1.058	0.290213
estudis_pareSecundària Postobligatòria	0.712994	0.123314	5.782	7.77e-09 ***
estudis_pareSuperior	1.057843	0.136305	7.761	9.90e-15 ***
estudis_pareno info.	-0.101378	0.116413	-0.871	0.383875
estudis_mareSecundària Obligatòria	0.042651	0.105850	0.403	0.687011
estudis_mareSecundària Postobligatòria	0.597589	0.112776	5.299	1.21e-07 ***
estudis_mareSuperior	0.907186	0.126998	7.143	1.02e-12 ***
estudis_mareno info.	-0.674335	0.168052	-4.013	6.08e-05 ***
atur_llarAlgun ocupat	0.209688	0.138538	1.514	0.130186
num_indiv	-0.110173	0.029485	-3.737	0.000188 ***
socio_eco_hfAlt	0.020958	0.077494	0.270	0.786827
taxa_atur_prov	-0.008762	0.039594	-0.221	0.824872
ProvinciaTarragona	-0.186838	0.100963	-1.851	0.064282 .
ProvinciaGirona	-0.097047	0.094067	-1.032	0.302262
ProvinciaLleida	-0.062233	0.146766	-0.424	0.671562
Any2018	0.176167	0.190849	0.923	0.356007
Any2017	0.165072	0.197750	0.835	0.403893
Any2016	0.221389	0.233283	0.949	0.342651
Any2015	0.077132	0.309743	0.249	0.803354
Any2014	0.050538	0.364955	0.138	0.889868
Any2013	-0.651325	0.421052	-1.547	0.121942
Any2012	-0.494192	0.447934	-1.103	0.269956
Any2011	-0.636369	0.372613	-1.708	0.087717 .
Any2010	-0.703549	0.293524	-2.397	0.016566 *
Any2009	-0.907811	0.265123	-3.424	0.000621 ***
Any2008	-1.007658	0.174341	-5.780	7.87e-09 ***
Any2007	-1.040563	0.204386	-5.091	3.67e-07 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Final model:

Coefficients:	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.39352	0.20966	6.646	3.28e-11 ***
GenereDona	0.56079	0.06333	8.855	< 2e-16 ***
NaixNacioNacio Espanya naix estranger	-0.22530	0.15336	-1.469	0.141866
NaixNacioNo nacionalitat	-0.85321	0.09151	-9.324	< 2e-16 ***
estudis_pareSecundària Obligatòria	0.12232	0.10739	1.139	0.254726
estudis_pareSecundària Postobligatòria	0.71891	0.12364	5.814	6.40e-09 ***
estudis_pareSuperior	1.06564	0.13624	7.822	6.13e-15 ***
estudis_pareno info.	-0.11000	0.11521	-0.955	0.339701
estudis_mareSecundària Obligatòria	0.05275	0.10543	0.500	0.616833
estudis_mareSecundària Postobligatòria	0.61119	0.11221	5.447	5.33e-08 ***
estudis_mareSuperior	0.92421	0.12638	7.313	2.96e-13 ***
estudis_mareno info.	-0.68681	0.16680	-4.118	3.88e-05 ***
num_indiv	-0.10403	0.02887	-3.603	0.000318 ***
ProvinciaTarragona	-0.20459	0.07904	-2.588	0.009667 **
ProvinciaGirona	-0.10769	0.07800	-1.381	0.167448
ProvinciaLleida	-0.03702	0.10809	-0.342	0.731996
Any2018	0.16804	0.18090	0.929	0.352975
Any2017	0.14750	0.17795	0.829	0.407190
Any2016	0.18581	0.18390	1.010	0.312363
Any2015	0.01741	0.17646	0.099	0.921393
Any2014	-0.02508	0.17354	-0.145	0.885084
Any2013	-0.73793	0.16470	-4.481	7.59e-06 ***
Any2012	-0.59775	0.16939	-3.529	0.000421 ***
Any2011	-0.71440	0.16822	-4.247	2.20e-05 ***
Any2010	-0.75597	0.16666	-4.536	5.85e-06 ***
Any2009	-0.95087	0.16625	-5.720	1.12e-08 ***
Any2008	-1.00876	0.16299	-6.189	6.47e-10 ***
Any2007	-1.00529	0.16799	-5.984	2.30e-09 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for quasibinomial family taken to be 1.000459)

Category of reference: Male, with Spanish nationality, who lives in a household where both parents have primary education, all the members of the household are unemployed, the period is 2019, the socio-economic level of the head of the family is not high and the province of residence is Barcelona.

