The Economic Value of Education: Three Statistical Approaches

by

Laura Requena

A thesis submitted to the University of Sheffield in partial fulfilment of the requirements for the Degree of Doctor of Philosophy in the Department of Economics

Juanuary, 2018

Abstract

This thesis deals with the topic of the economic value of education from three different substantive and statistical perspectives.

In Chapter 2 the effects on educational subject choice of the increase in tuition fees after the 2012 higher education reform in the UK are analysed at individual level. A multinomial logit model with a difference-in-differences approach is estimated with data from HESA, choosing Scotland, where the 2012 reform had no effect, as the control group. Several models are presented taking into account students' socioeconomic background and gender. Main results show that after the reform students are less likely to opt for Arts & Humanities and more likely to choose Health & Life Sciences.

Chapter 3 estimates the returns to education in Spain using an instrumental variables approach. With data from the 2011 Living Conditions Survey, a sample selectivity model is estimated in order to avoid possible selection biases derived from only considering wage-earners. Family background variables are used as instruments for schooling. Beside showing that, in agreement with pertinent literature, ordinary least squares coefficients are downward biased, the analysis confirms that more educated individuals have the largest returns, although in Spain educational returns are higher for women than for men.

In the context of increasing global migrations, recent research has shown relevant differences in education and skills between natives and immigrants in developed societies. With data from PIAAC, chapter 4 presents extended Mincerian equations to gauge differences regarding returns to skills and levels of education between natives and immigrants arrived to Denmark, France, Spain, and United Kingdom from OECD and non-OECD countries. The ordinary least squares analysis is complemented with a Bayesian estimation. Results suggest that there are significant differences between native and immigrant returns to education in some of the European countries considered.

Acknowledgements

I would like to thank my supervisors, Arne Risa and Steven McIntosh for their invaluable comments, time, effort, patience and their great attitude at all times during my PhD. I would like to thank the Department of Economics for funding my PhD and giving me the opportunity to gain teaching experience in a prestigious University.

Secondly, I would like to thank my fellows Emily McDool, James Oswald, and Dan Li who have been by my side since the start of the Masters Degree and to Damon Morris, Harry Pickard, Alessandro Sasso and the rest of the PhD students who helped make this experience unforgettable. I am also grateful to my friends from outside academia and the Mexican community in Sheffield, who have always been like a second family throughout the writing of this thesis.

Last but not least I must express my very profound gratitude to my family, in particular to my beloved husband Pepe, for his support, love and encouragement every single day during the PhD; to my admirable mother Maite, for always being positive, and to my father Miguel, for sharing his wisdom, for being there whenever needed, for his unceasing support and without whom this accomplishment would not have been possible.

Contents

List of Tables

Chapter 1 Introduction

1.1 Purpose and Relevance

A society's economic wellbeing depends on several factors such as the amount of goods and services that a country can produce. A relationship between inputs/outputs can be established through the production function. Relevant inputs are labour, human capital, physical capital and natural resources. Out of the latter, human capital could be considered a key input to achieve fast economic progress. A country could have very rich natural resources, but if the level of human capital stock is low, the economic progress could not be fast achieved. Raymond (2011) gives an example of this that occurred after the Second World War: several countries were destroyed but Germany and Japan, countries with a high human capital stock, recovered faster than others in spite of the destruction of their physical capital during the war.

Technological change is vital for the economic progress of a country. In order for technological change to occur in an economy, innovation and capability to imitate behaviours of successful economies are required. This cannot be achieved without accumulation of human capital. One of the ways to measure human capital stock is through individuals' education level that directly affects income generation.

Although the Economics of Education could be considered a very recent field of research, it is internationally recognised and its relevance is increasing over the years as Salas (2002) points out. This field rose to prominence when educational institutions faced economic difficulties, when education was demanded to cope with the technological change or when there was presence of high graduate unemployment rate. Economic analysis applied in the field offers solutions for these issues or some form of orientation on how to mitigate them.

The Economics of Education is a discipline that exists since the 60s when the economists Gary Becker and Theodore Schultz started to consider education as an investment. Education acts as a channel to increase productivity and consequently a higher salary could be obtained. Questions such as whether education was a better investment than others or what type of education will generate higher returns arise. In addition to the private returns to education, a new concept was introduced, the social returns to education. If a population is more educated, the aggregated income will increase, generating social rates of returns.

The Economics of Education literature has experienced changes since its origins. As Pineda Herrero (2000) argues, in the 60s, an apogee of the discipline occurred when the Human Capital Theory (Becker (1964)) emerged. During this period, education was considered as a springboard for social mobility and, because of this a higher spending on education and an expansion of the education system took place. Researchers were interested in measuring the contribution that education made to societies' and individuals' development and started to use economic models to efficiently explain this contribution. There was an increase in the demand for education until the 70s when, compared with the initial boom, it started to decline. At this stage the positive economic effect of education and its power to mitigate social inequalities was called into question.

Because of this, governments reduced spending on education and its expansion slowed down. The main focus during this period was the improvement of the education system and the effective use of scarce resources. Work in this field was more centered during this time on topics such as the demand for education, equity or the relationships between education and certain aspects of the labour market. Governments were interested in the role that education plays in finding a job. Along with these changes, a radical stream of economists that were not in favour of the Human Capital theory emerged such as Spence (1973) or Thurow (1975) and alternative education theories were developed. The 80s were a period of socioeconomic changes; with Neoliberalism, a reduction of expenditure in education took place again. The research on this topic decreased and there was a loss of interest in the role that education plays in the economy. In the 90s the research in this field focused on the evaluation of educational institutions to make them work efficiently and to guarantee successful results in order to determine an appropriate portion of public budget to finance education. Although research in this field was initially limited to the United States and the United Kingdom, it was rapidly extended to western Europe.

Education is a key element for development both at an individual level and at a country level. Education is the enginee to spread culture, values and knowledge across the population. Thanks to this, better life conditions can be achieved and a reduction of inequalities between rich and poor individuals can happen. Apart from the mentioned effects of education, there are papers that show wider benefits of education. Being more educated increases the probability of good health (Feinstein (2002)), produces higher income and this increases happiness (Blanchflower and Oswald (2004)), reduces population growth (Becker et al. (1990)), crime (Freeman (1996)) and poverty (Biosca et al. (2011)).

As Grao and Ipiña (1996) point out, the Economics of Education literature analyses the education's economic value to promote the economic progress and analyses the economic aspects of education such as costs, returns and efficiency. Some of the main research areas within the Economics of education literature are School Choice; Policy Evaluation; Skills, Vocational Education and Training; Returns to Education and Economic Impact of Education; Education and the Labour Market; Determinants of Academic Performance and School Failure; Education Market & Competition; Gender & Education; Finance, Management & Quality; Equity, Inequality & Demand for Schooling; Human Capital, Growth and Economic Development; Intergenerational Mobility; Education Mismatch among others.

Apart from the first work in the field mentioned before, there are highly influential papers such as Mincer (1974) that introduce the earnings equation relating the logarithm of earnings to schooling and experience. Another paper that can be considered a cornerstone in this field is Card (1999) with his revision of the work on the causal effect of education on earnings.

According to Machin (2014) an upsurge of research in the Economics of Education literature occurred in the last decade. Machin (2014) shows that from the 50s to the 00s, excluding the 80s the number of education publications in mainstream economic journals has increased, therefore there is a growing interest in the topic. Factors that contribute to this upsurge are expansion of the education system, the increased relevance that education has on economic and labour outcomes, influences that other disciplines have on the field, use of empirical methods taken from other disciplines, the need to evaluate recent education policy and the availability of more and richer data sets. The quality of the data has increased thanks to governments allowing access to administrative data. In addition, there are better survey data available, plus the linkage between different surveys and a large amount of available variables to establish relationships between education and other aspects not available before. An increase in the quality of the data occurred in the past years.

Another aspect that makes this field one of the most important stream in the economics literature is that governments, in order to improve life conditions and make progress, implement education policy and this literature serves as an evaluation mechanism. This is why the interest in this field has increased in the last decade and it is believed that will continue to grow in the next years.

1.2 Objectives and Structure of the Thesis

Chapter 1 introduces this thesis. The three topics addressed in the three main chapters contained in this thesis reflect different aspects of the Economics of Education literature. Chapter 2 focuses on the impact of the 2012 higher education reform regarding the increase in the tuition fees cap in the United Kingdom. This chapter belongs to the higher education policy evaluation literature. Chapter 3 covers another aspect of the economics of education literature, the returns to levels and years of education in Spain. Chapter 4 estimates returns to levels of education and numeracy skills with an emphasis on comparing natives to immigrants in four European countries. Chapter 3 and 4 belong to the returns to schooling literature but each of them focuses on different aspects within the returns literature. Chapter 5 concludes and offers a summary of the whole thesis and future research avenues.

The way that Chapter 2 fills a gap in the literature is by estimating the effect of the 2012 Tuition Fees on degree choices for the UK using data from two cohorts of students in Higher Education. In addition, a novel multinomial difference-in-differences econometric technique is applied to analyse the data and to evaluate the reform. Data come from the Higher Education Statistics Agency. Chapter 3 estimates the returns to schooling addressing the potential endogeneity bias in the coefficients by using different instrumental variables for each education level and justifying the choice of instrument for each level. This work is the first one using the particular choice of instruments for different education levels. Data from the 2011 Life and Conditions survey available from the Spanish National Statistics Institute were used. Chapter 4 estimates a linear model under two econometric approaches, a Frequentist and a Bayesian, and presents the returns to levels of education and numeracy skills for natives and immigrants from developed and developing countries and investigates if there are significant differences. To the best of my knowledge, there is no research work that applies the Bayesian methodology (Bayesian posterior estimation plus Bayesian Model Averaging) to compare immigrants and natives returns. Data from the Program for the International Assessment of Adult Competencies available from the OECD were used.

Overall, the thesis therefore contributes to the Economics of Education literature discussed above in two important ways. First it contributes new evidence on particular issues relating to participation in, and the value of, education. But second, the thesis also makes important methodological contributions related to the estimating of such effects, in terms of the use of instruments and of a Bayesian approach.

1.3 Research Questions

Chapter 2 investigates if students in higher education change their degree choices after the 2012 tuition fees cap increase. The main question addressed is whether students in higher education affected by the reform choose subjects that will provide better employment prospects such as rapid insertion in the labour market or a higher salary. If this effect after the tuition fees cap increase exists, is it equal for different students subgroups?

Chapter 3 addresses whether public expenditure in education and mother's education are valid instruments for attainment of low education levels. Moreover another investigated research question is whether father's occupation is a valid instrument for tertiary education attainment. These two questions are addressed for the Spanish context.

Chapter 4 focuses on the estimation of the returns to levels of education and numeracy skills in four European countries (Denmark, France, Spain and the United Kingdom) that have a variety of immigrant integration approaches. The difference between immigrants' and natives' returns are investigated under two approaches: a Frequentist and a Bayesian.

Chapter 2 Tuition Fees rises in the UK: Effect on degree choices

2.1 Introduction

We are all part of an "information age". Education and training have become more attractive in the last decade. As a result of this and in conjunction with the technological change, an expansion of Higher Education is taking place. As Barr (2004) points out, the government uses two key elements to control this expansion: price control and quantity control. On the one hand, universities in the United Kingdom can freely determine the amount of tuition fees to be charged for international students and for all postgraduates courses. On the other hand, there is a regulation that states the minimum and maximum amount of fees the universities can charge to European and home students. Regarding the quantity control, universities stipulate in accordance with the Higher Education Funding Council for England (HEFCE) to restrict the number of students taught HEFCE (2012). Universities could be penalized if they do not exactly follow the agreed recruitment quota. As a result of the growing higher education demand, the number of universities and number of students have increased. A consequence of this is a greater variety of subjects. Each university has a different funding scheme.

The way tertiary education in England was funded since 1998 implied that, to cover the costs of teaching, students pay tuition fees and the government provides universities with teaching grants. Students are entitled to fee and maintenance loans to be able to pay the main expenses when they attend university (tuition fees plus living expenses). Students whose families are on a low-income scheme, have the right to apply for extra economic help from the government or fee waivers from the universities. Following the Browne Review[1] (an independent review of tuition fee policy in the UK reported in October 2010) in 2012, the government introduced changes that affected how the system was functioning. The teaching grants that the government provide to universities were greatly reduced. In order not to pass that cost to the universities, the government allowed Universities to charge higher tuition fees. In addition to this, modifications took place in the the terms of the loans.

The last higher education reforms have been criticised for being the drivers of a shifting in the system: students now behave as real consumers purchasing a qualification and not engaging in the acquisition of knowledge and new skills (Maringe (2006)).

With respect to policy implications, the research question addressed in this chapter is linked to the skill composition of the labour force. If the 2012 higher education reform discouraged people from studying a particular subject, this could lead to a potential shortage of labour supply in some areas and an increasing concentration of graduate jobs taken by individuals from the same subjects. In this chapter a multinomial logit model is estimated for the subject choice among students domiciled in England and Scotland and average treatment effects under a difference-in-differences approach are used to study

[1] The Browne Review formal title is "Securing a Sustainable Future for Higher Education in England". This report is available in `https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/422565/bis-10-1208-securing-sustainable-higher-education-browne-report.pdf`

the impact of the higher education reform on subject choice. A robustness check is carried using weights in the regressions.

This chapter is divided as follows: section 2 gives an overview of the policy background regarding the higher education system in England and Scotland; section 3 reviews the relevant literature regarding subject choice; section 4 explains the data set used for the analysis and the variables included in the model; section 5 describes the methodology employed in the estimation; section 6 reports the results obtained; and section 7 concludes the chapter.

2.2 Policy Background and Education Systems

There is an ongoing debate concerning the 2012 reform in Higher Education and, in particular, the increase in tuition fees. There are three key variables related to Higher Education in the United Kingdom: tuition fees; means of economic support, maintenance grants or maintenance loans; and the repayment scheme for this money. The first form of tuition fees was introduced in all the United Kingdom in 1998. Since then, in England, the university tuition fees have not been constant over time. The tuition fees cap increased to £9,000 in September 2012. More than half of the universities made public their plan to charge the whole amount while the intention of the rest was to charge the minimum amount of £6,000. A brief list of the major changes that this reform entails is given: rise of deferred tuition fees cap from £3,375 in 2011 to £9,000 per year and the larger amount of the cost of the degree has been passed on to the students rather than the government; a change in the income threshold where students should start to pay back their loans from £15,795 to £21,000; a change in the number of years after which the loans are written off from 25 to 30 years; an increase in the maintenance grant from £2,984 to £3,250 per year if parental income is less or equal to £25,000; and an introduction of the National Scholarship Programme to help students from low-income backgrounds to enter Higher Education. According to HEFCE, students starting their full-time course on or after 1st of September 2012 could be charged an upper tier of £9,000 and a minimum amount of £6,000. In comparison with privately-funded or alternative providers, only publicly-funded institutions are affected by the tuition fees caps. Figure 2.1 from Crawford and Jin (2014) contains the changes explained in more detail of how the reform affected students commencing the academic year of 2012.

Policymakers and researchers are concerned about the effect of these Higher Education reforms on students in England. This worry could be explained considering that, if students pay higher fees, it has to be transparent for them how they will benefit after that higher payment. In order to enable students to be informed, universities must work in an efficient way and be transparent explaining financial worth in exchange of Tuition Fees.

HEFCE (2013) is report written in March about the impact of the 2012 reforms and contains information regarding subjects of study. An advisory group was created in 2005 to deal with the changes in demand and supply of subjects: Strategically Important and Vulnerable Subjects (SIVS). Subject previously designated as SIVS are Maths, Physics, Chemistry, Engineering, Modern Foreign Languages and Quantitative Social Science. These advisors have several aims to ensure that the amount of graduates meet the demand of the employers and society and to monitor which subject is at risk at any point in time.

Figure 2.1: HE funding system in England for students first enrolled in 2011/12 and 2012/13

	Students first enrolled in 2011–12	**Students first enrolled in 2012–13**
Fees	£3,375 in 2011; £3,465 in 2012, as announced; assumed to stay at £3,465 thereafter.	Up to £9,000 a year.
Student support		
Fee loan	All students may get a loan from the Student Loans Company (SLC) to pay the fees and must repay SLC after they graduate.	All students may get a loan from the Student Loans Company (SLC) to pay the fees and must repay SLC after they graduate.
Maintenance grant	In 2011, £2,906 if household income less than or equal to £25,000 p.a. Tapered away at 20% withdrawal rate for income between £25,000 and £34,000. Tapered away at around 7% withdrawal rate between £34,000 and £50,020. The maximum grant, the means-testing thresholds and the taper rates changed slightly for subsequent years.	In 2012, £3,250 if household income less than or equal to £25,000 p.a. Tapered away at around 18% withdrawal rate thereafter. No grant available when parental income exceeds £42,600. The maximum grant increases slightly in subsequent years.
Maintenance loan	The maximum loan is £3,838 for students living at home, £4,950 for others outside London, and £6,928 for those away from home and in London. The maximum loan is lower for the final year of study. Students lose 50p maintenance loan for every £1 they receive as maintenance grant. The loan is tapered away at 20% for household income above £50,778. All students are guaranteed at least 72% of the maximum loan. The parameters did not change in cash terms between 2011 and 2013.	The maximum loan is £4,375 for students living at home, £5,500 for others outside London, and £7,675 for those away from home and in London. The maximum loan is lower for the final year of study. Students lose 50p maintenance loan for every £1 they receive as maintenance grant. The loan is tapered away at 10% for household income above £42,875. All students are guaranteed at least 65% of the maximum loan. The parameters did not change in cash terms between 2012 and 2013.
Other student support	Universities have their own schemes. They are obliged to pay a minimum of 10% of fees to students who receive the maximum maintenance grant.	The National Scholarship Programme (NSP) was introduced to give at least £3,000 each to low-income students. Eligibility requires household income to be no more than £25,000. The award may be given as fee waivers. Universities determine the detailed criteria. Not all eligible students are guaranteed an award. The NSP has since been abolished for undergraduates.
Accumulation and repayment of student loans		
Real interest rate (relative to RPI)		
During study	0%	3%
After graduation	0%	0–3% depending on graduate income: 0% if below the repayment threshold, linearly increasing to 3% for income at or above the higher repayment threshold
Repayment rate	9%	9%
Repayment threshold	£15,795 in 2012 (above which 9% of income is to be paid)	£21,000 in 2016 (above which 9% of income is to be paid)
Higher repayment threshold	n.a.	£41,000 in 2016 (at which point the real interest rate is 3%)
Threshold indexation	Annually in line with RPI from 2012	Annually in line with national average earnings from 2017
Repayment period	25 years	30 years

This reform has not been uniform in all the countries of the United Kingdom. Nowadays, Scotland provides free higher education for "ordinary residents" domiciled students in Scotland for at least three years prior to the starting of the first academic year.

Scotland behaves differently from England in terms of the cost of Higher Education. Tuition fees were abolished with the Scottish Graduate endowment scheme in 2001. In this situation, the money collected from the endowment that graduates paid after their studies was devoted to provide poorer students with bursaries but not in the form of tuition fees. The graduate endowment was annulled for those students who graduated on 1st April 2007 or after.

A brief summary of the tuition fees costs depending on country of domicile before starting their course and depending where they choose to study is given. As mentioned before, students domiciled and residing in Scotland that pursue their degree do not pay any university tuition fees. In order to afford their living costs, from April once they graduate and if their annual wage is over £15,000, they begin paying back their maintenance loan. Economically-disadvantaged students are eligible for non-repayable bursaries. Students from Scotland that choose an English provider are subject to pay a variable tuition fee up to £3,375 if they started in 2011. These students can either apply for loans conditional on the income of the household or apply for the Students Outside Scotland bursary. The latter has changed over the last academic years. In 2007/08 this bursary consisted of the highest amount of £2,045 if the yearly family income was below £18,360 while if the income was above £32,515 the students were not eligible. In the academic year 2013/14 there were two available bursaries: the Young Students' Bursary, for students who are under 25 years old and a maximum bursary of £1,750 and the Independent Students' Bursary, for students aged 25 or older and a maximum bursary of

£750. All of these bursaries are for students studying elsewhere in the UK or the Republic of Ireland. Students domiciled in England but studying in Scotland followed the same maintenance grant scheme and loan scheme to pay the tuition fees as those studying in England. Students from England studying in Scotland, would pay a maximum of £3,375 in tuition fees if they started in 2011 and a maximum amount of £9,000 if they started in September 2012. Again, this group of students could have used the repayable loan to cover the fees. Repayable loan conditions differ across countries and across academic years.