Source: Own elaboration from the LFS data, 2007-2019 period.

Odds ratio summary of the final model (without interactions) including the dummy variables. The confidence intervals of the OR are calculated assuming normality. OR are ordered decreasingly.



Results of the Test F for the variables of the Final model:

```
Anova table: (Rao-Scott LRT)
svyglm(formula = ABAN ~ Genre, design = my_design4, family = quasibinomial())
      stats      DEff      df  ddf      p
Genre      87.0091    1.3719    1.0000  5858  2.224e-15 ***
NaixNacio  523.1413    1.6239    2.0000  5856 < 2.2e-16 ***
estudis_pare 529.6348    1.5140    4.0000  5852 < 2.2e-16 ***
estudis_mare 231.4660    1.4168    4.0000  5848 < 2.2e-16 ***
num_indiv    24.1674    1.6571    1.0000  5847  0.0001457 ***
Provincia     5.1734    0.8657    3.0000  5844  0.1123121
Any          291.7400    1.4488   12.0000  5832 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Demand for Tertiary education in Catalonia

Descriptive table of relative frequencies

Relative frequencies table - Year 2019

Variable	Categories	Tertiary education	
		No access	Access
Total		42,2%	57,8%
Gender	Male	51,7%	48,3%
	Female	32,3%	67,7%
Nacionality and place of birth	Nacionality and birth place - Spain	33,0%	67,0%
	Birth place - other	55,1%	44,9%
	Nacionality and birth place - other	77,3%	22,7%
Father's educational attainment	Primary	79,1%	20,9%
	Secondary compulsory	48,9%	51,1%
	Secondary post compulsory	33,8%	66,2%
	Tertiary	17,8%	82,2%
	No info.	51,4%	48,6%
Mother's educational attainment	Primary	70,6%	29,4%
	Secondary compulsory	60,4%	39,6%
	Secondary post compulsory	32,4%	67,6%
	Tertiary	19,8%	80,2%
	No info.	61,9%	38,1%
Socio-economic status	Others	39,6%	60,4%
	High	51,1%	48,9%
Unemployment household members	All members are unemployed	46,7%	53,3%
	Not all	42,1%	57,9%
Province	Barcelona	41,4%	58,6%
	Tarragona	43,6%	56,4%
	Girona	45,9%	54,1%
	Lleida	43,0%	57,0%

Output of the logit models

Full model:

Coefficients:	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.08955	0.31457	0.285	0.77589
GenereDona	0.65288	0.05177	12.611	< 2e-16 ***
NaixNacioEspanya naix estranger	-0.92170	0.12959	-7.113	1.26e-12 ***