Turning now to the education systems in England and Scotland, details for both countries will be given and some differences will be highlighted. As Machin et al. (2013) argue, we need to accept that although England and Scotland are two different countries, they are both part of the UK. Tax and benefit policies regarding the whole expenditure in education are determined at a national level rather than a country level. These authors summarise the main differences in both education systems: Scotland's curriculum is non-statutory[2]; the most common qualifications needed to access university are different with Standard Grades and Highers in Scotland to GCSE and A-Levels in England; in Scotland, local authorities are more influential in school management.

In considering the two education systems in these two countries of the UK, differences are next reviewed. Assessment differs in England and Scotland. In the former, pupils take key stage exams until General Certificate for Secondary Education (GCSE) in the start of the 11th year of school and at the end of year 12 when students are 16. Once students reach this point, they have two options, either leave school or continue education

[2] Not dictated by the Government.

for a couple of years. After this, then they take a common minimum of 3 A-Level exams to access university[3] at age 18. On the other hand, pupils in Scotland take the Scottish Standard Grades at age 15 and 16. Therefore, they have 11 years of compulsory education compared to 12 in England. In Scotland, they could either sit Scottish Highers at age 17 and Highers and Advanced Highers at age 18. To enter higher education, students aged 17 could then take five Highers subjects. These differences stem from the length of university courses: 4 years in Scotland while 3 in England. Thus in Scotland they often have one extra year of University instead of one year more of school.

In summary, pupils are allowed to leave school during S4, the final year of Scottish Standard Grades while in England, year 11 is compulsory for them before leaving after finishing two years of GCSEs. Scottish pupils are allowed to quit school at age 17 after one year of post-compulsory education.

2.3 Literature Review and Research Hypothesis

There are two strands of literature relevant for this chapter: first, the tuition fees literature focusing on the impact that the tuition fees increase has on several areas and second, literature about educational choice in Higher education, in particular the subject choice.

Regarding to the literature about the impact of higher tuition fees, Crawford and Jin (2014)

[3] Note on Admission: The most common route to enter higher education in England is obtaining the General Certificate of Education at Advanced (A) Level. Less common qualifications to access universities are listed: qualifications such as NQL Level 3 and HE Diploma, level 3 qualifications in Credit and Qualifications Framework for Wales, Welsh Baccalaureate or Scottish Highers, Advanced Highers or equivalent qualifications.

published an article focus on several points related to the 2012 higher education reform. They pay particular attention to the consequences of the debt that students have after finishing their studies. The old debt system was compared to the current system and the authors found that higher earners are more affected. These authors predicted that three-quarters of graduates will not repay the full loan due to not having enough earnings to pay back and they argue that if graduates become high income earners they pay more and if graduates become low income earners they pay less under the new system. They pointed out that on average, under the new system students finish their studies with a debt of £44,035 compared to £24,754 under the old system. Another source of information about the impact of the 2012 higher education reform is given by the HEFCE (2013) annual report. The latter focused on the impact of the reform in two aspects: students and higher education institutions. A brief list of relevant findings follows: in the academic year of the reform, there were 47,000 fewer full-time students compared with the year before. The number of part-time students at undergraduate and postgraduate level has decreased since 2010/11. There is a continuing growing trend in international students coming to the UK to study. There was a downward trend in participation of mature students (aged 20 and older).

Another interesting finding was the existence of participation gaps between subgroups of students. In the most deprived areas of England, girls aged 18 are 50% more likely to apply than boys; students from a high socioeconomic background are three times more likely to apply for tertiary education than students from a low socioeconomic background. Related to participation in higher education, Dearden et al. (2011) used the Labour Force Survey to create a pseudo-panel with cohorts varying from 1992-2007 to conclude that a £1,000 increase in tuition fees led a participation reduction of 3.9 percentage points.

On the other hand, a £1,000 increase in maintenance grants increased participation by 2.6 percentage points.

Turning to the literature about subject choice, in a study conducted by Lindley and McIntosh (2015) wage inequality is linked to subject of degree. Wage differentials were calculated by subject from 1994 to 2011. The study gives a potential explanation to the increase of wage inequality over the last decades: less occupational concentration of subjects and widening of cognitive skills of graduates. Lindley and McIntosh (2015) found that a larger variance of wages could be explained by a broad range of jobs being performed. This paper decompose the variance of the graduate log wage into the variance within subjects plus the variance between subjects.They claim that, as a result of the higher education expansion, more people became part of the higher education section and study a degree and as a consequence of this, more variability and dispersion of graduates in all subjects and a wider range of cognitive skills arise. Regarding the occupation analysis, all subjects became less concentrated.

Some authors focus on the returns to a particular degree such as Chevalier (2011) and Arcidiacono (2004). Chevalier (2011) uses a cohort of recent graduates from different degree subjects to perform his analysis. After reporting specific subject wage premiums, he defends that the introduction of subject specific tuition fees would be reasonable. Large heterogeneity was found in the mean wages of graduates once they join the labour market. In addition, large gender differences regarding the subject specific wage were present. The evidence seems to indicate that graduate women have a wage premium if they study a particular subject in comparison with graduate men. In general, women are better at jobs that require soft skills while men succeed more in hard skill jobs. This is why women tend to study degrees that involve communication, dealing with people, human resources, teaching, etc. and men opt for jobs that require more technical skills.

Although it could be thought that at the time of choosing their subject of study, future students are uncertain about their potential income, and that this fact would not influence their subject choice, Chevalier (2011), in contrast, maintains that a difference in the amount of tuition fees to be paid and potential future income do have an impact on student choices. An author that agrees with Chevalier (2011) in this idea is Sa (2014). One of her findings was that applications to courses with uncertain employment perspective are more likely to be affected by tuition fees changes. Her article compared two higher education reforms: the 2001 tuition fees reform in Scotland and the 2012 reform in England. The effect of fees on the demand for higher education, university attendance and course choice was examined between these two countries. Regarding the subject choice, Sa (2014) ordered the subjects by employment prospects quartiles and by expected salaries quartiles. The difference in differences econometric technique was employed to analyse the effect of the higher education reform in the separate countries. In relation to Chevalier's (2011) ideas about the subject specific tuition fees, Sa (2014) suggested that students would be well-disposed to pay higher fees to attend particular universities or to study particular subjects that offer better employment opportunities.

Other authors arguing that different fields of study have different labour market payoffs are Kirkeboen et al. (2016) for Norway, Britton et al. (2016) and Walker and Zhu (2013) for the UK. Kirkeboen et al. (2016) compare payoffs attending a selective institution to payoffs by field of study. After using an instrumental variables approach, findings show that the latter payoffs are larger than the former; students choose subjects in which they have a comparative advantage. Related to this topic, Britton et al. (2016) argue that there is considerable variation in earnings observed across different subjects and institutions.

These authors find that subjects such as Medicine or Maths yield higher premiums than other subjects, while Arts delivers earnings typical of non-graduates. They also argue that the decision-making process regarding the subject occurs before the student enter Higher Education.

Walker and Zhu (2013) present evidence of the impact of higher education on lifetime net earnings and analyse the decision-making of students taking into account prospective futures. Although these authors acknowledged that the robustness of the estimates broken down by subject cannot be performed, they find that recent changes in the loan system benefit students from a low socioeconomic background and balance the effect of the increase in the tuition fees.

Valbuena (2012) suggests a model for subject choice depending on several aspects: for instance, personal characteristics, family background, attitudes and behaviors, past educational attainment or university costs. Large differences are found between advantaged and disadvantaged background students in terms of Higher Education decisions (such as choosing to study science or attending a prestigious institution). The paper indicates that much of the socio-economic gap in Higher Education participation rates is caused by reduced participation rates of students at the lower end of the income distribution.

Similar to Valbuena (2012), Chowdry et al. (2013) find that students from lower socioeconomic background are less likely to pursue a Higher Education degree than students from a high socioeconomic background.

Regarding linking college major and tuition fees, Walker and Zhu (2011) study the effect of an increase in tuition fees in the internal rate of returns in terms of the quality of the investment in higher education. The analysis gives information across different subjects. Evidence from their paper indicates that the large rise in tuition fees will not generate a significant substitution across subjects. Estimates of college premium are provided as well as graphs that show earning profiles by degree major for men and women. The authors point out that unobserved differences between graduates by major could emerge and this could limit the econometric analysis due to the available variables in each dataset. This paper stands out for calculating age-earnings profiles across subjects and gender in addition to the allowance of alternative tuition fees in their regressions

Several studies have explored the relationship between subject choice and other factors. For example, Purcell et al. (2008) draw attention to the influence of socio-economic and educational factors on the subject of degree decision. This work gives context and background of several variables used as controls in the regression of this chapter. These authors have drawn attention to the existence of regional differences in likelihood of acceptance to a particular place in the United Kingdom. They investigate as well how applicants chose their courses and subjects and they find that personal characteristics and previous experiences affect course choices. A survey that they carry out about subject decision-making reveals that interest in the course and employment prospects were the most common reasons to study an specific subject. They find that age, socioeconomic background, ethnicity and gender have an effect on subject choice. Another study that agrees with this is Leppel et al. (2001).

There are authors such as Berger (1988) and Beffy et al. (2012) that try to establish a relationship between subject choice and future earnings. The first author shows that expected initial earnings have less impact on graduates in comparison with their expected flow of future earnings. Beffy et al. (2012) using data from France obtain the effect of expected earnings on the likelihood of choosing a particular subject. They correctly argue that gender and parental profession are related to the choice of major. In their study science degrees are more likely to be chosen by male students while humanities and social science are more female predominant. Regarding the parental background, individuals who had a father who is a blue-collar worker have a lower probability of studying science and a higher probability of studying humanities and social science degrees. If the mother has a white-collar occupation, students are less likely to choose majors in humanities and social science in comparison with those students who have an executive mother. However, if the mother holds an intermediate profession, a farmer occupation or is a tradeswoman, they less probably choose degrees in social science in comparison to those from science. Going back to the main purpose of the paper, the authors discovered that the elasticity of major choice to expected earnings, although statistically significant, is very low at the same time and therefore it can be concluded that the choice of subject is essentially influenced by non monetary factors. The effect of expected earnings on major choice is higher in humanities and social science than in science. The authors maintain that the major choice is mainly driven by schooling preferences and ability.

The way this work fills the gap in the literature is by combining the following strands: the effect of the tuition fees increase and the subject choice. In order to do this, a model for the choice of university subject will be proposed and the effect of the tuition fees increase on the subject choice will be investigated. Given the reviewed literature and its main concerns, the aim of this chapter is to address the following research question: did the 2012 tuition fees reform affect subject choice? How are student choices influenced by the cost of their studies? How is this effect distributed across subgroups of students?

2.4 Data Set and Variables

Data used in this analysis was provided by HESA[4], which is the agency in charge of collecting yearly student data from publicly funded Universities in the United Kingdom[5]. Pooled cross-sections of two academic years, 2010/11 and 2013/14, were used in this chapter. 2011/12 was the academic year just before the tuition fees reform took place while 2013/14 was the year after the tuition fees increase took place. Students from the 2011/12 academic year were not used because the participation rates in the year before the reform would be 'artificially high' and the participation rates in the year of the reform would be 'artificially low'. The reason why the participation rates will not be the same as in a random year is because those who intended to take a gap year before starting university did not take it in order to avoid the higher tuition fees.

Although international students are a key element for Higher Education in the United Kingdom, they are excluded from this analysis because they face different tuition fees to home students. European students are excluded as well from the analysis because HESA does not collect data on them for the variables used. Because of the econometric model used, only individuals domiciled in England or Scotland are taken into account.

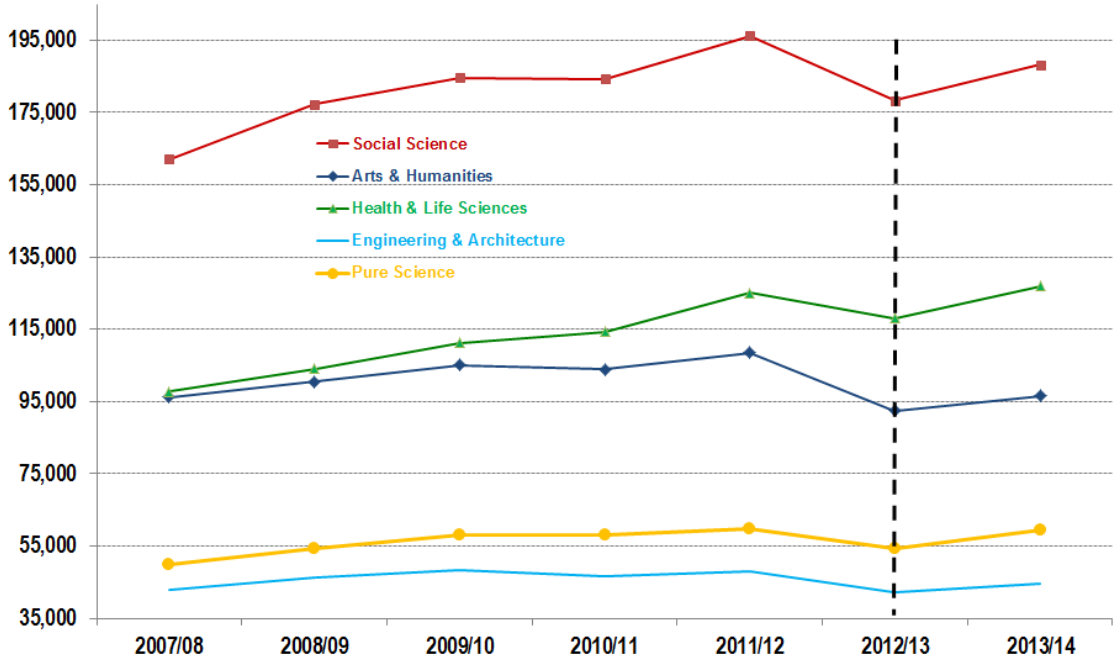
As mentioned above, two cross-sections are used for this study: 2010/11 and 2013/14. The first cross-section corresponds to the 'Pre-Treatment' period that comprises students

[4] The Higher Education Statistical Agency (HESA) webpage includes all the necessary documentation on data used, namely: specification of the data to be returned, XML Schema Definition (XSD) files, Supporting documents, and Additional guidance for users (`https://www.hesa.ac.uk/collection/c10051`).

[5] A huge majority of students were enrolled in universities publicly funded. According to HESA data, in 2015 there were in England 132 Higher Education Institutions publicly funded, 241 Further Education Colleges, and 115 Alternative Providers of higher education. Out of a total of 1,597 million students enrolled in these institutions, 89% were studying in HEI publicly funded.

commencing their studies before the tuition fees cap increase. The 'Post-Treatment' period corresponds to the academic year after the reform occurred. A control group is needed in order to see how the tuition fees cap increase affects the subject choice.

Figure 2.2: First Year, First Degree students by Subject of Study (HESA)



According to the Higher Education Statistics Agency, and looking at the seven years trend graph (Figure 2.2), we can see that there is a drop in the participation in all subjects before the reform and then a recovery in raw numbers after the reform. Data from HESA shows that 521,990 students started first degree courses at UK HE providers in 2013/14. This is 5% higher than in 2012/13, but still 5% lower than in 2011/12 - the last year before the £9,000 fee cap was introduced. If the academic year 2010/11 is compared to 2013/14, groups of subjects such as Social Sciences, Health and Life Sciences and Pure Sciences experienced an increase in participation rates of 2, 10 and 2 percent respectively while Arts & Humanities participation rates comparing these two academic years have decreased by 7 percent. Subjects such as Medicine or Biological Science, both Health and Life Science subjects, experienced the highest percentage growth in applicants over all the period represented. It can be seen that Engineering & Architecture and Pure Science, over this 7 years period, and comparing the year before and the year after the reform, are the most stable regarding the number of entrants. This graph represents individuals of all ages pursuing their first degree in their first year of study.

Students with the following characteristics are included in the sample: full-time, first degree students, domiciled in the United Kingdom (England/Scotland), who study in an institution in England/Scotland[6], who study a pure subject (students pursuing a combined subject degree are excluded) and who are younger than 21 (to be able to use the parental occupation instead of their own as a proxy of their social status). A list of variables is given:

Outcome Variable:

- Subject Choice (Y_i): takes the value of 0 if Arts & Humanities, 1 if Social Sciences, 2 if Pure Sciences, 3 if Life Sciences and Health, 4 if Engineering & Architecture. This variable is the result of a grouping of subjects using the two digits Joint Academic Coding System (JACS).[7]

Control Variables:

- Male dummy: takes the value of 0 if the individual is a female student, 1 if male student.
- Parental Occupation: takes the value of 0 if Unemployed, 1 if Managers &

[6] If students who chose to study in other countries such as Wales or Northern Ireland are included, the overall results remain.

[7] The detailed list of all the subjects grouped in each category is given in the Appendix.

Professionals (Senior Officials, Associate Professional and Technical occupations), 2 if Intermediate Level occupations (Administrative, Secretarial and Skilled trades occupations), 3 if Elementary occupations (Sales, Customer and Personal service occupations and Process, plant and machine operatives).

Three dummies are created for the models where the base category corresponds to the Managers & Professionals group. Individuals who have retired parents[8] have been dropped from the sample because the category of the variable could be misleading when interpreting the results since the previous occupation of the parents is unknown. This variable indicates the occupation of the parent, step-parent or guardian who earns the most and it is a recode of the major groups of the Standard Occupational Classification (SOC2000).

- Ethnicity dummies: each corresponding dummy takes either the value of 1 if Asian, Black or Other Ethnicity including mixture where the reference category is White.

- Parental Highest Education dummy: takes the value of 1 if at least one of the student's parents has higher education qualifications, 0 otherwise.[9]

- Associated Tariff Points (ATP): Score that indicates two aspects. The first aspect is the type of qualification gained and the second is the grade in that particular gained qualification. For each student, this variable is a sum of all the tariff points that they have obtained in total. This is the closest proxy for ability that is available in the data. This variable was standardised to control for grade inflation observed in Table

[8] The total number of students whose parents are retired represent 0.36% of the full sample.

[9] Missing Values Note: Students provide the Universities with information. Missing data could take the form of not stated information, not known or information refusal. The values from the variables Parental Occupation, Parental Highest Education and Ethnicity could be missing because these are non compulsory fields of the survey.

2.1 (the significant positive difference in ATP between the scores before and after the reform).

- Treatment Variable (D): takes the value of 1 if the individual is domiciled in England, 0 if domiciled in Scotland. Although those from Scotland who decided to study in England have to pay the same tuition tees than those domiciled in England, they are considered not treated because they have the option of not paying for higher education if they stay in Scotland and therefore no higher education cap in tuition fees is imposed on them.

- Time Variable (T): takes the value of 0 if the student belongs to the 'Pre-Treatment' period and starts their undergraduate studies in the academic year of 2010/11. The variable takes the value of 1 if the students belongs to the 'Post-Treatment' period or to the academic year of 2013/14.

- Interaction ($T \times D$): takes the value of 1 if the student is domiciled in England and starts to pursue their first degree in the 'Post-Treatment' period.

Students by Country Domicile and Country of Study in the Post-Treatment period

		Country of Domicile		
		England	Scotland	Total
Country of Study	England	201,292	884	202,176
	Scotland	2,822	17,033	19,855
	Total	204,114	17,917	

The table above shows the number of treated and not treated students from HESA database. The treated belong to the Post-Treatment period ($T = 1$) and they are domiciled in England ($D = 1$) no matter what country of study they select to pursue their undergraduate studies. Recall that students domiciled in England who choose a Scottish higher education provider are subject to a maximum fee level for a degree course of

£9,000 as well. Therefore these 204,114 students, are the treated ones in the model presented in the Methodology section.[10] Additionally, the previous table shows evidence of limited student mobility between these two countries.

After cleaning the data and imposing all the restrictions mentioned before, Table 2.1 shows that the sample comprises 395,015 students domiciled in England and 34,701 domiciled in Scotland; slightly under 52% of those domiciled in England started their studies after the reform in the academic year 2013/14. The corresponding percentage for those domiciled in Scotland is similar.

Table 2.1 presents descriptive statistics, distinguishing between students domiciled in England and those domiciled in Scotland, and pre/post reform within these groups. On the whole, students domiciled in England are similar to students domiciled in Scotland. In particular, regarding the subject choice variable, Table 2.1 columns (1) and (2) show that, taking into account both academic years, for some subjects such as Arts and Social Sciences, the percentage of students is larger for students domiciled in England than those domiciled in Scotland. The opposite occurs with subjects such as Pure Science, Health and Engineering; these subjects are relatively more popular among students domiciled in Scotland. However, if all the subjects are taken into account, apart from those who chose to study Arts, the percentage of students between domiciled in England and domiciled in Scotland does not differ by more than 6%. Even though differences in characteristics exist between the two groups of students, it can be seen that these differences only occur in some of the control variables. For instance, it is unsurprising that the percentage of university students aged 17 and under is higher for those

[10] Results remain the same when students domiciled in Scotland who chose an English Higher Education provider are excluded from the analysis.

domiciled in Scotland than for those domiciled in England. This difference stems from the fact that in Scotland, they can progress to University after college at the age of 17 while in England the most common age to start University is 18. Regarding the older age groups, there are fewer older students (aged 20) domiciled in England than in Scotland while the opposite happens for students aged 19. With respect to the ATP (the proxy for ability), students domiciled in England have, on average, a higher total score. This means that either they have attained more qualifications than students domiciled in Scotland or those qualifications have been attained with a better grade. Notice that the ATP score increases the more qualifications the student has or the better the grade is in each of these qualifications. On the other hand, parents of students domiciled in Scotland are more likely to have attained higher education in comparison with parents of students domiciled in England. Regarding ethnicity, whites are more predominant amongst Scottish domiciled students and the contrary occurs for Asians and Black students. In general, with respect to ethnicity, there is more diversity amongst the students domiciled in England. Scotland has a lower proportion of ethnic minorities and is relatively poorer compared to England (ONS, 2002).

It is widely known that nowadays, there are more girls than boys pursuing higher education. This phenomenon is observed in all the subsamples irrespective to the period of study or the region of domicile. Although parental occupations distribution is not even within country of domicile, parental occupation proportions are similar across students from both countries of domicile with slight differences in some of the occupational groups. These summary statistics suggest that, on the whole, there are few differences in the composition of students domiciled in Scotland and England.

Migration Flows Between Countries in the UK

According to Simpson and de Sheridan (2014) the overall number of applications and acceptances to the UK higher education institutions have decreased since 2012. These authors, basing their analysis on student acceptances, argue that the 2012 Higher Education reform had a different impact on applicants depending on the country where they live before starting their course. In particular, as Higher Education is free for Scottish domiciled students if they study in Scotland, but not in England, it is natural to think that the 2012 Higher Education reform that took place in England could act as a financial incentive to pursue their degree in a Scottish Institution rather than an English one. However, further evidence from a different dataset is next presented regarding stable mobility trends in the United Kingdom over time.

Using actual data of students from the Education Information Database for Institutions[11] (HEIDI) and plotting students from all levels of study, full time, first degree and domiciled in England and Scotland, Figure 2.3 shows that the time trends on the number of students studying in different countries from they were domiciled are stable over time. More precisely, there is a slightly larger number of Scottish students who stay in their country of domicile to study their degrees in the year of the reform (90%) compared to the year before (89%). The contrary occurs for English domiciled students who stay in their country of domicile to pursue their degrees in the year of the reform (95%) compared to the year before (96%). To summarise, the bar chart in Figure 2.3 shows that the student flows between administrations do not greatly vary over time.

Mosca and Wright (2010) analysed migration flows for undergraduate students. Cohorts

[11] HEIDI: web-based management information service run by HESA. HEIDI provides a rich source of aggregated information about higher education statistics in the UK. HEIDI provides aggregated statistics for students with certain characteristics such as region of domicile.

of graduates are used between 2002/03 and 2006/07. Using data from HESA, these authors demonstrate that the majority of graduates chose to study their degrees at their home country. Specifically, out of all the students domiciled in England, 95.4 % chose to study in England and 1.4% chose to study in Scotland. Of those domiciled in Scotland, 6.8% chose to study in England while 93% chose Scotland.

To conclude, Heat maps made by HESA are shown in Figures 2.4 and 2.5 and they reveal that in the 2013/14 academic year, students tend to stay and pursue their first degree in their country of domicile. These figures represent full-time first degree students by region of higher institution provider and region of domicile. According to HESA, in the academic year of 2013/14, 95% of students domiciled in England and Scotland remained there to pursue their first degrees. This again shows evidence of little student mobility across countries.

Figure 2.3: Students Migration Flows Between Administrations Over Time

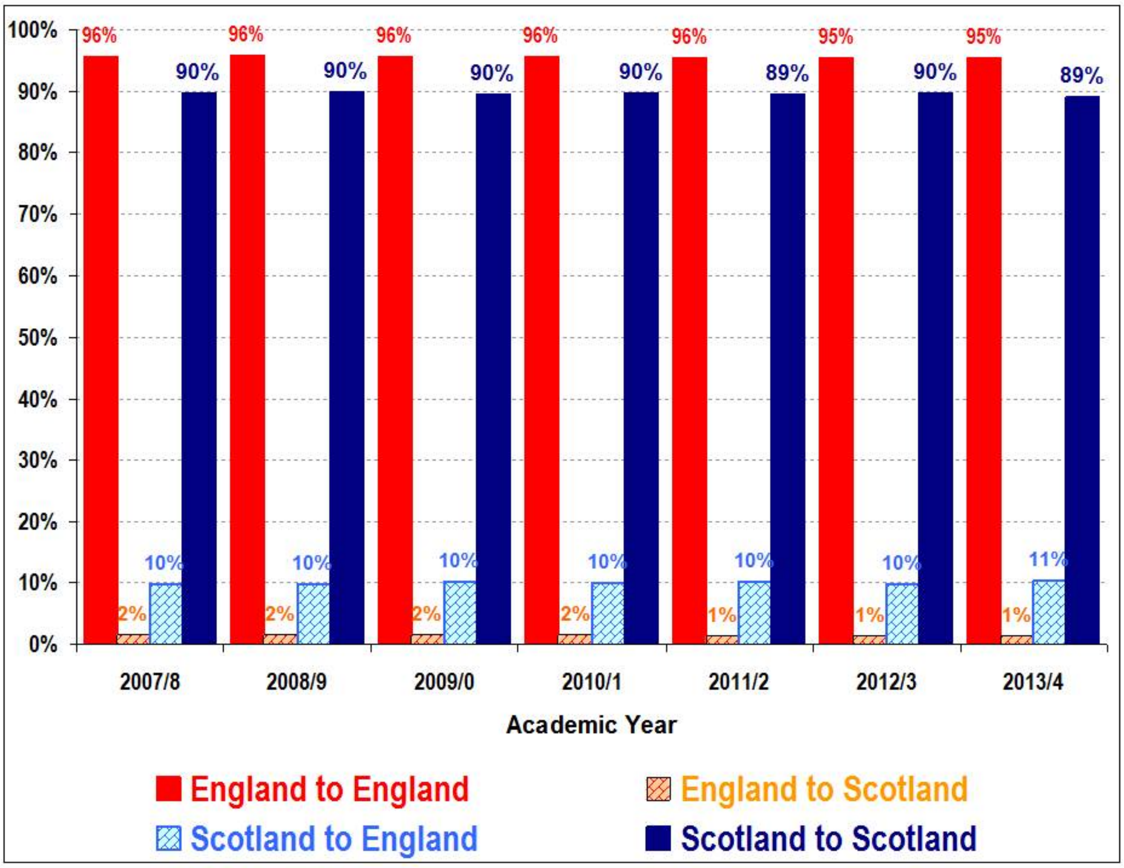


Figure 2.4: **Students Domiciled in England by Region of HE provider 2013/14**

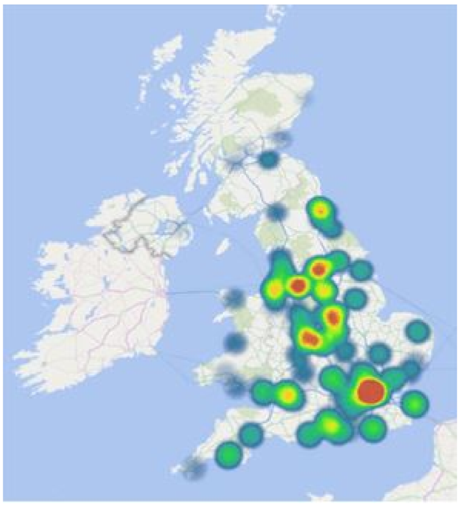
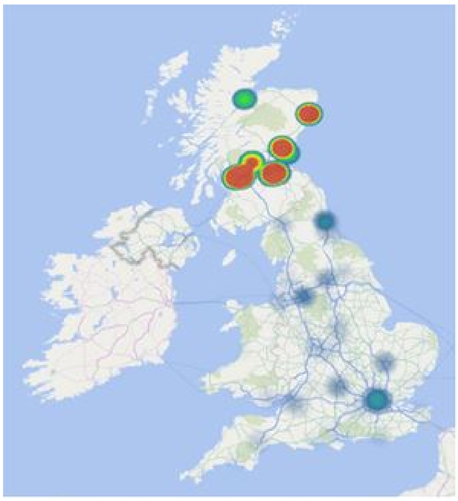


Figure 2.5: **Students Domiciled in Scotland by Region of HE provider 2013/14**



Financial consideration is a key element for school-leavers when they decide whether to obtain higher education and where to pursue their degree. There is no doubt that student migration across countries in the United Kingdom was present before and after the increase in the tuition fees cap in 2012, albeit in small numbers. However, Figure 2.3 reveals that migration flows between England and Scotland have not changed drastically over time. This happens despite the 2012 reform and the fact that the Scottish group are eligible for free higher education. It could still be thought that the 2012 reform has had an effect on the control group (students domiciled in Scotland), but the figures shown give evidence that the migration flows between administrations do not reflect this.

2.5 Methodology

Puebla and Velasco (2006) pointed out that when a student finishes college and decides to pursue a higher education degree, she has to select one out of several alternative subjects. It is assumed that students choose the subject that maximises their utility.

After the 2012 Tuition Fees increase reform, most students have to pay back the tuition fees and maintenance loan. As a result of the 2012 tuition fees cap increase, the borrowed amount (debt) has increased approximately £6,000 in comparison to the previous years. The underlying assumption of the econometric model is that, as a result of the higher education reform, students will choose subjects that most likely ensure the repayment of the loan in the future. As a consequence, students change their subject choice and maximise their utility in making that decision, taking into account prospective futures. As mentioned in the literature review, papers supporting different graduate premium by subjects include Britton et al. (2016), Kirkeboen et al. (2016) and Walker and Zhu (2013).

In broad terms, following Greene (2003), the reason why an individual i chooses to study subject j is because, out of all the the alternatives (other possible subjects), subject j represents the maximum utility, U_{ij}. According to this author, the econometric model is driven by the probability that choice j is made, and can be expressed as:

$$\text{Prob}(U_{ij} > U_{ik}) \quad \text{for all other} \quad k \neq j \tag{2.1}$$

The utility function takes the form of

$$U_{ij} = \theta_j X_i + \alpha_j T_i + \beta_j D_i + \gamma_j T_i D_i + \varepsilon_{ij} \tag{2.2}$$

Y_i, the outcome variable, is defined as a random variable that indicates the chosen subject j. The student has a multiple choice among: 0=Arts & Humanities, 1=Social Science, 2=Pure Science, 3=Life Science and Health and 4=Engineering & Architecture. X is the vector containing control variables including parental and individual characteristics such as parental highest education, ethnicity, ATP and other independent variables. As mentioned in the Data and Variables section, T and D, are time and group indicators respectively. The Multinomial Logit Model (MNL) is used to estimate the probability of studying a particular subject:

$$P(Y_i = j) = F(\alpha_j T_i + \beta_j D_i + \gamma_j T_i D_i + \theta_j X_i) = \frac{exp(\theta_j X_i + \alpha_j T_i + \beta_j D_i + \gamma_j T_i D_i)}{1 + \sum_{k=1}^{J} exp(\theta_k X_i + \alpha_j T_i + \beta_j D_i + \gamma_j T_i D_i)} \quad \text{where} \quad j = 0, 1 ... J \tag{2.3}$$

Difference-in-differences in the MNLM

The control group is required as we cannot observe the outcome for the same individual if they were not exposed to the tuition fees increase or in this context, what subject would have been chosen if the tuition fees increase had not occurred. Because at a given point in time the same individual could not have two simultaneous existences, a good estimate of the counter-factual outcome is needed for those students who are affected by the 2012 tuition fees increase reform. The fact that the Higher Education reform has not affected different UK countries in the same way helps us to find a good control group: students domiciled in Scotland who study in an institution based in Scotland are not affected by the 2012 tuition fees cap increase. In fact, the Scottish domiciled students are eligible for free higher education and are not required to pay university tuition fees.

The reason why Scottish students are likely to be a satisfactory control group is that they share common characteristics with the English students, except for the alleged treatment. The main difference between these two groups of students is where their higher education institution is located. Therefore, the treatment group will be defined as those students who are domiciled in England while the non-treated group will be students whose region of domicile is Scotland.

In order to evaluate the impact of the higher education reform in 2012, the difference-in-differences approach is used in conjunction with the subject choice model. This evaluation method has a main advantage: temporal effects in the subject choice can be taken into account. An assumption of the model is that any unobserved heterogeneity between students that is time invariant is controlled for. Puhani (2012) shows how to calculate the difference-in-differences estimator for a nonlinear model but this author only considers binary models. Extending that methodology to the MNLM[12] allows us to calculate the average treatment effect on the treated (ATET). Depending on the treatment, there are two potential outcomes, $Y_i^{(1)}$ and $Y_i^{(0)}$. The first potential outcome $Y_i^{(1)}$ denotes the outcome if the student receives treatment, while the second potential

[12] The model presented in this chapter is a MNL with five possible choices. In order to check its robustness by means of a logit-DID, these categories have been collapsed into two: 1 if the student chooses a STEM subject (Pure Science, Life Science Health, Engineering), 0 otherwise (Arts Humanities or Social Science). Although the logit ATET is not directly comparable to the ATETs in the MNL model since subjects are collapsed in the logit model, the interaction term and the estimated ATET effect (TxD) is equally statistically significant in both.

outcome, $Y_i^{(0)}$, denotes the outcome if no treatment was received. Following Puhani (2012), the difference in the potential outcomes of the dependent variable is:

$$\tau(T_i = 1, D_i = 1, X_i) = E[Y_i^{(1)} | T_i = 1, D_i = 1, X_i] - E[Y_i^{(0)} | T_i = 1, D_i = 1, X_i]. \qquad (2.4)$$

As we can see in equation (2.4) the treatment effect is obtained by the difference in two expectations. Equation (2.5) indicates the participation in the treatment (I)

$$I = 1[T_i = 1, D_i = 1] = T_i \times D_i \qquad (2.5)$$

where $1[\cdot]$ is the indicator function taking the value of 1 if the expression in brackets is true and 0 otherwise. I will be equal to 1 if the student is domiciled in England and starts their undergraduate degree in the academic year 2013/14. Assuming common trends in U_{ij} it can be shown that the conditional expectations of the potential outcomes $Y_i^{(0)}$ and $Y_i^{(1)}$ respectively are

$$E[Y_i^{(0)} | T_i, D_i, X_i] = F(\alpha_j T_i + \beta_j D_i + \theta_j X_i) \qquad (2.6)$$

which is the counterfactual and

$$E[Y_i^{(1)} | T_i, D_i, X_i] = F(\alpha_j T_i + \beta_j D_i + \gamma_j + \theta_j X_i). \qquad (2.7)$$

Substituting equations (2.7) and (2.6) in (2.4) gives the treatment effect in the 'difference-in-differences' multinomial logit model,

$$\tau(T_i = 1, D_i = 1, X_i) = F(\alpha_j + \beta_j + \gamma_j + \theta_j X) - F(\alpha_j + \beta_j + \theta_j X_i). \qquad (2.8)$$

In order to calculate the ATET for a particular subject, ATET_{P_j}, we need to average the treatment effect obtained in equation (2.8) over the treated sample which gives

$$\text{ATET}_{P_j} = \frac{1}{N^1} \sum_{i=1}^{N} T_i D_i \{F(\alpha_j + \beta_j + \gamma_j + \theta_j X_i) - F(\alpha_j + \beta_j + \theta_j X_i)\} \qquad (2.9)$$

where $N^1 = \sum_{i=1}^{N} T_i D_i$. [13]

Regarding the statistical significance of the estimate of the ATET, the delta method or bootstrapping must be applied in order to obtain correct standard errors.

Robustness test using weights.

In order to test the ATETs obtained for the general model, entropy balancing will be performed. Hainmueller and Xu (2013) describe the way to implement this multivariant re-weighting method to obtain a balanced sample. One of the advantages of the method is that it decreases model dependency for the analysis of treatment effects and provides weights that satisfy a set of balance constraints. For this analysis, as the majority of covariates are binary variables we only need to balance on the first moment of the covariate distributions in the treatment and control group. Once the weights are created, the means of the covariates are very similar (balanced) among the control group and the treated group. The particular advantage of this method to obtain the weights is that, based on the covariates chosen, control units who have more similar units in the treated group will be given a higher weight. After obtaining the weights following this procedure, we run the MNL model but this time adding as sampling weights the weights

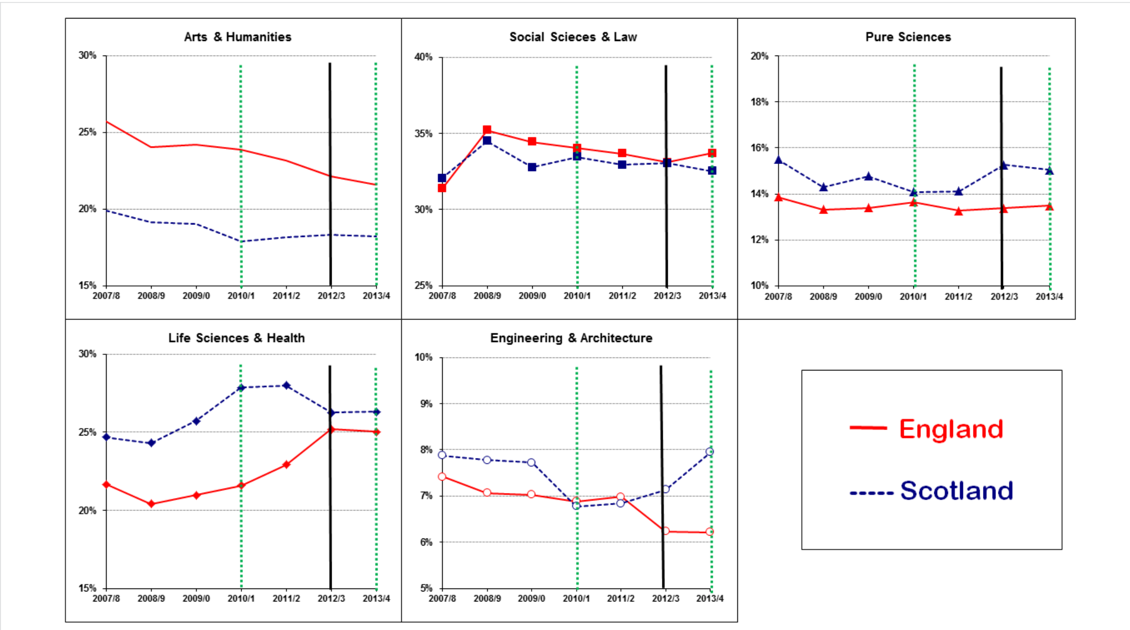
[13] Lechner (2011) proposes an alternative method to calculate the ATET_{P_j} using an indicator of the observed outcome rather than the predicted probabilities for the first term in equation (2.8). Due to the property of the multinomial logit model specified in equation 15.8 in Cameron and Trivedi (2005), the approach proposed by Lechner (2011) will give the same coefficient to that followed by Puhani (2012) to calculate ATET_{P_j}.

obtained following Hainmueller and Xu (2013). The covariates chosen to match the moments are explained in the results section. If the ATETs coefficients are similar, this can increase confidence in the results obtained.

2.6 Results

Subject Choice Time Trends in England and Scotland

Figure 2.6: Percentage of Students by Subjects and Country of Study



The use of the difference-in-differences methodology requires common pre-treatment trends in the treated and control groups (the methodology section contains detailed information regarding these groups). Aggregated data from HEIDI is used to check the subject choice time trends in England (the treated group) and Scotland (the control group) in Figure 2.6. The population represented are students with the following characteristics: full time, first degree, first year and from the United Kingdom. The main subject categories include the following groups:

- Arts & Humanities: Creative Arts, Design, Historical and Philosophical Studies and Languages.
- Social Science & Law: Business and Administration, Law, Mass Communication, Documentation, Sociological Studies and Education.
- Pure Science: Computer Science, Mathematical Science and Physical Science.
- Life Science & Health: Agriculture and Related, Biological Science, Medicine, Dentistry, Subjects allied to Medical Studies and Veterinary Science.
- Engineering & Architecture: Building, Planing, Engineering and Technology.