NaixNacioNo nacionalitat	-1.30254	0.08723	-14.932	< 2e-16 ***
estudis_pareSecundària Obligatòria	0.20531	0.08806	2.332	0.01975 *
estudis_pareSecundària Postobligatòria	0.79758	0.09764	8.169	3.72e-16 ***
estudis_pareSuperior	1.31064	0.10129	12.940	< 2e-16 ***
estudis_pareno info.	0.10041	0.09359	1.073	0.28337
estudis_mareSecundària Obligatòria	0.07563	0.08534	0.886	0.37550
estudis_mareSecundària Postobligatòria	0.57473	0.08872	6.478	9.98e-11 ***
estudis_mareSuperior	1.28272	0.09909	12.945	< 2e-16 ***
estudis_mareno info.	-0.30402	0.11467	-2.651	0.00804 **
atur_llarAlgun ocupat	0.02472	0.10859	0.228	0.81994
num_indiv	-0.13938	0.02647	-5.265	1.45e-07 ***
socio_eco_hfAlt	-0.10260	0.06427	-1.596	0.11045
taxa_atur_prov	-0.02287	0.03265	-0.700	0.48366
ProvinciaTarragona	-0.25185	0.08115	-3.104	0.00192 **
ProvinciaGirona	-0.17810	0.07446	-2.392	0.01679 *
ProvinciaLleida	-0.03080	0.12548	-0.245	0.80608
Any2018	0.10067	0.14597	0.690	0.49041
Any2017	0.13208	0.15696	0.841	0.40010
Any2016	0.08372	0.18201	0.460	0.64557
Any2015	0.11146	0.24602	0.453	0.65052
Any2014	0.21237	0.29127	0.729	0.46597
Any2013	0.19111	0.34790	0.549	0.58279
Any2012	0.20743	0.36708	0.565	0.57203
Any2011	-0.01578	0.30455	-0.052	0.95868
Any2010	0.05789	0.23830	0.243	0.80806
Any2009	-0.20175	0.21522	-0.937	0.34857
Any2008	-0.25947	0.14940	-1.737	0.08247 *
Any2007	-0.39673	0.17034	-2.329	0.01989 *

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Final model:

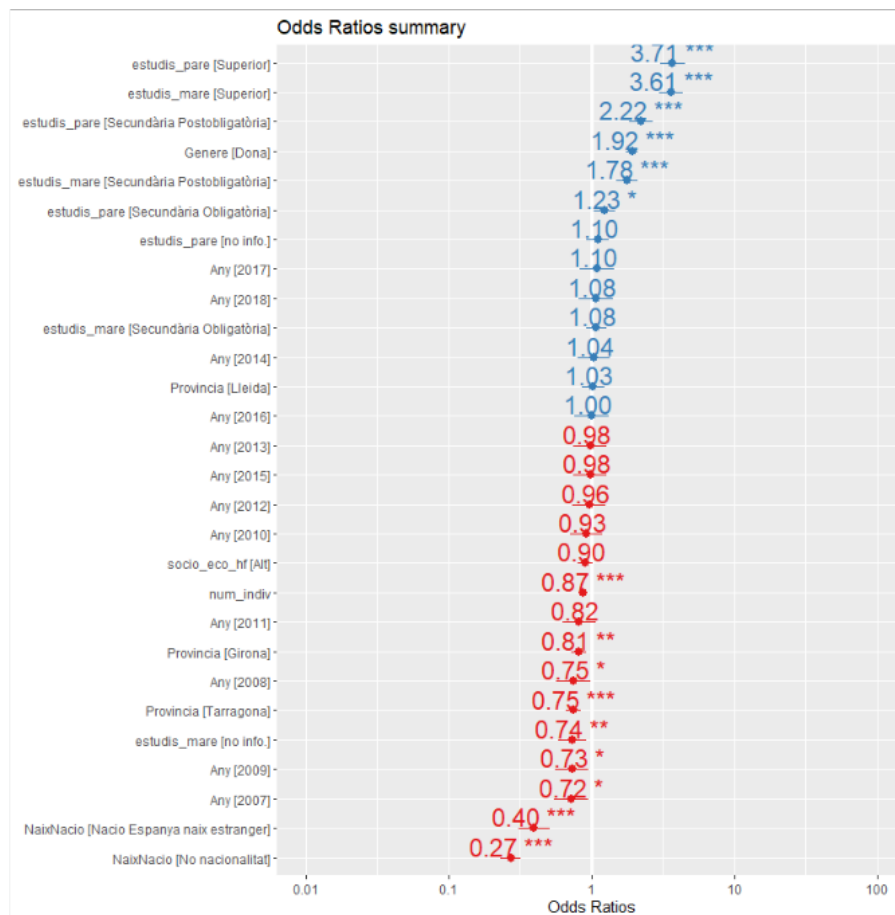
Coefficients:	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.072731	0.164541	-0.442	0.65849
GenereDona	0.653450	0.051657	12.650	< 2e-16 ***
NaixNacioEspanya naix estranger	-0.923572	0.129460	-7.134	1.08e-12 ***
NaixNacioNo nacionalitat	-1.302907	0.087233	-14.936	< 2e-16 ***
estudis_pareSecundària Obligatòria	0.205162	0.088014	2.331	0.01978 *
estudis_pareSecundària Postobligatòria	0.797571	0.097608	8.171	3.65e-16 ***
estudis_pareSuperior	1.310441	0.101238	12.944	< 2e-16 ***
estudis_pareno info.	0.099571	0.093564	1.064	0.28728
estudis_mareSecundària Obligatòria	0.076233	0.085235	0.894	0.37115
estudis_mareSecundària Postobligatòria	0.575729	0.088584	6.499	8.68e-11 ***