A few comments regarding Figure 2.6 are given. The graphs express the percentage of students studying a particular subject out of the number of students in each country. It can be seen that, for both countries, Arts & Humanities subjects have become less popular over time. This might be explained by a shift in the young people's decisions. Before the outburst of the economic crisis, when job prospects were not limited, young people who were uncertain about what they really wanted to study tended choose Arts & Humanities. However, nowadays, young people face a different situation. Having a degree in science or maths, for example, may allow them to access a wide range of jobs in the future. Graphs showing the percentage of students studying Social Science & Law and Engineering & Architecture (Figure 2.6) in 2009/10 reveal the same pattern: a decrease in the number of people studying those subjects. The reverse occurs in the graphs corresponding to Pure Science and Life Science & Health. On the whole and over the years represented on the graphs, trends for Arts & Humanities, Social Science & Law and Pure Sciences are flatter and, more importantly, have negative rather than positive slopes than Life Sciences & Health and Engineering & Architecture. For the vast majority of the analysed years, the percentage of students doing Arts & Humanities and Social Science & Law is lower in Scotland than in England while the contrary is true for Pure Science, Life Science & Health and Engineering & Architecture. What needs to be emphasised is that, before the higher education reform, there is presence of common trends and, more importantly, the graph regarding Health & Life Sciences and Arts & Humanities supports the research hypothesis that we are testing: students after the reform in England are more likely to choose Health & Life Sciences and are less likely to choose Arts & Humanities. We will see that this argument is supported as well by the ATETs coefficients reported later. Overall, Figure 2.6 shows that the trends do not dramatically differ between countries before the reform.

Table 2.2 presents the average treatment effects on the treated for the different subjects. These effects, obtained using equation (2.8), give the difference in outcomes on the treated and control group and indicate the impact of the higher education reform on the subject choice. Rows 1-5 focus on the possible outcomes in the multinomial logit model. Table 2.2 is divided into 5 sets of columns with 3 columns in each set. Estimates from the first set of columns present the subject choice model for all parental occupations while in the other 4 sets of columns, the model is presented for a specific group of parental occupations, used as a proxy for socioeconomic status. The motivation for estimating several models depending on the parental occupation is that it might be the case that the increase in the tuition fees cap had a different effect across individuals from distinct socioeconomic backgrounds. In each set of columns, the model for the career choice is firstly estimated on all the students, and subsequently on male and female students to investigate whether the higher education reform in 2012 had different effects on boys and girls due to their different attitudes towards debt and risk aversion.

If we focus on the ATET for the alternative subjects in rows 1-5, a general pattern arises. Due to the nature of the model, any of the ATET using the difference-in-differences approach in this particular framework could be interpreted as follows: for example, for the first coefficient, the fall in Arts & Humanities is larger in England following the increase in fees than in Scotland. After the increase in the tuition fees cap, students are less likely to study Arts & Humanities and Social Science while they are more likely to study Health & Life Science. This could be explained by the fact that students wish to ensure that the high tuition fees payment will lead to high expected future earnings and better employment prospects as Sa (2014) argues. According to HESA, out of all UK domiciled full-time university graduates who obtained first degree qualifications and entered full-time paid work in the UK in 2013/14, those who studied a degree in Medicine and Dentistry, Veterinary or Engineering earn, on average, the highest annual first salary. [14] On the other hand, students who hold a degree in Arts and Design receive the lowest salary after graduation. This information could explain why the signs of the coefficients are generally negative and significant for the ATET in Arts & Humanities and positive in Health & Life Sciences and is in accordance with what Sa (2014) found about students considering expected future earnings and employment prospects when choosing a degree subject.

Regarding the sign of the coefficients for the other subjects, there is a less clear pattern due to coefficients changing sign depending on the different parental occupations. Turning now to the ATET for all the subjects, it is clear that out of all the statistically significant coefficients, the ones from the unemployed parents are the largest in size. This implies that the higher education reform in 2012 had a larger effect on the probability of choosing a

[14] Averages obtained from HESA Destinations of Leavers from Higher Education survey.

particular degree for students with unemployed parents than students with parents holding an occupation.

Although there seems to be a pattern for the ATET signs for all the students, there are differences between boys and girls' choices. The existence of gender differences when choosing a degree subject is supported by Leppel et al. (2001), but that study did not control for any tuition fees cap increase. In this study, gender differences are more present in Pure Science and Engineering than in the remaining subjects. It can be seen that, although the coefficients are not statistically significant, a negative coefficient pattern arises for girls when estimating the ATET on Engineering and Architecture. After the reform, girls are less likely to study these subjects than if no higher education reform occurred. Purcell et al. (2008) show gender differences figures: women prove career planning with a short term view while men are more intrinsically-oriented to the entire career and further employment opportunities when choosing a subject. As a consequence, female presence in engineering degrees is often poor. Some campaigns such as WISE[15] have been developed in the UK to mitigate gender differences regarding participation in Engineering and to encourage girls to pursue a career in technical subjects. An explanation of a different reaction to fees change for these particular subjects could be related with gender and debt aversion. Davies and Lea (1995) find that men are more likely to be in debt than women. As mentioned before there was already an existing trend of low female participation in technical subjects in the past years. The majority of degrees that contain technical modules such as Engineering & Architecture, on average require more effort and time: effort in terms of the high grades in A-Levels

[15] Campaign that inspires girls and women to study and build careers using science, technology, engineering and maths (STEM).

required to access these degrees and time in terms of length of study. After finishing their degree in Engineering & Architecture, students could become chartered. This requires extending the study period to gain professional competencies. Girls might not be willing to choose longer courses that will accumulate more debt over time. Although a degree in Engineering or Architecture leads to a high graduate salary, girls are less likely to study Engineering degrees after the reform while the contrary occurs for boys. In the future, if this pattern continues, inequality in the workforce could arise and have policy implications.

Going back to the statistical analysis and turning the attention to students with unemployed parents, although the majority of the ATETs are not statistically significant, there are no differences in the signs for the coefficients between girls and boys; therefore, the higher education reform impacted on the probability of studying a particular subject in the same direction for both genders among those with unemployed parents.

In absolute terms, the size of the treatment effect is larger for Health & Life Science than for the other subjects. Regarding Health & Life Science, the effect is larger and statistically more significant for female students. Girls affected by the reform are less likely to study Arts & Humanities relative to girls domiciled in Scotland from the control group with the effect being significant at the 1 percent level.

On the other hand, when splitting the sample according to parental occupation, it can be seen that Managers and Professionals' daughters are around 3 percent less likely to study Arts & Humanities and around 1 percent less likely to study engineering after the reform, while the sons are around 2 percent less likely to pursue a career in Pure Science after the reform. Sons and daughters are together more likely to choose a degree in

Health & Life Science, with girls being around 2 percent more likely to study these subjects than boys. Although Purcell et al. (2008) argue that students of higher socioeconomic status choose more enjoyable subjects or subjects where students have been successful, these results show that, after the higher education reform, individuals are more likely to chose subjects that will lead to better employment prospects. Similarly, students whose parents hold an intermediate occupation are around 4 percent more likely to choose a degree in Health & Life Science after the reform. According to Wilson and Homenidou (2011), if qualification patterns by occupation are reviewed, it can be seen that the most frequent highest qualification that individuals in intermediate occupations hold are intermediate vocational or school leaving qualifications. It might be the case that sons or daughters of parents in intermediate occupations behave according to the Human Capital Theory and proceed one step further than their parents pursuing higher education and therefore, are more likely to study a degree that will ensure a higher paid job such as those in Health & Life Science. This argument is related to the negative and less statistically significant coefficients for those students choosing a degree in Social Science, a heterogeneous subject group including pursuers of degrees that offer very attractive job prospects such as Economics and Law or less attractive like Mass Communication and Documentation[16].

There is no clear and straightforward explanation for the coefficients that refer to students with parents in elementary occupations, probably due to the latter being a very heterogeneous group comprising occupations from sales to process plant and machine operatives. Column 10 shows that students are around 2 percent more likely to study a

[16] According to the DLHE survey out of all the social science subjects, Mass Communications & Documentation has the lowest mean salary for UK domiciled full-time leavers who obtained first degree qualifications and entered full-time paid work in the UK in 2013/14.

degree in Engineering & Architecture and column 11 shows that, in particular, boys are 4 percent more likely to study these degrees after the reform.

To conclude this section, it is worth mentioning that, after the higher education reform in 2012, students with unemployed parents who could come from disadvantaged backgrounds or may not be well informed about higher education choices and career prospects are around 8 percent more likely to choose a degree in Health & Life Science. In particular boys are 10 percent more likely to choose one subject within this group of subjects than if no reform was implemented as respectively columns 13 and 14 from Table 2.2 show.

Table 2.3 shows the robustness test for the models. In order to get these ATETs coefficients the weights have been generated after running the initial model for the subject choice. Once we perform entropy balancing on the sample means of the treated and control group and the weights are generated according to this process, then the ATETs are obtained using a weighted multinomial logit model. The covariates chosen to balance the sample are gender, parental occupation, ATP, ethnicity and parental highest education. Age is not used because depending on the country of domicile and due to differences in the education system in England and Scotland, this variable differentiates the sample rather than homogenises it. We can see in Table 2.3 that after balancing the sample the results are very similar to the initial ones from column 1 in Table 2.2 (the statistically significant ATETs from Table 2.3 are 0.4 percentage points smaller than the initial ones). According to the weighted regression students after the reform are less likely to study a degree in Arts & Humanities and more likely to study Health & Life Sciences.

Marginal effects of the control variables from the MNLM are shown in Table 2.4. These coefficients represent the marginal effect that each of the control variables have on the probability of choosing a specific subject of study. The way they differ from the ATET is that these marginal effects are obtained for all the student population, not only for the treated. As a general comment, it can be seen that the majority of coefficients are statistically significant, therefore the chosen control variables are relevant for the subject choice. When students from both gender are included, there are gender differences when choosing a degree subject and males are less likely to choose Arts & Humanities, Social Science or Health & Life Sciences subjects and more likely to choose Pure Sciences or Engineering & Architecture. If ATP is used as a proxy for ability, high ability students tend to choose less a degree in Arts & Humanities or Social Science and are more likely to choose Health & Life Sciences, Engineering & Architecture or Pure Sciences. Having at least one of the parents or both who have attained higher education has an impact on the probability of their sons/daughters choosing a particular subject except for girls that choose Pure Sciences or Health & Life Sciences. Given that the main scope of this chapter is to evaluate the effect of the higher education reform in 2012 on subject choice no more attention is drawn to these coefficients.

2.7 Conclusions

This paper exploits the 2012 increase in tuition fees in higher education cap to estimate its causal effect on English domiciled students' subject choice. A multinomial logistic regression is used to model the subject choice and two cross-sections are employed for the statistical analysis —one in 2010/11 before the reform occurred to avoid any interference from anticipation effects, and another in 2013/14 after the reform occurred. As Scotland was not affected by this higher education reform, it was used as a counterfactual case for the difference-in-differences methodology.

The main results of this study show that, as a result of the 2012 increase in the tuition fees cap, English domiciled students —no matter their socioeconomic status— are 2 percentage points less likely to study a degree in Arts & Humanities and around 3 percentage points more likely to study a degree in Health & Life Sciences. This evidence strongly suggests that after the higher education reform students take into account employment prospects when deciding the subject of study. In particular, English students with unemployed parents have experienced the largest impact of the education reform in terms of ATET's size in comparison with students whose parents hold an occupation.

From a methodology point of view, this paper extends the work presented by Puhani (2012) in the context of non-linear models and combines a multinomial logistic model with a difference-in-differences approach in order to evaluate the impact of the latest higher education reform on subject choice among English students. Controls for sex, parental occupation, ethnicity, parental education, as well as students ability (derived from their associated tariff points) are included in the model. Even though available data did not allow to formally test for common time trends by means of dummy variables for different years, and despite the limited number of variables included in the model, the results obtained seem to be solid. The trends are graphically tested using aggregated data from HEIDI and they do not seem to differ one from another.

Several papers have dealt with different dimensions regarding the impact of the 2012 higher education reform in England, but the present study throws new light on its possible effects on subject choice. If the Government decides to continue increasing the tuition fees cap in future years, it might consider these effects apart from the ones already studied in the literature such as those related to widening participation. A concentration of students in certain subjects favored by the reform could arise and, therefore, affect the potential labour force supply in England.

Chapter 2 Tables

Table 2.1: Characteristics of English Domiciled and Scottish Domiciled

	Domiciled England	Domiciled Scotland	Domiciled England		Domiciled Scotland	
			Pre-Reform 2010/11	Post-Reform 2013/14	Pre-Reform 2010/11	Post-Reform 2013/14
	(1)	(2)	(3)	(4)	(5)	(6)
Subject						
Arts & Humanities	0.234	0.124	0.243	0.226	0.124	0.125
Std. Dev.	0.424	0.330	0.429	0.418	0.329	0.331
Social Science	0.314	0.299	0.317	0.311	0.299	0.299
Std. Dev.	0.464	0.458	0.465	0.463	0.458	0.458
Pure Science	0.130	0.144	0.129	0.131	0.141	0.147
Std. Dev.	0.336	0.351	0.335	0.338	0.348	0.354
Health & Life Science	0.238	0.296	0.225	0.250	0.299	0.293
Std. Dev.	0.426	0.456	0.418	0.433	0.458	0.455
Engineering & Architecture	0.083	0.137	0.085	0.081	0.137	0.137
Std. Dev.	0.276	0.344	0.279	0.273	0.344	0.343
Male	0.464	0.439	0.470	0.458	0.440	0.438
Std. Dev.	0.499	0.496	0.499	0.498	0.496	0.496
Age Dummies						
17, 16 & Under	0.002	0.288	0.003	0.002	0.292	0.284
Std. Dev.	0.049	0.453	0.053	0.044	0.455	0.451
18 (base)	0.601	0.455	0.594	0.608	0.462	0.448
Std. Dev.	0.489	0.497	0.491	0.488	0.498	0.497
19	0.299	0.156	0.303	0.294	0.154	0.158
Std. Dev.	0.458	0.363	0.460	0.456	0.361	0.365
20	0.097	0.101	0.099	0.095	0.091	0.110
Std. Dev.	0.296	0.301	0.299	0.294	0.288	0.313
Associated Tariff Points	453.083	435.949	430.817	472.907	415.907	454.640
Std. Dev.	200.911	208.015	202.049	197.795	207.486	206.770
Ethnicity						
White (base)	0.755	0.933	0.771	0.739	0.941	0.925
Std. Dev.	0.429	0.249	0.419	0.438	0.234	0.263
Asian	0.131	0.041	0.125	0.136	0.038	0.045
Std. Dev.	0.337	0.199	0.330	0.343	0.190	0.208
Black	0.061	0.006	0.055	0.066	0.006	0.007
Std. Dev.	0.239	0.079	0.228	0.249	0.076	0.082
Other (including Mixed)	0.053	0.019	0.049	0.057	0.015	0.023
Std. Dev.	0.224	0.136	0.215	0.232	0.121	0.149
Parental Higher Education	0.536	0.653	0.549	0.525	0.658	0.649
Std. Dev.	0.499	0.476	0.498	0.499	0.474	0.477
Parental Occupations						
Managers & Professionals	0.588	0.635	0.599	0.578	0.636	0.635
Std. Dev.	0.491	0.481	0.489	0.493	0.480	0.481
Intermediate Occ.	0.178	0.176	0.184	0.173	0.182	0.171
Std. Dev.	0.383	0.381	0.387	0.379	0.386	0.376
Elementary Occ.	0.204	0.173	0.190	0.217	0.168	0.177
Std. Dev.	0.403	0.378	0.392	0.412	0.374	0.381
Unemployed	0.029	0.015	0.027	0.032	0.013	0.017
Std. Dev.	0.169	0.122	0.163	0.175	0.113	0.130
Observations	395,015	37,701	190,901	204,114	16,784	17,917

Table 2.2: Difference-in-Differences

	All Occupations			Managers & Prof.			Intermediate Occ.			Elementary Occ.			Unemployed		
	All	Boys	Girls	All	Boys	Girls	All	Boys	Girls	All	Boys	Girls	All	Boys	Girls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
ATET$_{S_1}$: Arts & Humanities	-0.019***	-0.005	-0.029***	-0.016*	-0.000	-0.028**	-0.025	-0.008	-0.034	-0.028	-0.014	-0.035	-0.089	-0.128	-0.055
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.08)	(0.14)	(0.09)
ATET$_{S_2}$: Social Sciences	-0.008	-0.012	-0.006	-0.006	-0.011	-0.003	-0.029*	-0.046*	-0.018	-0.001	0.018	-0.007	0.030	0.022	0.024
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.02)	(0.00)	(0.03)	(0.02)	(0.07)	(0.11)	(0.09)
ATET$_{S_3}$: Pure Sciences	-0.005	-0.011	0.000	-0.008	-0.019*	0.001	0.013	0.037**	-0.004	-0.003	-0.023	0.006	-0.038	-0.034	-0.025
	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.00)	(0.02)	(0.01)	(0.05)	(0.10)	(0.05)
ATET$_{S_4}$: Health & Life Sciences	0.028***	0.013*	0.041***	0.029***	0.021**	0.039***	0.038***	0.002	0.060***	0.017	-0.021	0.037**	0.083**	0.105***	0.051
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.00)	(0.02)	(0.02)	(0.04)	(0.03)	(0.06)
ATET$_{S_5}$: Engineering & Architecture	0.004	0.015***	-0.006**	0.000	0.009	-0.009**	0.004	0.014	-0.004	0.015***	0.040***	-0.001	0.014	0.036	0.005
	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.02)	(0.05)	(0.01)
Observations	146,762	65,432	81,330	86,120	39,571	46,549	25,150	10,952	14,198	31,020	12,982	18,038	4,472	1,927	2,545

Notes: A MNLM is used to estimate the predicted probabilities based on treated and control individuals. These probabilities are employed to calculate the ATET$_{P_j}$ using equation (2.8) for the treated sample. The delta method is used to obtain the standard errors. The significance levels are * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. There are three different specifications of the model. Control variables in all the models are age, ethnicity, parental education dummies, atp, T, D and the interaction term TD. Column 1 contains gender and occupation variables as well. Apart from the control variables included in all models, columns 2 and 3 contain parental occupation dummies. The models in columns 4, 7, 10 and 13 control for gender and parental occupation.

Table 2.3: ATET using a weighted MNLM

	All	
ATET_{S_1}: Arts & Humanities	-0.023	***
	(0.01)	
ATET_{S_2}: Social Sciences	0.001	
	(0.01)	
ATET_{S_3}: Pure Sciences	-0.005	
	(0.01)	
ATET_{S_4}: Health & Life Sciences	0.024	***
	(0.01)	
ATET_{S_5}: Engineering & Architecture	0.003	
	(0.00)	
N	146.762	

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.4: Average Marginal Effects from the Multinomial Logit models

	Arts & Humanities			Social Sciences			Pure Sciences			Health & Life Sciences			Engineering & Architecture		
	All (1)	Boys (2)	Girls (3)	All (4)	Boys (5)	Girls (6)	All (7)	Boys (8)	Girls (9)	All (10)	Boys (11)	Girls (12)	All (13)	Boys (14)	Girls (15)
Male	-0.084***			-0.059***			0.137***			-0.113***			0.120***		
	(0.00)			(0.00)			(0.00)			(0.00)			(0.00)		
Unemployed	-0.014**	-0.005	-0.020***	0.013**	-0.007	0.029***	0.021***	0.047***	-0.000	-0.013***	-0.018***	-0.008	-0.008**	-0.017***	-0.001
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
Intermediate Occ.	-0.010***	-0.010***	-0.009***	-0.012***	-0.021***	-0.005	0.011***	0.019***	0.004**	0.006**	0.003	0.008**	0.005***	0.009***	0.002
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Elementary Occ.	-0.016***	-0.016***	-0.015***	-0.008***	-0.022***	0.000	0.017***	0.036***	0.003	0.010***	0.007**	0.013***	-0.003**	-0.005*	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
17, 16 & Under Years Old	-0.022***	-0.031***	-0.015*	-0.012*	-0.024**	-0.005	0.023***	0.035***	0.016***	-0.004	-0.007	-0.004	0.016***	0.028***	0.009***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
19 Years Old	0.068***	0.047***	0.085***	-0.030***	-0.001	-0.054***	-0.029***	-0.042***	-0.019***	-0.005**	0.005*	-0.012***	-0.004***	-0.009***	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
20 Years Old	0.041***	0.049***	0.034***	-0.040***	-0.027***	-0.050***	-0.013***	-0.014***	-0.015***	0.013***	-0.002	0.029***	-0.001	-0.005	0.002
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Associated Tariff Points	-4.8$x10^{-5***}$	-7.9$x10^{-5***}$	-1.8$x10^{-5***}$	-2.3$x10^{-4***}$	-1.4$x10^{-4***}$	-3.2$x10^{-4***}$	1.3$x10^{-4***}$	1.4$x10^{-4***}$	1.2$x10^{-4***}$	1.1$x10^{-4***}$	3.5$x10^{-4***}$	1.1$x10^{-4***}$	0.4$x10^{-5}$ ***	4.3$x10^{-5***}$	3.6$x10^{-5***}$
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Asian	-0.173***	-0.156***	-0.185***	0.054***	0.059***	0.049***	0.008***	-0.006	0.019***	0.081***	0.054***	0.103***	0.030***	0.049***	0.014***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Black	-0.139***	-0.112***	-0.161***	0.094***	0.083***	0.103***	-0.024***	-0.035***	-0.016***	0.034***	0.012**	0.053***	0.036***	0.052***	0.021***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
Other Ethnicity	-0.036***	-0.040***	-0.032***	0.023***	0.024***	0.023***	-0.015***	-0.024***	-0.008***	0.008**	0.019***	0.001	0.020***	0.021***	0.017***
	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Parental High Educ.	0.033***	0.020***	0.044***	-0.045***	-0.031***	-0.055***	-0.002	-0.006**	0.001	-0.004**	-0.007***	-0.002	0.017***	0.024***	0.012***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Ac. Year 2013/14 (T)	0.011*	0.002	0.018**	0.007	0.006	0.010	0.003	0.017**	-0.007	-0.015***	-0.007	-0.022***	-0.007**	-0.018***	0.002
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
Treatment (D)	0.122***	0.095***	0.143***	0.018***	0.043***	0.003	-0.001	-0.005	0.002	-0.083***	-0.016**	-0.138***	-0.057***	-0.116***	-0.010***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
Interaction ($T \times D$)	-0.019***	-0.005	-0.029***	-0.008	-0.012	-0.006	-0.005	-0.011	0.000	0.028***	0.013*	0.041***	0.004	0.015***	-0.006**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
Observations	146,762	65,432	81,330	146,762	65,432	81,330	146,762	65,432	81,330	146,762	65,432	81,330	146,762	65,432	81,330

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1. Subject grouping using JACS. *Arts & Humanities*: Broadly-based programs within languages, Linguistics, Comparative literary studies, English studies, Ancient language studies, Celtic studies, Latin studies, Classical Greek studies, Classical studies, Others in linguistics, classics & related subjects, French studies, German studies, Italian studies, Spanish studies, Portuguese studies, Scandinavian studies, Russian & East European studies, European studies, Others in European languages, literature & related subjects, Chinese studies, Japanese studies, South Asian studies, Other Asian studies, African studies, Modern Middle Eastern studies, American studies, Australasian studies, Others in Eastern, Asiatic, African, American & Australasian languages, literature & related subjects, Broadly-based programs within historical & philosophical studies, History by period, History by area, History by topic, Archeology, Philosophy, Theology & religious studies, Heritage studies, Others in historical & philosophical studies, Broadly-based programs within creative arts & design, Fine art, Design studies, Music, Drama, Dance, Cinematics & photography, Crafts, Imaginative writing, Others in creative arts & design; *Social Science*: Broadly-based programs within education, Training teachers, Research & study skills in education, Academic studies in education, Others in education, Broadly-based programs within social studies, Economics, Politics, Sociology, Social policy, Social work, Anthropology, Human & social geography, Development studies, Others in social studies, Broadly-based programs within law, Law by area, Law by topic, Others in law, Broadly-based programs within business & administrative studies, Business studies, Management studies, Finance, Accounting, Marketing, Human resource management, Office skills, Hospitality, leisure, sport, tourism & transport, Others in business & administrative studies, Broadly-based programs within mass communications & documentation, Information services, Publicity studies, Media studies, Publishing, Journalism, Others in mass communications & documentation; *Pure Science*: Broadly-based programs within physical sciences, Chemistry, Materials science, Physics, Forensic & archaeological sciences, Astronomy, Geology, Science of aquatic & terrestrial environments, Physical geographical sciences, Others in physical sciences, Mathematics, Operational research, Statistics, Others in mathematical sciences, Computer science, Information systems, Software engineering, Artificial intelligence, Health informatics, Games, Computer generated visual & audio effects, Others in computer sciences; *Health & Life Science*: Broadly-based programs within medicine & dentistry, Pre-clinical medicine, Pre-clinical dentistry, Clinical medicine, Clinical dentistry, Others in

medicine & dentistry, Broadly-based programs within subjects allied to medicine, Anatomy, physiology & pathology, Pharmacology, toxicology & pharmacy, Complementary medicines, therapies & well-being, Nutrition, Ophthalmics, Aural & oral sciences, Nursing, Medical technology, Others in subjects allied to medicine, Broadly-based programs within biological science, Biology, Botany, Zoology, Genetics, Microbiology, Sport & exercise science, Molecular biology, biophysics & biochemistry, Psychology, Others in biological sciences, Broadly-based programs within veterinary sciences, agriculture & related subjects, Pre-clinical veterinary medicine, Clinical veterinary medicine & dentistry, Animal science, Agriculture, Forestry & arboriculture, Food & beverage studies, Agricultural sciences, Others in veterinary sciences, agriculture & related subjects; *Engineering & Architecture*: Broadly-based programs within engineering & technology, General engineering, Civil engineering, Mechanical engineering, Aerospace engineering, Naval architecture, Electronic & electrical engineering, Production & manufacturing engineering, Chemical, process & energy engineering, Others in engineering, Minerals technology, Metallurgy, Ceramics & glass, Polymers & textiles, Materials technology not otherwise specified, Maritime technology, Biotechnology, Others in technology, Broadly-based programs within architecture, building & planning, Architecture, Building, Landscape & garden design, Planning (urban, rural & regional), Others in architecture, building & planning.

Chapter 3 Returns to Education in Spain. Revisiting Family Background Variables as Instrumental Variables

3.1 Introduction

To continue assessing how important education is today, we quote some words by the OECD Deputy-Director of Education, Schleicher (2009): ‘It is shown that those from disadvantaged backgrounds have still low chances to succeed in education; even in the richest nations in the OECD, the chances of moving up to high education are only about half if the parents didn’t succeed in school’. It can be seen that those who didn’t complete high school have experienced rapidly deteriorating prospects. Since the turn of the century, highly educated people have seen their job prospectives grow dramatically. The medium and low educated have seen theirs decline.

Why study the Returns to Education in Spain and not the causes of the economic growth in Ghana? After reviewing the literature, it can be seen that there is not an agreement about the exact returns to education. Why is it still a challenge to find the definitive returns? Depending on the models used or the variables included, the coefficients change (sometimes dramatically). The reason why Spain was chosen is that recently it has suffered from the economic crisis and the education system has changed a few times, so

I thought it is an interesting scenario. Also, as Arrazola and De Hevia (2008) say, Spain has high unemployment rates and the access to employment is competitive. One of the causes of the high youth unemployment rate in Spain is young individuals' lack of experience when seeking a first job.

Spanish politicians are concerned about education. In fact, they believe that education is the engine that promotes the development of a country; if a citizen acquires more education, this will enable him/her to get to a higher position in the workplace and this will help the economic growth of the country. What the politicians are trying to avoid is an increase in the drop out rate and youth unemployment.

The exact coefficient for the private returns to education and what estimation method should be used is today still a topic of discussion. Researchers debate what variables should the model include as covariates in the wage equation and which estimation method is the one that performs best.

In this chapter, the two stages least squares method is used to overcome the endogeneneity problem in the schooling variable. What it is empirically explored is which instrumental variables perform best to explain the difference in attained education. The vast majority of papers in the Spanish returns to education literature focus more on instrumenting years of schooling. In this study both types of education variables are instrumented: years and levels. By finding suitable instruments for the levels of education attained we can see what variables affect the attainment of the consecutive higher level of education.

This chapter is divided as follows: first, a literature review regarding returns to education is given in section 2; section 3 explains the data set used for the analysis and the variables included in the models; section 4 describes the methodology employed in the estimation;

section 5 reports the results obtained; section 6 concludes the chapter.

3.2 Literature Review and Research Hypothesis

3.2.1 Theory

To understand the returns to education we need to talk about the beginning, the Human Capital Theory. It is believed that education is the engine that promotes the economic progress of a country. Education stimulates economic growth. There are authors such as Schultz (1961) and Denison (1962) that believe that education promotes more economic growth in comparison with a rise in the stock of material capital. They suggest that a significant percentage of the economic growth in the U. S. between 1929 and 1957 was due to a rise in the education of the labour force. Education also contributes to reduce a country's crime rate and encourages people to be more interested in public affairs.

What the Human Capital Theory promotes is that if an individual invests in education, he/she will become more productive and will benefit from future higher earnings. The development of the Human Capital Theory made governments more willing to spend a larger part of economic resources on education. One of the causes of the success of the Human Capital Theory is that it provides a consistent explanation for the empirical evidence on education profitability. A method do this is using Mincer's equation and estimating the coefficient on education in the wage equation.

3.2.2 Estimating the returns to education and some international evidence

It is well known that there is a huge amount of literature on returns to education. Mincer's wage equation is one of the most used tools to estimate these returns. For example Psacharopoulos (1994) and Harmon et al. (2003); the first paper compares returns among different countries; estimation of returns by gender and sector of employment are presented. The second mentioned paper explores functional forms, measures of schooling and social and private returns to education. In this study private returns to education are chosen to be estimated rather than social returns.

Researchers and politicians are interested in both type of returns, social and private. Education is important to single individuals because it can be considered as a channel for them to perceive higher earnings, have a better standard of living and belong to a higher social class. This also has a benefit for the Government in terms of taxes, the higher the individual earns, the higher he/she contributes. Private returns are estimated rather than social returns because of the availability and accuracy of earnings.

Mincer (1974) proposed a semi-logarithmic wage equation to estimate the private returns to education. In this equation, wages vary in a linear manner with the time invested in education and in a quadratic manner with work experience:

$$\ln(W_i) = \beta_0 + \beta_1 S_i + \beta_2 E_i + \beta_3 E_i^2 + \varepsilon_i \tag{3.1}$$

W_i is a wage measure for individual i, S_i is the amount of schooling that individual i

has, E_i is work experience, E_i^2 its square and ε_i represents the error term, assumed to be independent of the rest of variables in the equation and captures other variables not included in the model. Assumptions followed in this model are listed: (i) no direct costs associated with the investment in education (only the opportunity cost of not receiving a salary because of studying), (ii) the individual stays constantly in the labour market (they do not leave the labour market); time that they stay in the labour market is independent of the level of schooling attained and (iii) individuals start to work right away after they finish their studies (no existence of gap years).

The estimation of the returns to education have been controversial since Mincer proposed this equation. Criticisms emerged due to econometric problems that researchers can experience if they use Mincer's method to estimate the returns to schooling. Griliches (1977) lists potential problems that could bias the coefficients: (i) existence of omitted variables such as ability; (ii) the non correct measure of education (measurement error) and (iii) education treated as an exogenous variable i.e. not affected by anything inside the estimation equation. A large part of the economics of education literature focuses on solving these econometric problems in order to get unbiased estimates. Some papers that stand out for doing this are Blackburn and Neumark (1995), Angrist and Krueger (1991, 1992) and Card (1999). Blackburn and Neumark (1995) explore whether the Ordinary Least Squares (OLS) returns to schooling estimates are downwards bias. They studied the omitted-ability bias. In order to do this, the National Longitudinal Survey Youth Cohort (NLSY) was used. Test scores were used to isolate the two ability components: academic and non-academic. They conclude that the education effect (OLS estimate), is lower when controlling for ability. If ability is not taken into account and instrumental variables techniques are used to solve the endogeneity problem, the returns to education are higher than the OLS ones. An explanation of this is given later in the chapter.

Angrist and Krueger (1991) wrote a very well known paper within the economics of education literature using data from United States. The returns to schooling were estimated using month of birth as an instrumental variable for education. These authors propose that the quarter of birth produces a difference to the future amount of education attained. The idea behind this instrument relays on the fact that a person born in the first quarter of the academic year starts school later and reaches the minimum school leaving age before those born in later months, therefore these individuals would have less schooling than those born at the end of the year. Following this, a direct relationship between the attained education and quarter of birth could be established. As an advantage, there is no reason to think that quarter of birth would have a direct effect on earnings therefore no endogeneity problem arises. Several researchers question the validity of this paper. One of these is Bound et al. (1995). They found that Angrist and Krueger's reported estimates suffer from finite-sample bias and therefore are not consistent as expected. Another paper that criticized Angrist and Krueger's work is Hoogerheide and Dijk (2006) that argues that the coefficient obtained in Angrist and Krueger is mainly determined by data from a few Southern states and also, they conclude that the quarter of birth instrument does not affect all the population with the same strength; individuals with 9-13 years of schooling are not that affected in comparison with those who have at most 8 or at least 14 years of schooling. The mechanism explaining how quarter of birth works as an instrument was updated in Hoogerheide and Dijk (2006). The initial idea was that this instrument affects only those about to leave school (who normally would have about 9-13 years of schooling).

Angrist and Krueger (1992) exploit the fact that a person entering university has a lower probability of participating in the Vietnam War. This can be used as an exogenous effect over the demand for schooling. The instrumental variable suggested for this study was the random number given to each person to enter the draft lottery. This paper is another illustration of IV estimates exceeding OLS estimates.

Card (1999) reviews the literature on the causal effect of education on earnings. Four main areas are addressed: theoretic and econometric progress, recent studies that exploit the institutional features of the education system to obtain the returns to education using instrumental variables, recent studies using samples of twins and new studies to evidence the presence of heterogeneity in the returns to education. Twins studies suggest that OLS results are about 10 percentage points upwards biased because of ability. When institutional changes in the education system are used as instruments, the bias between the OLS and the IV approaches ranges between 20 - 40 percentage points. This occurs because of higher marginal returns for certain groups, especially for those with low education levels in comparison with the average marginal returns in the full population.

The papers listed above are some of the "classic" ones in the economics of education literature. Due to the extensive literature related to this topic, this section is divided into two subsections, separating streams related to the Human Capital Theory and Spanish literature review concerning the returns to education.

3.3 Streams related to Human Capital Theory

There was a boom in the number of papers published in the sixties regarding the economics of education. In the seventies, there were economists not in agreement with the Human Capital Theory that promotes education as a way to create a more just and equitable society rather than a highly productive one. Several streams arise against the Human Capital Theory (Becker (1964)). A few of them are reviewed briefly as follows. The first one is the “credentialism” or “signaling theory” in which the level of education is used as a filter or screening mechanism. As the employers have no perfect information beforehand about how productive the candidate is, years of education and the qualifications obtained by the candidate are used as an indicator of productivity. The employer will then offer the individual a proffered job in a company and therefore a higher wage. The education level of an individual acts as a signal of ability, (Spence, 1973).

An advocate of another stream that critiques Human Capital Theory is Thurow (1975). This stream was called “institutionalism” and states that institutions play a more important role in society in comparison to single individuals. The productivity is associated with the workstation rather than individuals and wages depend on the job position held rather than required tasks or activities.

The last stream headed by the Marxist economists Bowles and Gintis (1976) is the most radical one against the Human Capital Theory. These authors believe that the Human Capital theory does not take into account social classes. The individual’s success in education and in the workplace is considered a result of his/her social class of origin. In turn, this success helps to reproduce the division of society into social classes and this occurs generation by generation. Education is a segmentation mechanism of the labour force and can produce inequality in the society.

3.3.1 Empirical evidence from Spain

There are two important pieces of work that review the Spanish literature on this topic. The first one is written by Salas (2002) that reviews the origins of the Economics of Education and devotes a section to explain the proliferation of this field in Spain. The second one is a recent meta-analysis written by Pino (2014).

Pino (2014) review the studies that are already published about the returns to education in Spain. This work compiles more than 80 papers on returns to education in Spain. As these authors mention, the first study about the returns to education in Spain was published in 1978 and since then, there are new papers published every year regarding this topic. To write this meta-analysis, several search engines and databases were used. The study encompasses papers written in English and Spanish.

Pino (2014) believe that what is estimated should be called an education wage premium rather than returns to an investment in education. The explanation given is the following: in Mincer's equation, in order to calculate the returns to an additional year of schooling, several things should be taken into account such as the direct cost of education, the chance to have a part-time job while studying, the length of working life and the exogenous growth of productivity. None of these things are taken into account once the schooling coefficient is calculated.

After performing the meta-analysis and because of data availability, the authors argue that the majority of Spanish papers use a continuous measure for education (years rather than levels or degree attainment), the wage premium is larger for men than women and the coefficients obtained by using Ordinary Least Squares are smaller than the ones obtained using the Instrumental Variables estimation method.

In Raymond (2011) the returns to education are studied from 1995 to 2006 and the author shows evidence that during this period, there was a fall of approximately 2 percentage points in the returns for Spain. In addition to the change in returns, there was a decrease in the average wage of the population. This decrease was mainly due to the change in the population structure reflected on a larger proportion of women and immigrants. The Spanish labour market suffered from a demographic transformation including an increase in the female labour participation rate and in the general level of education in society (with higher levels for women).

The accumulation of human capital is relevant because it leads to economic progress. For this accumulation to happen, there must be an incentive to attain further levels of education. The accumulation of human capital will only take place if the more educated can find an optimal position in the labour market and achieve a significant wage for what they do. If technological progress is affecting the production system, a larger proportion of qualified manpower is needed.

Some influential papers on this topic from the Spanish literature are summarized. Arrazola and De Hevia (2008) use both measures of education (discrete and continuous) and control for endogeneity using instrumental variables. Mincerian wage equations are presented after performing OLS and IV, they observe a downward bias of 4 percentage points in the OLS schooling coefficient for men. Along with Arrazola and De Hevia

(2008), Jimeno and Fuente (2005) estimate private returns to education across the different regions in Spain and draw attention to the implications for the public finances. They argue that an additional year of schooling increases the probability of employment by between 0.5 and 2 percentage points. They showed that the returns to education across regions are very different.

Corugedo et al. (1992) argue the idea that signaling and human capital theory are not two exclusive theories. Once the individual joins the labour market, and the signaling mechanism (through education) has already occurred, there is a human capital accumulation in the workplace through the experience gained each additional year. After using a sample from 1988 they conclude that women benefit more from education than men and that secondary education served more as a selection mechanism. These authors believe that if the aim of education is to make people be more productive, a strengthening of the quality of education is needed.

There is a group of papers that studied the returns in the period 1980 to 1991. This period was of interest to many researchers because a change in the wage structure occurred and as a consequence of this, the returns to education changed. Important events happened during this period such as the expansion of schooling, a change in the economy from industrial based to service based, a large increase in the number of educated women and the use of education as a mechanism to become employed. One of these papers is Lassibille and Gomez (1998). They investigate if different average wages during this period can be explained by features of the working population or can be explained by distinct pay structures. Oaxaca's decomposition is applied to the average earnings differential being the first group males surveyed in 1980-1981 and the second group males surveyed in 1990-1991. The authors show evidence of variation in wages over time that depends on level of education attained.

Other papers that studied the returns to education during this period are are Hidalgo (2010) and Pastor and Serrano (2008). The former highlights some issues regarding the available data bases in that period such as only gathering information regarding employees, therefore self-employed workers were not present and only collecting information on heads of families and women were underrepresented.The returns obtained for the periods 1980/81, 1990/91, 2000/01 in the paper were 7.4, 6.4 and 5.8 percentage points respectively.

Pastor and Serrano (2008) focus on comparing the returns to education across regions and sectors of activity. Human capital endowments were very heterogeneous across provinces in Spain due to autonomy on the decisions made by each Autonomous Community. The indirect relationship between education, sector of activity, productivity and wages was studied. They show that the returns in Spain were stable across the years, around 8 percentage points since the middle of the the nineties.