estudis_mareSuperior	1.283797	0.098877	12.984	< 2e-16 ***
estudis_mareno info.	-0.304445	0.114699	-2.654	0.00797 **
num_indiv	-0.138924	0.026329	-5.277	1.36e-07 ***
socio_eco_hfAlt	-0.102866	0.063862	-1.611	0.10729
ProvinciaTarragona	-0.288941	0.060719	-4.759	1.99e-06 ***
ProvinciaGirona	-0.204743	0.062971	-3.251	0.00115 **
ProvinciaLleida	0.027186	0.091139	0.298	0.76549
Any2018	0.076433	0.139706	0.547	0.58433
Any2017	0.091505	0.139772	0.655	0.51270
Any2016	0.003222	0.137940	0.023	0.98137
Any2015	-0.021537	0.137986	-0.156	0.87597
Any2014	0.037957	0.132164	0.287	0.77397
Any2013	-0.021365	0.136533	-0.156	0.87566
Any2012	-0.035849	0.138080	-0.260	0.79516
Any2011	-0.199211	0.136643	-1.458	0.14492
Any2010	-0.077809	0.133755	-0.582	0.56077
Any2009	-0.316023	0.136286	-2.319	0.02044 *
Any2008	-0.284976	0.140626	-2.026	0.04276 *
Any2007	-0.325747	0.141208	-2.307	0.02109 *

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Category of reference: Male, with Spanish nationality, who lives in a household where both parents have primary education, all the members of the household are unemployed, the period is 2019, the socio-economic level of the head of the family is not high and the province of residence is Barcelona.

Source: Own elaboration from the LFS data, 2007-2019 period.

Odds ratio summary of the final model including the dummy variables. The confidence intervals of the OR are calculated assuming normality. OR are ordered decreasingly.



Results of the Test F for the variables of the Final model:

```
Anova table: (Rao-Scott LRT)
svyglm(formula = SUP ~ Genere, design = my_design3, family = quasibinomial())
  stats      DEff      df  ddf      p
Genere    164.837    1.31705    1.00000    6407 < 2.2e-16 ***
NaixNacio 1008.277    1.67302    2.00000    6405 < 2.2e-16 ***
estudis_pare 961.492    1.40844    4.00000    6401 < 2.2e-16 ***
estudis_mare 472.985    1.41602    4.00000    6397 < 2.2e-16 ***
num_indiv    39.748    1.40729    1.00000    6396 1.213e-07 ***
Provincia    21.573    0.76365    3.00000    6393 1.060e-05 ***
Any           41.684    1.41571   12.00000    6381 0.003617 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```