Vila and Mora (1998) examine changes in the returns to education in Spain during the eighties. Marginal rates of return are obtained for 1981 and 1991. In 1981, higher education returns were higher for men in comparison to women's returns (9.2 percentage points and 5.5 percentage points respectively) while in 1991, the returns for women were higher than men's returns (11 percentage points and 8.3 percentage points).Vila and Mora (1998) estimate returns for men were higher than those estimated by Hidalgo (2010).

Oliver et al. (1999) wrote a chapter that includes a collection of results on the returns to education stemming from miscellaneous data sets. These returns range between 5 and

7 percentage points and they vary because of the use of different variables and wage measurements. A curiosity mentioned is that in 1999, papers controlled for self-selection and sector choice bias but no work was done trying to overcome the endogeneity problem of the coefficients on education using instrumental variables. The authors supported the signaling theory.

Signaling serves as a first selection mechanism for hiring. Signaling and sheepskin effects[1] together serve to understand differences in wages and education. The authors highlight again the pattern of female returns being larger than male in the majority of the studies reviewed.

García-Prieto et al. (2005) use stochastic frontiers econometric technique to estimate the returns. They argue that there is a clear positive effect of schooling on the observed wage. As García-Prieto et al. (2005), Hernández and Serrano (2013) use the Survey of Adult Skills (PIAAC) to analyze how education affects wages and the probability of having a job. The authors indicate that thanks to education, individuals could become more capable and more productive workers. Key results presented were ceteris paribus, the fact that the individual has completed the compulsory level of education increases the probability of participation in the labour market by 7 and 20 points in the case of having higher education. Returns to a year of schooling are 7.1 percentage points without controlling for reading and maths score and 6 percentage points if controlling for the mentioned variables.

Rodríguez and Albert (2004) relate the returns to education with health outcomes. They explain that education affects so many aspects such as fertility, formation of couples,

[1] Sheepskin effects refer to an increase in wages associated with the acquisition of a diploma or degree.

happiness and health. A complication of estimating the health status arises when several variables that will partially explain health are endogenous. The authors argue that higher education gives better health status.

Alba-Ramirez (1993) calculate the returns to years of education and tests whether there is presence of over-education in the Spanish Labour Market. Controlling for educational mismatch he obtains a return of 7.3%. Ramos and García-Crespo (2003) using the random effects technique show that the returns to education once controlling for educational mismatch is 7.9 . Rather than looking at years of education, Abadie (1997) use the "robust to outliers" method of quantile regression to estimate the returns to levels of education. He obtains returns of 17, 44 and 72 percentage points for primary, secondary and university levels.

In order to support the idea of using family background variables as instrumental variables and specifically parents' education, a list of papers are mentioned: Dearden et al. (1997), Uusitalo et al. (1999), Aslam et al. (2012), Trostel et al. (2002), Brunello and Miniaci (1999) and Levin and Plug (1999). Apart from the international evidence, Arrazola and De Hevia (2008) and Pons and Gonzalo (2002) are examples of papers that use Spanish data and parents' education as instrumental variables to solve the endogeneity problem.

After reviewing the empirical evidence from Spain we can conclude that there is limited availability of data sets to study the relationship between education and earnings. However, the reviewed papers offer an initial idea of what is the rate of return to education in that country.

The way this paper fills the gap in the literature is by instrumenting dummy variables representing levels of education and by trying different sets of instrumental variables for each level checking which one adjusts best to explain specific levels of education.

According to Requena and Bernardi (2005), the growth of the education system in Spain in the last forty five years has been mainly due to two factors: education reforms and an increase in education resources. Although the implementation of the reforms was gradual and not performed at an equal pace in all the regions, thanks to them eventually an expansion in the education system took place in Spain. Today, the economic system has grown and education, particularly higher education, has become more accessible to everyone, no matter their socioeconomic status. Moreover, the persistent fall of fertility rates in Spain in the last decades has reduced the number of students and therefore, as public investments in education increased over time, the expenditure in education per student in now higher than before. Some basic characteristics that define modernization of the Spanish education system during the 20th century are: (1) the almost complete disappearance of illiterate people (there are now only a few old people who do not know how to write or read); (2) the total elimination of child labour (almost all the recent generations have received education from 4 to 16 years); (3) the growing gender equality in terms of allowing women to go further in education; and (4) an easier access to higher education. A convenient way to summarize the heart of this issue is to say that 80 percent of Spaniards of all the cohorts born after 1960 have reached, at least, the first stage of secondary education.

Therefore, taking into account all the changes in the education system and the intense process of social modernization that Spain has experienced in the second half of the last century, the main hypotheses of this chapter are formulated as follows: since (a) at the macroeconomic level, the rise in the proportion of people acquiring secondary education was promoted by growing public investment in education, and because (b) at the microeconomic level, traditional Spanish mothers spent more time with their children than fathers did (which is explained in more detail in the results section), it can be expected that public expenditure in education and mother's education will be valid instruments for the attainment of the lowest levels of education among Spanish born in the second half of the 20th century. Likewise, the achievement of higher levels of education should be explained –according to the main tenets of the Human Capital Theory– by the socioeconomic status of the student's origin family Becker (2009). The variable that better describes the socioeconomic status for the generations studied is father's occupation, which will be used as an instrument for the attainment of higher levels of education. These hypotheses will be tested and discussed in the results section.

3.4 Data Set and Variables

The survey used is the Living Conditions Survey, collected by the National Institute of Statistics in Spain. The Living Conditions Survey emerged as a continuation of the European Community Household Panel (ECHP). This survey contains information about households and single individuals. Although the total number of observations (N=29,211) is small in comparison with the population of the country, the sample is designed as a nation-wide representation of all the people. Weights are applied to the data to adjust for the sample design and obtain population results in the descriptive statistics. The weight that is initially given to each observation in the sample will be modified. In order to do this, the following formula is used to cancel the expansion factor:

$$\text{Weight} = \frac{\text{Original weight}}{\frac{\text{Population aged 18-65 in 2011 (38808754)}}{\text{Survey Size (29211)}}}$$

Only wage-earners will be taken into account in this study for several reasons. First, some researchers such as Vila and Mora (1998) believe that self-employed workers do not give reliable information about their own income. Second, if the net wage per worked hour is used as a measure of income to calculate the returns to education, it will be more appropriate to study wage-earners because self-employed workers choose their own time to work so it will be more difficult (or less precise) to calculate the wage per hour for this group. Third, as we can see in the fourth column of Table 3.3, wage-earners comprise 84 percent of the whole working age population so they represent a significant percentage in comparison with the other groups such as self-employed and others (16 and less than 0.01 percent respectively).

Table 3.3 shows the distribution of the labour force in Spain by gender. We can see that the proportion of males participating in the labour force is larger than female participants (81 in comparison with 68 percent). In Spain the percentage of self-employed, of the employed population, for male and female (12 and 6 percent respectively) is very low in comparison with the high percentage of wage-earners, although superior to the northern and central european countries. Estimating the returns to education for the self-employed in Spain would not be very representative of the whole working population. To conclude the analysis of the labour force, the proportion of male wage-earners (51 percent) do not differ much to the proportion of female wage-earners (46 percent).

Table 3.4 contains the average net monthly earnings of wage-earners by gender and maximum education level attained. At all education levels, male earnings are significantly greater than female earnings. The gender gaps in earnings are not very different across education levels. The highest gap between males and females occurs in the secondary second stage with a difference in earnings of 424.3 euros in 2011. In relative terms, on average male earnings are 23 per cent higher than female earnings.

In many Spanish studies only male workers are considered to calculate the returns to education. This is to avoid the sample selection problems that female individuals can cause in the estimation. These problems arise from the fact that female careers are affected and sometimes stopped by caring duties or birth of children. Therefore, regarding the estimation, these differences could cause sample selection issues. In this study female participants are also taken into account because there seem to be very similar numbers of males and females in this sample (out of all the wage-earners in the sample we have around 53 percent males and 47 percent females). It could be argued that it is necessary to control for male selection issues too.

The large number of variables are classified into six different categories: sociodemographic, education, labour, health, income, inter-generational transmission of poverty. Information is available from the year 2004 to 2013. Each year, the survey contains a special subsection with new variables added. The year 2011 is used because the subsection called "intergenerational transmission of poverty" provides us with variables about the socioeconomic background of each individual. This special subsection focuses on individuals that are between 25 and 59 years old. The questions asked refer to the period when the individual was a teenager (14 years old). When family background variables are included in the regressions the analysis is restricted to individuals aged 25 - 59 years while if none of these variables are used, the estimation sample consists of individuals older than sixteen years and younger than 65 (age of retirement).

In order to derive a continuous variable for education, we follow the common approach:

YEARS OF EDUCATION = year that highest level of education was attained

– year of birth – 6

Children in Spain start primary school at the age of 6. If after creating this variable following the formula above we look at the summary statistics we find that the minimum value is 0 years and maximum 52. Obviously this procedure is not taking into account if the individuals have stopped their education and years after go back, or if they resit some years for example. In order to avoid these discontinuities, a procedure is followed. In this database we also have a variable that reports the highest level of education attained that takes values: 0 no education at all, 1 if primary education, 2 secondary first stage, 3 secondary second stage and 4 higher education. The average[2] years of education for those who have less than 22 years of schooling (6 years of primary education, 4 secondary education, 2 years of pre-university education, 4 years of degree in university, 2 of Masters and 3 of PhD) is calculated. These values are tabulated with levels in the table below. Now, putting together all the individuals (the ones who have less than 22 years of education (21,20, etc.) and the ones with more than or equal to 22 years of education) we truncate the cases where the individuals have more than 21 years and we

[2] The median years of education, more robust than the mean, has been tried but does not alter the estimation results.

replace the number of years with the one from the table below according to the education level that they report. So, for example, for an individual that has 25 years of education and reported a secondary second stage level we assign them the value 12.33 rather than 25. The continuous variable for education is called Years of Education.

Level of Education attained	**Mean of Years of Schooling**
No Education	0
Primary	7.62
Secondary First Stage	9.50
Secondary Second Stage	12.33
Higher Education	16.41

The variables used for the levels of education attained are:

- No education: the individual has not attained any level of education at all.
- Primary: the individual's highest education attainment is primary education.
- Secondary First Stage: the individual's highest educational attainment is secondary first stage education or equivalent.
- Secondary Second Stage: the individual's highest educational attainment is secondary second stage education or equivalent.
- Higher Education: the individual's highest educational attainment is an undergraduate degree, postgraduate, etc. or equivalent.

In order to make the analysis simpler, the variables that represent the region where the individual is living in 2011 are grouped into categories:

- North West: Galicia, Principado de Asturias and Cantabria.
- North: País Vasco, Comunidad Foral de Navarra, La Rioja and Aragón.

- East: Cataluña, Comunidad Valenciana and Baleares' Islands.

- South: Andalucía, Región de Murcia, Extremadura, Ciudad Autónoma de Ceuta, Ciudad Autónoma de Melilla and Canarias.

- Inner: Castilla-La Mancha and Castilla y León.

- Madrid: the capital of the country.

The following listed variables are used to estimate the labour market participation equation and the wage equations:

- Employee: This is a dummy variable that takes the value of 1 if the individual was employed (wage-earner) in the month before the interview and 0 if the individual was in one of these situations: unemployed, student, schoolchild or in training, permanently disabled and unable to work, working in the home, looking after children or other persons or another type of economic inactivity (not receiving a wage).

- Experience: number of past years in paid work[3].

- Experience^2: experience squared (years squared).

- Income rest household: annual gross income of the rest of the household, measured in Euros.

[3] As a robustness check, results were also obtained for potential experience defined as (age year when achieved the highest education + year of birth) for people with some education and defined as (age - 16) if "No Education" was attained. In the wage equations, the returns to education estimated with potential experience were slightly higher, but the interpretation of models is the same than the one presented in the text.

- Dependent children: dummy variable that takes the value of 1 if there is one or more economic dependent children in the household, 0 otherwise.

- Married: dummy variable that takes the value of 1 if the individual is married and 0 otherwise.

- ln(wage): the logarithm of net wage per hour. The wage is measured this way because the distribution of log(hourly earnings) is close to a normal distribution. Heckman and Polachek (1974) tried several transformations of earnings and this one was the one finally chosen. Also it is convenient when it comes to interpreting the results.

- Male: dummy variable that takes the value of 1 if the individual is male, 0 if female.

And the last set of variables are the instruments for schooling (all these variables give information from the period when the respondent of the survey was a teenager):

- Father's Higher Education: dummy that takes the value of 1 if the father's highest level of education is higher education, 0 otherwise[4].

- Mother's education: The mother could have attained one of these levels:

  -No education: the dummy takes the value of 1 if the mother is illiterate, 0 otherwise.

  -Secondary First Stage: the dummy takes the value of 1 if the mother's highest educational attainment is secondary education, basic education or a level below this.

[4] For the sake of parsimony, only results using father's higher education are presented in the analysis. Results including the rest of father's education categories did not pass the econometric test and this is the main reason why they are not used.

-Secondary Second Stage: the dummy takes the value of 1 if the mother's highest educational attainment is post-secondary education.

-Higher Education: the dummy takes the value of 1 if the mother's highest educational attainment is higher education.

- Father's Occupation: a recode of the OECD Occupational Classification ISCO-88 has been used grouping the occupation into three groups. The dummies take the value of 1 if the father had one of the following occupations, 0 otherwise.

  -Managers & Professionals: legislators, senior officials, managers, professional, technician, associate professionals.

  -Skilled Worker: clerk, service worker and shop and market sales worker, skilled agricultural and fishery workers, craft and related workers.

  -Unskilled worker: plant and machine operators and assemblers (elementary occupations).

- Public Expenditure. This is a macro economic series of the total public expenditure on education in Spain in the year when the individual was 12 years old. This variable is measured in millions of pesetas (currency used in Spain before the Euro). Data are recorded on this variable from the year 1944 to 1994.

- No Financial Difficulty. Economic difficulty for the family of making ends meet when the adult was a teenager. Takes the value of 1 if not difficult and 0 if difficult.

- Financial Situation. Economic household situation when the adult was a teenager. Takes the value of 1 if good financial situation and 0 if not good[5].

[5] Financial Difficulty and Financial Situation are two different variables. An illustrative example is that a family with high income and therefore with a good financial situation could be in a financial difficulty because of not being organised with the household expenses.

3.5 Methodology

3.5.1 Sample Selectivity Model

There is only a wage variable available for those who receive a salary. This leads to a sample selection of individuals. If we obtain the estimates of the returns without taking this into account, any variable that influences the fact of being a wage-earner (and therefore belonging to the sample) is potentially correlated with the error term in the wage equation. If this is not taken into account the coefficients will be biased. Being a wage-earner is not a random feature of the population so must be taken into account. In order to proceed, two equations are estimated.

Participation Equation

The equation that indicates the propensity of belonging to the sample:

$$Z_i^* = \psi' W_i + u_i \tag{3.2}$$

where u_i follows a bivariate normal distribution with mean equal to 0, W_i is a vector of observed characteristics that influence the probability of being in the paid labour force, and Z_i^* is a latent variable. The observed variable, Z_i, is defined as follows:

$$Z_i \begin{cases} 1 & \text{“wage is observed” if } Z_i^* > 0 \\ 0 & \text{“wage not observed” if } Z_i^* \leq 0 \end{cases} \tag{3.3}$$

$\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and distribution functions respectively. The probabilities of participating and not participating in the paid labour force are

$$\text{Prob}(Z_i = 1|W_i) = \Phi(\psi' W_i),$$

$$\text{Prob}(Z_i = 0|W_i) = 1 - \Phi(\psi' W_i)$$

respectively.

<u>Main equation</u>

The earnings equation for those who receive a salary is:

$$\ln(Y_i) = \beta' X_i + \varepsilon_i \tag{3.4}$$

where Y_i is the wage and is only observed if $Z_i = 1$. X_i is a vector of observed variables that affect earnings such as education, experience, squared experience, etc. β is a vector of parameters to be estimated. ε_i is a normal random error term disturbance with 0 mean and constant variance σ_ε^2. This random disturbance encompasses all the variables that affect wages and are not included in the main equation. The selection model assumes that ρ, the correlation between the two error terms, ε_i and u_i is not zero.

So, to recap, W_i and Z_i are obtained for all the individuals (those who have a wage and those who do not) while Y_i is only observed when $Z_i = 1$. Following Greene (2003) it is then possible to derive the following expectation to obtain the model that applies to the observations in our sample:

$$E[Y_i|Y_i\text{observed}] = \beta' X_i + \rho \sigma_\varepsilon \lambda_i \tag{3.5}$$

This expectation is obtained using the theorem about the Moments of the Incidentally Truncated Bivariate Normal Distribution; σ_ε is the standard deviation of the random wage disturbance and λ_i is equal to:

$$\lambda_i = \frac{\phi(\psi' W_i)}{\Phi(\psi' W_i)} \tag{3.6}$$

λ is also called the Inverse Mills ratio or selection hazard. This ratio should be included in the main regression to obtain unbiased coefficients. As Greene (2003) says least squares regression of Y on X and λ would be a consistent estimator of the coefficient on X, but if λ is omitted, then the specification error of an omitted variable occurs.

Now in order to estimate the missing parameters, we use the Heckman (1979) method. This method comprises two steps:

1. A probit estimation is performed by maximum likelihood in order to obtain an estimate of ψ. Also, the $\hat{\lambda}$ value needs to be calculated for each observation in the selected sample:

$$\hat{\lambda}_i = \frac{\phi(\hat{\psi}' W_i)}{\Phi(\hat{\psi}' W_i)} \tag{3.7}$$

2. Estimate β and the product $\rho\sigma_\varepsilon$ by OLS regression of $\ln(Y)$ on X and $\hat{\lambda}$. Following this procedure, we will obtain consistent coefficients in the wage equation and therefore a consistent return to schooling will be obtained as well.

A last thing to mention regarding this method is that at least one variable that appears in the selection equation must not be included in the wage equation as an explanatory variable. This means that at least one of the variables affects participation but has no effect on the wage in this case. In this study the variables not included are: the income of the rest of the household and the presence of an economic dependent child in the household. If the exclusion restriction variables[6] are not included, collinearity problems are very likely.

3.5.2 Endogeneity, causes and explanation

This part is devoted to explaining the endogeneity problem that could be faced when estimating the returns to education. The model to be estimated takes the form:

$$\ln(Y_i) = \beta_0 + \beta_i X_i + \varepsilon_i \tag{3.8}$$

One of the Gauss-Markov assumptions is that the expected value of the error term conditioned on observed variables equals 0, so $E[\varepsilon_i|X_i] = 0$. There are three different scenarios where this assumption does not hold when analyzing cross-section data:

1. Omitted Variables. If an omitted variable is correlated with variables contained in the vector of independent variables or regressors, this causes a violation of the Gauss-Markov assumption and the expected value of the error term would not equal 0, so $E[\varepsilon_i|X_i] \neq 0$ and there would be an endogeneity problem. In the case of studying the returns to education, a classic omitted variable is ability.

2. Measurement error in the independent variables.

3. Reverse causality. If the primary process of interest is how X_i affects Y_i but there might be also a way in which Y_i causes X_i.

[6] The use of these variables is convincingly justified in Arrazola and De Hevia (2006).

The problems or consequences of this violation are briefly explained as follows. The most important problem is that the expected value of the coefficient $\hat{\beta}$ no longer equals the true parameter β, therefore $E[\hat{\beta}_{OLS}] \neq \beta$. In other words, OLS is biased. Furthermore it is not only biased; as the sample size tends to become larger, $N \to \infty$, then the estimated parameter does not tend to the true parameter, $\hat{\beta}_{OLS} \nrightarrow \beta$, so the estimator is also inconsistent. Therefore the estimates obtained using OLS would be incorrect because of these issues. They will not be centered around the true population value and, as the sample size increases towards the population size, there is no convergence to the true parameter.

The process of interest is how X_i affects Y_i. As X_i increases, if $\beta > 0$ we tend to see increases in the dependent variable Y_i as well. The problem is that, under endogeneity, if X_i increases, this also will have an effect on the error term ε_i. In the case of endogeneity caused by **omitted variable**, if this variable is positively correlated with the independent variable X_i, increases in X_i will cause increases in the omitted factor, which will cause some increase in Y_i. At the end then, changes in Y_i are not just due to changes in X_i. There is also a component of Y solely due to changes in the omitted factor. Changes in Y_i could be decomposed into two effects. The problem with endogeneity is that X_i cannot just be changed without changing some of the omitted factors or without causing some change in the error term. Instrumental Variable (IV) estimation allows us to untangle X_i from the error term and gives at least consistent estimates of β.

If the cause of the correlation between explanatory variable and disturbance term is **measurement error** in the independent variable, then again $E[\hat{\beta}_{OLS}] \neq \beta$ and therefore $\hat{\beta}_{OLS}$ will be biased. In the attempt to estimate the returns to education, the problem arises when the schooling variable (and independent variable in the wage equation) is not measured correctly. In order to explain this, we need to make a distinction between two variables: actual measured schooling MS_i and true schooling S_i. These two variables are related using the following equation:

$$MS_i = S_i + v_i \tag{3.9}$$

where v_i is an error term. The correlation between schooling and the wage variable is weaker now. If there is a large measurement error and wages are regressed on schooling, it might be the case that OLS would not pick any effect at all so $\hat{\beta}_{OLS} \to 0$.

To recap, if either of these things happen $E[\varepsilon_i|X_i] \neq 0$ or $cov(\varepsilon_i, X_i) \neq 0$, then this leads to bias in the OLS coefficients. Now the direction of the bias caused by measurement error is explained. In order to do this, we need to solve for the true schooling in equation 3.9 and substitute schooling in the wage equation:

$$\begin{aligned} \ln(W_i) &= \beta_0 + \beta_1 S_i + \varepsilon_i \\ &= \beta_0 + \beta_1 (MS_i - v_i) + \varepsilon_i \\ &= \beta_0 + \beta_1 MS_i + (\varepsilon_i - \beta_1 v_i) \end{aligned} \tag{3.10}$$

where $(\varepsilon_i - \beta_1 v_i)$ will be called the composite error term. What is of interest here is the covariance between the independent variable and the composite error term in the last equation, $cov(\varepsilon_i - \beta_1 v_i, MS_i)$. If this covariance does not equal 0, then this cause a violation of the Gauss-Markov assumption of no endogeneity and OLS is biased. Using properties of the covariance, the following approximation could be made:

$$\begin{aligned} cov(\varepsilon_i - \beta_1 v_i, MS_i) &\simeq -\beta_1 cov(v_i, S_i + v_i) \\ &= -\beta_1 cov(v_i, v_i) \\ &= -\beta_1 \sigma_{v_i}^2 \end{aligned}$$

So as said before, if this last term ($-\beta_1 \sigma_{v_i}^2$) does not equal 0, endogeneity is present in the model and OLS coefficients are biased. The direction of the bias is given by $-\beta_1$ because the variance is always positive ($\sigma_{v_i}^2 > 0$). To conclude this explanation, when there is measurement error, the coefficients are biased downwards and the degree of the bias is going to be increasing in the level of variance of the error term v_i.

3.5.3 Instrumenting Schooling

In order to obtain consistent estimates of β_i in the wage equation when S and ε are correlated instrumental variables are needed in the model. The schooling equation is

$$S_i = \alpha_0 + \alpha_1 E_i + \alpha_2 E_i^2 + \alpha_3 IV_1 + \alpha_4 IV_2 + u_i$$

where: $E(u) = 0$, $Cov(E,u) = 0$, $Cov(E_i^2, u) = 0$ and $Cov(IV_k, u) = 0$ for $k = 1,2$. To identify the model, the two coefficients, α_3 and α_4, have to be non zero.

Going back to the wage equation,

$$\ln(Y_i) = \beta_0 + \beta_1 S_i + \beta_2 E_i + \beta_3 E_i^2 + \varepsilon_i$$

we have the explanatory variables S, E and E_i^2 and the error term ε. It is assumed that: $E(\varepsilon) = 0$, the experience and its square are exogenous (not correlated with ε) but what is not sure is if S is also not correlated with the error term. It could be the case that such correlation exists and therefore $Cov(S,\varepsilon) \neq 0$.

Another test that is needed to be performed is the endogeneity test or Hausman test. To carry out this test the following steps are followed:

1. The schooling equation (or reduced form) is estimated.
2. The residuals from the schooling equation are saved, $\hat{u}$;
3. The wage equation is estimated now but including these residuals as another regressor:
$$\ln(Y_i) = \beta_0 + \beta_1 S_i + \beta_2 E_i + \beta_3 E_i^2 + \delta \hat{u} + \varepsilon_i$$
4. The last step is to test the null hypothesis $H_0 : \delta = 0$. If the null hypothesis is rejected, the schooling variable is endogenous and therefore the 2SLS method is to be used rather than OLS.

There are two possible scenarios: the schooling variable being exogenous, in which case OLS estimation will give consistent estimates and the schooling variable being endogenous, in which case OLS will give wrong estimates. In the second scenario, 2SLS is needed to overcome the endogeneity problem. This procedure comprises two steps:

1. In the first step, instrumental variables are used in order to predict the amount of schooling that each individual has accumulated. In order to explain the method, assume that two instrumental variables are used. The following equation is

estimated:

$$S_i = \alpha_0 + \alpha_1 E_i + \alpha_2 E_i^2 + \alpha 3 IV_1 + \alpha_4 IV_2 + u_i \tag{3.11}$$

Once equation 3.11 is estimated (called the reduced form), the predicted values $\hat{S}_i$ are saved.

2. In the second step, the wage $\ln(Y_i)$, is regressed on E_i, E_i^2 and $\hat{S}_i$ using OLS. If this procedure is followed, the coefficient on schooling would be a consistent estimate of the returns to education if the instruments are valid.

The most difficult task in the procedure is to find good instrumental variables. Each instrumental variable has to meet two criteria: in order to be valid instruments, they must not be correlated with the error term of the wage equation (ε) and they must be highly correlated with the endogenous variable, which in this case is schooling. In this study, the used instrumental variables are the following: the level of education of the parents, the public education expenditure in Spain and proxies of the economic situation of the household when the individual was a teenager. Depending on the form (discrete or continuous) of the endogenous variable different set of instruments will be used.

3.5.4 Testing the Instrumental Variables

If more than one instrument is used in the model, and therefore we have more variables to identify the model than needed, then an over-identification test can be performed. It is very important to point out that this test assumes that at least one instrument is valid. What is needed to perform this test is:

1. The wage equation is estimated by 2SLS and the residuals are saved ($\hat{\varepsilon}$).

2. Then a regression of $\hat{\varepsilon}$ is made on the rest of exogenous variables (the ones from the wage equation) plus the instruments used; in the example from before the variables E_i, E_i^2, IV_1 and IV_2 are used. Then the R^2 from this last regression is obtained.

3. Under the null hypothesis that all the instrumental variables are uncorrelated with the error term, nR^2 follows a Chi-Squared distribution with q degrees of freedom (χ_q^2) where q is the number of instrumental variables in the model minus the number of endogenous variables and n the number of observations. If the test statistic does not exceed the χ_q^2 critical value then we do not reject the null hypothesis and there is confidence that the set of instruments used pass this robustness test.

In order to test if the instruments are highly correlated with the endogenous variable, we need to test the joint significance of the coefficients on the instruments in the reduced form; in this case a rule of thumb can be applied Staiger and Stock (1997): if an F statistic smaller than 10 is obtained this means that the set of instruments used are potentially weak. If F is larger than 10, it can be assumed that education (S) is partially correlated with IV_1 and IV_2. Weak instruments could induce bias in the coefficients and size distortions in the hypothesis tests.

Stock and Yogo (2005) formalised the Staiger and Stock (1997) procedure of the rule of thumb. Based on the Cragg-Donald Wald F test, Stock and Yogo (2005) developed two precise and quantitative definitions of weak instruments: 1) a group of instruments is weak if the IV estimator bias, relative to the OLS bias exceeds a particular threshold, b (for example 10%); 2) instruments are weak if the Wald test based on IV statistics has an actual size that could exceed a certain threshold r, for example $r = 10\%$ when $\alpha = 5\%$.

A last robustness check is to estimate the model using Fuller-Limited Information

Maximum Likelihood (Fuller- LIML) instead of 2SLS. Fuller-LIML estimation is more robust than 2SLS. What is needed to be controlled is that the estimates on the returns to education once Fuller-LIML is used do not change too much in comparison with the ones obtained by 2SLS.

As Davidson (1993) explain, Fuller proposed a modified LIML estimator where a is a positive constant. In comparison to the LIML estimator, that has no finite moments, Fuller's modified estimator has all the moments finite provided the number of observations is large enough. Also the F test of the first stage (the schooling equation) must be checked. As Dickson (2009) says "the modified LIML estimator introduced by Fuller, with the Fuller parameter (a) set to one is regarded as most robust to any potential weakness of the instrument".

Stock et al. (2002) say, when instruments are used in the analysis and the errors are normally distributed, the Fuller-k estimator with the parameter a set equal to one is the best unbiased to second order. This means that among all the second-order bias-corrected estimators, it has the smallest mean squared error. When there is presence of weak instruments, LIML is a more robust estimation method in comparison with 2SLS. One of the reasons why LIML is superior is because of the thresholds needed to know if the instruments are weak or not. In LILM the thresholds do not increase depending on the number of instrumental variables used while if 2SLS is used as the estimation method, the thresholds increase with the addition of more instrumental variables. As mentioned before, even if the F statistic is larger than ten, it is cautious to check the coefficients using LIML just in case they change dramatically.

3.6 Results

3.6.1 Mincerian Earnings Equations

OLS estimations are reported on Table 3.5. Following Blackburn and Neumark (1995), industry and occupation controls are omitted for now because part of the returns to education can be captured by the variation in industry or occupation of employment.

As we can see, columns (a) and (c) are for the returns using the continuous measure of schooling, years, while columns (b) and (d) are for the returns using the discrete level of education attained. The coefficient on years of schooling gives us the estimated return to an additional year of education. The coefficients on the education levels gives us the return to the difference between having primary education in comparison with higher levels of education or no education at all. The drawback of using years rather than levels is that each additional year is assumed to have the same effect; one year of primary education will account for the same as one year of Masters studies. Comparing the rate of return among genders, males have lower returns to education than females. One more year of education will increase males' wage by 5 percent and 6 percent for females. These results imply that there are gender difference in the returns to education in Spain. In the levels specification the reverse happens. Men have higher returns to secondary second stage education (28 percent for males and 24 percent for females[7]) and higher education (69 percent for males and 82 percent for females). For females, the coefficients on no education and secondary first stage are not statistically significant. Having attained no

[7] These values are calculated using the inverse function of the natural logarithm from Table 3.5, $e^{0.246}$ and $e^{0.222}$, respectively.

education or secondary education does not make a difference relative to females with primary education. For males we observe that the coefficients on levels rise with the higher levels attained. Finally, there is more of an increase in the premium for females between the levels secondary second stage and higher education, 38 percent, while for males there is an increase in the premium of 28 percent.

3.6.2 Probit model: labour market participation

Table 3.6 presents the results for the binary probit estimation where the dependent variable is a dummy for Employee. As can be seen, there are some disparities in how the different variables affect the probability of being an employee among genders, although the effect of education is very similar for both genders. All the education levels have coefficients that are individually statistically significant. This reflects the fact that education increases the probability of being employed.To interpret the results, the marginal effects are used.

Marriage affects in different ways the probability of receiving a wage among genders. For men, being married has a positive effect on the probability while for a women, it has a negative effect. Being married increases the probability of participation by 8 percentage points for males while for females this probability is reduced by 12 percentage points. This difference may be due to financial responsibility; married males feel more responsibility while the reverse happens for females according to Aslam et al. (2012). Ferrada and Zarzosa (2010) estimate female labour participation in Chile and obtain for different regions always a negative coefficient for the marriage variable. If the husband has a high level of education, higher earnings for the household are expected and less need of an ‘extra’ income in the household; this would generate a lower female participation rate for married women. It was common that a woman quits her job once she gets married or during the first years of marriage. The fact that she quits her job could be determined by her age or the presence of children in the household. There is a stream of researchers that argue that the husband's characteristics are not useful to explain the probability of a woman quiting her job after being married although, the variables related to the maternity are important to determine this probability. Blau and Kahn (2000) give an explanation of the signs of the coefficients on the marriage variable saying that in patriarchal family structures, women are expected to turn into good mothers and homemakers while men are expected to be breadwinners and the head of the household. Now in Spain this is changing because Spanish women are participating more and more in the labour market, but for the sample studied here, Blau and Kahn (2000)'s argument applies to some extent. As seen in Table 3.6 marital status makes a difference in being in the labour market or not for both genders.

As expected, the significant coefficient on income of the rest of the household has a negative effect for both genders. This is because if there are more sources of income in the household, and these provide enough income for the family to live, there is no need for the individuals to work and therefore to contribute with an additional wage. In agreement with popular expectation, the number of dependent children has a positive and significant impact on the probability of being employed. This occurs for both genders. The effect is more attenuated in males than in females. Having at least one economically dependent child, raises the probability of participating by 2.8 percentage points while for females this probability rises by 3.5 percentage points. This fact also can be explained with the financial responsibility; having a child entails expenses in education, clothes, etc. In order to afford these additional expenses, there is need for a source of income.

Regarding the statistical significance of the regional coefficients, they differ by gender. The reference category is Madrid, the capital of Spain. For males the only statistically significant zone that affects the probability of employment is the south of Spain. In comparison with Madrid, living in the south decreases the probability of being in paid work by 8 percentage points for males. For females, the same effect occurs but this time with a larger impact of 9.6 percentage points. In the female case, more zones compared to Madrid alter the probability of being in paid work. Even though there are disparities in the individual statistically significance of the single coefficients for the zones, the Chi-Squared test on the joint significance of all the regions is performed. We can see that these variables, altogether, make an impact on the probability of receiving a wage for males and females (they are jointly statistically significant).

3.6.3 Two-Step Heckman model

As explained in the methodology section, a possible bias could arise due to self-selection and the OLS coefficients will be biased. The results contained in Table 3.7 are corrected for sample selection using the two-step Heckman Selection Model[8]. Lambda (the inverse Mills ratio) is included as a regressor in the wage equation. The rest of the regressors in the earnings function are education (measured as a continuous variable in columns (a) and (c) using years, and as a discrete variable in columns (b) and (d) using levels of education attained), experience, squared experience, a dummy for being married and the grouped regions dummies. As exclusion restrictions the following variables are used:

[8] To sidestep the labour market participation issue, results have also been obtained for only prime-aged men (aged 24-55). The results are robust to looking at just prime aged men: the coefficients on the returns to education for these models do not vary substantially in comparison to the ones presented in this thesis.

income of the rest of the household and presence of dependent children. These variables are expected to affect the probability of being an employee but not directly affect earnings as Puhani (2000) points out.

As we can see, the coefficient on Lambda is statistically significant in the four models and has the same sign (positive) for all the cases. This can be interpreted as follows: wage earners for whom we observe a salary have higher salaries than a randomly selected individual. It can be seen that the coefficients on education (both specifications, years and levels) do change between to Table 3.5, OLS estimates without taking into account the sample selection, and Table 3.7 controlling for sample selection. Including Lambda, increases all the education coefficients (education levels and years) in comparison with the OLS estimates. The coefficient (years) for males increases from 5 percent to 7 percent and for females this increase is more accentuated, from 5 to 8 percent. Also there is a slight increase in the coefficients for some of the education levels in Table 3.7 for males and females. In the levels specification, for females, the OLS coefficient on the first education levels, no education and secondary first stage are not statistically significant, while in the Heckman model, the secondary first stage level becomes individually statistically significant. Note that there is a large difference for females in the case of the education levels, column (d) of Table 3.7, for secondary second stage and higher education if we compare with the reference category, women with primary education. While these coefficients in the OLS are 24 and 82 percent respectively, in the Heckman model they are 57 and 162 percent (column (d)). In general, the results suggest that OLS underestimates the returns to schooling when not allowing for the selection into employment.

Whether the returns are estimated using the OLS or Heckman approach, the coefficients on the returns to education remain larger for females than males (using years rather than levels). After controlling for the sample selection bias, women have a return of 8.5 percent for each year of education that they complete while men have a lower return, 7 percent (columns (a) and (c)). The difference in these coefficients is 1.5 percentage points using the Heckman model while using OLS this difference is 1 percentage points. The experience variables have the expected signs for both genders: positive for experience and negative for squared experience. Experience is increasing albeit but at a diminishing rate. In terms of regions, as expected, wages are higher in the capital Madrid, because all the coefficients on regions that are significant are negative except the North one for males and females in Table 3.7. This can be explained because the cost of living in the capital is more expensive so people do need to have higher wages to face this. This increase difference also probably is due to different occupations. Wages on average may be lower in Madrid than in the north of Spain because of a large presence of a powerful industrial sector there. In fact several provinces in Spain with the highest income per capita are located in the north of the country.

3.6.4 Instrumenting Years of Schooling

As explained before, if the schooling variable is correlated with the error term, OLS estimates will be biased. Measurement error in the schooling variable may also bias the estimates downwards. Instrumental Variable estimation is a technique that controls for the possible bias due to omitted variables (such as ability) or measurement error. From this point onward we do not control for self-selection because the main focus is to correct for the mentioned biases[9].

Years of education are instrumented in Table 3.8 (Males) and Table 3.9 (Females). Three specifications for each gender are estimated: using father's education , mother's education and both instruments at the same time. Coefficients of the two stages of 2SLS are reported. The first stage is reported in columns (a), (c) and (e) while the earnings equation (second stage) is reported in columns (b), (d) and (f). When family background variables such as parents' education are used, only individuals aged 25-59 are taken into account (sample restrictions). Family background questions are only asked to individuals aged 25-59. Unfortunately, this subset of the population is not a random draw and this could lead to sample selection that cannot be controlled for.

All the instruments have the expected sign and are individually statistically significant. Parental education affects always positively own educational attainment. By observing the first stage coefficients we can say that the instruments are relevant.

The results suggest that the OLS coefficient on education is biased downwards. The change in the coefficients using OLS and 2SLS is for men 5.4 - 7.1 percent and for women 8.5 - 9.6 percent. We can see in Table 3.8, for males, father's education has a larger impact in comparison with mother's education. This also happens for females if both instruments are included at the same time, column(e) of Table 3.9.

In order to see if the instruments are correlated with schooling a joint significance F test is performed for each of the first stages. All the F tests statistics (Table 3.8 and Table 3.9)

---

[9] It is assumed that the self-selection bias when using education as a categorical variable and as a continuous variable is the same; therefore the estimates obtained not controlling for this bias in both cases are comparable.

are above 10, so there is evidence that the instrumental variables used are relevant instruments and not potentially weak. It can be seen that if we compare the C-D Wald F test to the critical values tabulated by Stock and Yogo (2005) the null hypothesis of the maximum relative bias being at least 10% can be rejected for the models presented, males (Table 3.8) and females (Table 3.9). These instrumental variables are able to identify the wage equation. When using both instruments, regarding the p-values, 0.895 and 0.353, from the overidentification tests, the null hypothesis is not rejected. So we can confirm that when both instruments (mother's education and father's education) are used together the overidentifying restrictions are valid. The smallest partial R^2 of the effect of the instrument on years-of-schooling having partialled out the effect of the other covariates is 0.026 which is high relative to the guidelines given by Bound et al. (1995). The p-values associated with the endogeneity test are all less than 5 percent so it is suggested that is better to use 2SLS as an estimation method rather than OLS. It is believed that education is endogenous rather than exogenous.

Wage equations are obtained using Fuller-LIML to check the robustness. For males (Table 3.10) and using father's education as an instrumental variable the coefficient on years of schooling has the same size using both estimation methods (Fuller-LIML and 2SLS). The same happens with females (Table 3.11). When using the Fuller estimation method the coefficients on years of education also still remain statistically significant. As explained in the methodology section, estimating the regressions by using Fuller-LIML and checking that the coefficients on schooling do not vary is a robustness check.

Borrowing Murray (2006)'s words regarding instrumental variables: we need to avoid "bad, weak and ugly instruments". 'Bad' instruments are those correlated with the disturbances. 'Weak' instruments, are those little correlated with the troublesome explanator (in our case schooling). 'Ugly' instruments are those that yield results uninformative about what we are interested in.

It seems that parental education is a commonly used instrument for schooling for the Spanish case. Even if father's education, mother's education or both instruments are used, the coefficients on the returns between the three specifications do not vary much (7.2 percent for males and 9.6 percent for females from Table 3.8 and Table 3.9).

Those individuals that have more ability are the ones who tend to acquire more education. Due to the non-observed ability, the ability is gathered by the error term in the regression. This will cause the error term to be correlated with the education variable. Basically, if more educated parents pass on to the children the tendency of acquiring more education but do not pass innate ability, parents education will likely be a valid instrument.

There are researchers who do not support parents' education as an instrumental variable for education. They argue that parental education should be included as a control in the wage equation rather that using it as an instrumental variable. Barceinas (2001) say that it could be argued that if a family is in a favourable economic condition, this will affect in a positive way the level of education of the children, but this also will affect in a direct way the future wage of the children so this will make the instruments invalid. Basically the idea is that parents with higher levels of education and, as a consequence, a higher level of income have a wide social network that enables the future work conditions of the children. Several papers still use family background variables as instruments for education, however.

Hoogerheide et al. (2012) use data from the 2004 German Socio-Economic Panel and Bayesian analysis to provide confidence in the use of family background variables as instruments in income regressions. They say that as however, if one's father's education affects these circumstances and attitudes, then it is implausible that one's own education would have no (or a smaller) effect. In other words, if education has a causal effect on earnings, then it is implausible that an additional year of education will benefit one's son or daughter, but not (or to a lesser extent) oneself.

3.7 Conclusion

Several models are presented in this chapter: Mincerian Earnings Equations to estimate the returns to education, a probit model to study the participation in the labour market for both genders, a two-step Heckman selection model for both genders and models instrumenting schooling (separated specifications depending if the schooling variable is discrete or continuous).

Regarding the Mincerian Earnings Equations, men have lower returns than women. Their return to an extra year of education is around 1 percentage point lower than for females. However, the reverse happens when the coefficient on returns to an extra level of education are calculated. In the latter regression, males have higher returns than females.

When a probit model is estimated for the labour market participation, education has a statistically significant impact on the participation. More educated people (males and females) are more likely to be employed than those with lower education. As the coefficients show, more educated individuals are more likely to receive a wage. Married men are more likely to work while the reverse happens for females. As expected, the higher the rest of the income of the household is, the less likely the individuals are to participate in the labour market, while if there is at least one dependent child in the household, they are more likely to be in a paid job.

Using Heckman's selection model and including the inverse Mills ratio in the wage equations indicates that a Spanish individual with sample average characteristics who is selected into a paid job obtains a higher hourly wage than an individual selected by random from the whole population. Heckman's model suggest that the coefficients obtained by OLS on the returns to education are biased downwards.

Returns to an extra year of education are calculated for individuals aged 25-59. 2SLS approch is used as the estimation method. Parental education is used as an instrumental variable. Mother's education has a larger effect on the education of the offspring in comparison with the father's education. Again, the OLS education coefficients are biased downwards suggesting that there is a need to control for the measurement error. Measurement error would appear to outweigh any ability bias, which would be expected to bias the estimated returns upwards.

Returns to education in Spain have been estimated. It is believed that parents' education is a good instrumental variable to solve the endogeneity problem that arises if OLS estimation method is used to obtain the estimate of the returns. Using the Spanish sample, OLS estimates are biased downwards. If years of schooling are used as the schooling variable, female returns to education are always higher.

The preferred instruments for the levels specifications are: mother's education and public expenditure for the lower levels and mother's higher education and father's occupation for the middle and highest levels. The reason why they are preferred is because they pass all the validity test and they are good proxies for the socioeconomic status of the family when the individual was a teenager. The socioeconomic status is the main predictor of the decision of attaining further levels of education.

Although there are alternative approaches to deal with the identification issues such as Propensity Score Matching, bounding, etc., the instrumental variables econometric technique was employed to correct for the possible biases in the coefficients. This technique was applied to categorical variables, i.e. levels of education, making a methodological contribution to the extant literature. In order to conclude, two main points should be remarked: the importance of considering various biases in the coefficients when estimating the returns to education as well as an appropriate election of instrumental variables for the levels of education to correct for these biases.

Appendix 1: Instrumenting Levels of Education

In this appendix the regression models are separated depending on levels of education attained. The instrumented education variable in this case is discrete.

First of all, three dummies will be created. The variable level of education can take one of these: Primary Education, Secondary First Stage, Secondary Second Stage and Higher Education. Each dummy will take a value of 1 if the individual has the higher level of education between two consecutive levels and then 0 if the individual has the level below. As can be seen, there are three dummies D_1, D_2 and D_3. The first one, D_1, will capture the individuals that have primary education and those who have attained secondary first stage; the second one, D_2, will capture the individuals that have secondary first stage and secondary second stage; the third one, D_3, will capture those who have attained secondary second stage and have attained higher education. A dummy with primary education versus no education is not created because there are not enough observations in the sample that have no education. In the whole sample of wage earners there are only 68 individuals that do not have education at all.

The criteria used to report the final set of instruments is:

- They have to be relevant (individually statistically significant) for the instrumented education level on its own.
- They have to be relevant for the instrumented education level when using both instruments.
- They have to pass the validity tests (F test and overidentification) and the

endogeneity test.

The reason why the regressions for the instrumented levels of education are not divided by gender is because when years of education (a continuous variable) is used as the schooling variable, the sample size is much larger (9131 observations from Table 3.1) than the sample used when estimating the model for education levels. For example, the model with the dummy for those individuals that have primary education versus secondary first stage would be estimated on 3106 observations. As mentioned before, when using any instrumental variable the sample is restricted to individuals older than twenty four and younger than sixty. The statistical significance of the coefficients depends much on the sample size used. If the regressions for the instrumented level of schooling are split by gender, the models do not perform that well as when estimating a regression for both genders[10]. That is the explanation why we only estimate one regression (including both genders) for each of the three dummies. Two regressions are run for each dummy. OLS and IV will be used here. In the 2SLS approach, the endogenous variable will be the dummy for the level of education.

In order to make the instrumental variable approach work, we need variables (instruments) that

- Explains part of the variation in education.
- Are not correlated with the unobservables (such as ability).

PRIMARY EDUCATION VS SECONDARY FIRST STAGE (D_1)

[10] The models in Tables 3.12 3.19 have been calculated including a dummy for gender. Results do not change substantially from the ones presented.

For the lower levels dummy, the set of instruments chosen are: the specific level of education of the mother when the individual was fourteen years old and the Spanish public expenditure on education in the year when the individual reached twelve.

Mother's education is a classic instrument for education and is used in this study to instrument this lower level. It is expected that Spanish mothers spent more time with her child than fathers did when the child was young. This could be explained because of the traditional family structure that applies for the generations analyzed: some mothers quit their jobs when they had a child or they stopped working when they got married so they devoted their time more to childcare duties while the father was the head of the household. Therefore is expected that the education of the mother would affect more the education of the offspring than father's education when the individuals in the sample were children. The coefficients on mother's levels of education have the expected sign. The coefficients increase in size: the higher the level of education achieved by the mother, the larger the effect on the education of the son/daughter (Table 3.12 column (a)). If the mother is more educated, it is more likely that the individual achieves secondary first stage level of education.

The reason why the age of twelve is chosen to get the data on public expenditure on education relies on the approximate age when the individuals are expected to finish primary education and start secondary education. It would have been more accurate if the expenditure in education variable refers to a regional level rather than the national expenditure but unfortunately there are no old data of the expenditure in education at a regional level and the region where the individual lived when the individual was a teenager is also unknown. Bear in mind that the sample analysed comprises individuals subject to different education systems. Different education systems were introduced at a different time in each region. This means that the exact year when the individual goes into the next level of education and the region where it happened are unknown; therefore a raising of the school leaving age is not a feasible instrument in this chapter.

In order to support the idea of choosing the national public expenditure on education when the individual was twelve as an instrumental variable a brief explanation is given. The national public expenditure in other years was tried without obtaining any statistically significant coefficient for the expenditure variable in the first stage equation. The variables tried were: the national public expenditure when the individual was born and the national public expenditure when the individual was six years old. None of them work well to instrument these lowest levels of education. The idea of using the public expenditure as a variable to explain education comes from Barceinas (2003). As he explains, when an economy is in recession, governments tend to diminish the national social expenditure. A part of the national social expenditure is devoted to education. Therefore the education budget is reduced. This adjustment could affect the individual decisions of continuing studying in the following way: a reduction of the national public expenditure could cause a decrease in the number of scholarships offered or a decrease in the amount of money invested in educational infrastructure and resources. According to this explanation, the expected sign for the coefficient should be positive. Table 3.12 column (c) shows that the public expenditure used on its own as an instrument for this level has a small size but statistically positive and significant coefficient.

As proved in columns (a) and (c) of Table 3.12 both instruments are separately valid. They are statistically significant in the first stage equation. For these low levels it is better to use the mother's education disaggregated dummies indicating all the possible levels rather than using the dummy that takes the value of 1 if the mother has higher education and 0 otherwise. Using a dummy indicating if the mother has attained higher education (where the reference category groups the rest of the categories available for mother's education) to instrument these low levels does not work well because it is not statistically significant in the first stage schooling equation.

The size of the effect of the expenditure in education on the achievement of secondary first stage level when using both instruments is quite small (because being measured in small units) but the sign of the coefficient is as expected. If the mother has higher education, this has a larger impact than if she has primary education again using both instruments. Again, using both instruments, if the government increases the education expenditure when the individual was twelve, this would have a positive effect on the acquisition of secondary first stage education. Individuals who attain secondary first stage education level have a wage premium of around 39 percent points in comparison with those who only get primary education.

If we look to the regression that includes both instruments, it can be seen that they all pass the tests for instrument validity (F test and overidentification). Out of several sets tried to instrument these low consecutive education levels, these are the combination of variables that are suitable together as instruments for this particular education level. Further checks have being made for the validity of education expenditure as an instrument for the Spanish case. If 2SLS is used in the wage equation to obtain the coefficient for the education level, and we introduce the education expenditure in the wage equation as a control variable, the coefficient of education expenditure is not statistically significant. It can be seen that this variable does not affect directly wages (Table 3.13 column (b)). If we observe the standard error of the coefficient of education in column (f), when using both instruments, is lower

(0.07) than when the single instruments are used, 0.12 in column (b) and 0.11 column (b) from Table 3.12. This means that the extra variation in education is more precise when both instruments are used.

To conclude the analysis of the low levels of education, several combinations of instruments were tried and do not pass the validity test when used together. Note that if father's education is used instead of mothers education the coefficient on education in the wage equation is not statistically significant anymore (which contradicts Human Capital Theory). Instrument sets discarded include:[11] mother‘s education combined with father's occupation, type of household tenancy, number of children in the household younger than eighteen (equivalent to number of siblings), number of people in the household that worked or household composition. They do not serve as instruments to reflect the difference between attaining a primary education level and a secondary first stage because they are not significant for education in the first stage equation or they do not pass the validity tests.

SECONDARY FIRST STAGE VS SECONDARY SECOND STAGE (D_2)

For the middle level dummies, two specifications are presented. The first one, in Table 3.14, contains this set of instruments: mother's higher education and father's occupation. The alternative specification, in Table 3.16, contains the wage equations estimated by 2SLS using two single instruments that represent different aspects of the financial status of the household when the individual was fourteen years old.

[11] Instrument sets discarded available upon request.

The difference between these two education levels, in terms of the years that it takes to attain the higher level (secondary second stage education) is not that large. People who only have secondary first stage education have attended all the courses of compulsory education (no matter what education system applies to them). They are supposed to acquire basic cognitive skills and general knowledge. Instead, secondary stage holders go further with their education and they do the pre-university courses which take approximately two years.

The first specification chosen explains the difference between these two consecutive levels through the mother's higher education and father's occupation variables. The reason why the specific mother's education level is not used as it was in Table 3.13 is because, if the different dummies are used for the specific mother's education level, the model does not pass the overindetification test. The only specific mother's education attained level that can be used as an instrument is higher education. Note that if father's education is used as an instrument for these levels, it is statistically significant in the first stage. The reason why it is not reported is that if it is going to be combined with father's occupation, it can be argued that father's occupation could depend on the level of education achieved by him. Mother's higher education and father's occupation also work as a good set of instruments for the next levels of education (higher education versus secondary second stage versus). Why these instruments only explain variation in the attainment of higher levels could be explained by thinking that they do not impact the offspring's education until he/she becomes older and has to decide between just obtaining secondary second stage education or tertiary education. If mother's occupation is used as an instrument instead of father's occupation the endogeneity test is not passed. This could be due to sample size given the traditional low level of labour participation of Spanish women.

When father's occupation is used, the regressions are on around 3500 observations while if mother's occupation is used, the regressions are reduced to 1000 observations because of missing values contained in the mother's occupation variable. This could be due to the fact that, as said before, for the generations studied, there is a large percentage of mothers that are housewives, so they do not report any occupation.

As can be seen in Table 3.14, the coefficient on the father's occupation dummies have the expected sign. It can be seen that father's occupations that require more education levels are the ones that have a larger effect on the offspring's own education. Fathers who are managers, professionals and skilled workers would have a positive larger impact than those fathers who hold an elementary occupation on the probability of the child getting further education (specifically higher levels than primary education). Further tests are carried out on in support of father's occupation. Table 3.15 shows that if the father's occupation dummies are included as control variables in the wage equation, none of their coefficients are statistically significant. It can be said that father's occupation has no further independent effect on wages.

The second specification specification in Table 3.16, explains the difference between individuals that attain secondary first stage education and secondary second stage education through the economic characteristics of the household. The instruments are financial difficulty and financial situation. It can be seen that they are both equivalent. Using one or the other gives a similar coefficient for education, a wage premium of around 104 percent in column (b) and around 122 percent points in column (d) from Table 3.16.

Regarding the statistical analysis, both coefficients on the financial household variables are statistically significant in the first stage. These financial variables are valid as single instruments for all the education levels. This means that the difference between attaining one level or another could be explained by the financial situation of each household when the adult was a teenager. This makes complete sense because if an individual wants to continue studying, he/she (or their families) needs enough economic resources to buy stationery and books, to pay for transport to commute if needed or to pay the annual fees if the school or education center is not fully owned by the state (public schools) not to mention the opportunity cost of keeping on studying. The reason why they are chosen to be reported to explain these levels is to show that even if the difference between the consecutive levels is short in terms of the number of years that it would take to acquire the highest level (secondary second stage in this case), nevertheless the variables still serve to explain the variation. If financial variables that describe the economic situation of the household when the adult was a teenager are used as instruments, the models pass all instrumental variables validating tests. They are individually statistically significant at the first stage and their coefficients indicate that, if the family do not have financial difficulties to make ends meet and the financial situation is good, this would positively affect the probability of attaining the secondary second stage level.

Bear in mind that the Human Capital Theory continues to hold in the case of Spain if levels of education are used as schooling variables. The higher the level of education attained is, the higher the wage premium that an individual gets. This can be checked if the education coefficients are compared between the lowest levels, Table 3.12 column (f) where the education coefficient is 39 percentage points when using both instruments and Table 3.14 column (f) with a premium of 92 percentage points for having attained the secondary second stage level in comparison with the ones that attain secondary first stage.

Taken together, the results just presented suggest that both instruments, mother's higher education and father's occupation are relevant variables for the highest level (secondary second stage) because they are individually statistically significant in the first stage. Also when used together, they are still relevant for the difference in education. Financial situation and financial difficulty are also valid instruments for these levels. To conclude the analysis of the middle levels of education, instrument sets discarded include: father's education, father's higher education and type of household tenancy, number of children in the household younger than eighteen, household composition. They do not pass the validity tests to reflect the difference between attaining a secondary first stage education level and a secondary second stage.

SECONDARY SECOND STAGE VS HIGHER EDUCATION (D_3)

For the highest levels dummies, the set of instruments chosen are: mother's higher education and father's occupation when the adult was fourteen. It is expected that a combination of two variables about the family socioeconomic status, mother's higher education dummy and father's occupation, it is sufficient to explain the variation between obtaining one of these levels.

Table 3.17 column (a) shows that the mother's higher education instrument has the expected sign in the first stage regression. The fact that the mother has attained higher education would affect positively the probability that the son/daughter attains higher education too. According to MECD (2012), a third of all the students that pass the university access exam have a father and/or mother who has attained higher education. It is reasonable to think that parents' education will affect the offspring's higher education achievement because in Spain, individuals tend to stay at home studying for their university degrees and therefore living at the parents' home reduces the cost of studying. MECD (2011) shows evidence of a very low residential mobility on account of higher education in Spain. The co-residence of the students with their parents until the end of the university period influences the student's decisions on education.

The signs of the coefficient on the instruments related to father's occupation are as expected. Father's occupation could proxy for the socioeconomic status of the family or how wealthy the family is. If the father is a manager or professional, his wage will be higher than a father who has a skilled worker occupation or an unskilled worker occupation. Managers and professionals will earn more money and therefore the family will be wealthier. Wealthy families could spend more money on the education of the children. Therefore, if the individual has a father whose occupation is manager or professional, this increases the probability of achieving higher education in comparison with those whose father has a skilled worker occupation or an elementary occupation. This can be seen in column (c) and (e) of Table 3.17. Father's occupation pass all the specification checks. As can be seen in Table 3.17 column (c) and (e) father's occupation is individually statistically significant in the first stage. If both instruments, mother's higher education and father's occupation, are used together, column (f) Table 3.17, the F test of joint significance and the overidentification test are passed. As done for the lower levels, if father's occupation is included as a control variable in the wage equation, Table 3.18, it is not statistically significant so does not affect wages.

Bear in mind that if father's education is used as an instrument combined with the type of household tenancy when the adult was fourteen, this set also works and passes all the robustness checks. There are several combinations of instruments that are valid to explain the attainment of higher education but, according to further tests the best combination of instruments is mother's higher education and father's occupation. Instruments that are not valid to explain the achievement of higher education are: number of children in the household younger than eighteen, number of people in the household who worked when the adult was fourteen or household composition. They are not valid because they do not pass some validity tests or they are not individually significant in the first stage.

The returns to attain a higher education level are 136 percentage points in comparison with those who attain only secondary second stage education. The higher education holders include graduates from several subjects such as doctors, architects, engineers, etc. and also individuals who have done a postgraduate program such as a Masters.

To conclude the results section, it is worth saying that if levels of schooling are instrumented using sets of variables as it is done in Tables 3.12, 3.14 and 3.17, there is also a downwards bias in the OLS coefficients for education (8.6, 15 and 39.7 percent respectively for each level) reported on Table 3.19. For the lowest levels, comparing the 2SLS coefficient on education from column (f) Table 3.12 with the first column in Table 3.19 the magnitude of the bias is of 30 percentage points. For the medium levels, comparing the 2SLS coefficient on education from column (f) Table 3.14 with the third column in Table 3.19 the magnitude of the bias is even larger, 77 percentage points. For the higher levels, comparing the coefficient of education of the 2SLS from column (f) Table 3.17 with the second column in Table 3.19 the magnitude of the bias is 97 percentage points. The bias in the schooling coefficients increase as the levels attained increases. The difference in the coefficients obtained by OLS and 2SLS is statistically significant. The p-value of the endogeneity test is always 0.000 meaning that the null hypothesis can always be rejected for all levels of education. The null hypothesis of education being an exogenous variable could be rejected, therefore 2SLS is a more adequate method in comparison with OLS.

Chapter 3 Tables

Table 3.1: Frequencies for the schooling variables (individuals aged between 25-59)

Schooling Variable	**Freq. (Observations)**
Years of Schooling	9131
No Education	68
Primary	971
Secondary First Stage	2135
Secondary Second Stage	2177
Higher Education	3811

Table 3.2: Sample Summary Statistics. Wage-Earners (17-65)

Variable	Mean	St. Dev	Min	Max
Experience	18.303	10.658	0	55
Experience^2	448.578	467.690	0	3025
Married	0.591	0.492	0	1
Income rest household	20509.420	18731.500	-40000	181373.8
Dependent Children	0.502	0.500	0	1
Employee	0.538	0.498	1	0
ln(wage)	2.148	0.425	0.069	4.080
Male	0.544	0.498	0	1
Public Education Expenditure	443291.9	348285.2	3855	1118000
Financial Difficulty	0.607	0.488	0	1
Financial Situation	0.724	0.446	0	1
EDUCATION				
No Education	0.019	0.136	0	1
Primary	0.112	0.315	0	1
Secondary First Stage	0.228	0.420	0	1
Secondary Second Stage	0.246	0.431	0	1
Higher Education	0.402	0.490	0	1
Years of Schooling	12.687	4.095	0	21
MOTHER'S EDUCATION				
No education	0.051	0.219	0	1
Secondary First Stage	0.825	0.380	0	1
Secondary Second Stage	0.068	0.252	0	1
Higher education	0.055	0.228	0	1
FATHER'S HIGHER EDUCATION	0.108	0.310	0	1
FATHER'S OCCUPATION				
Managers & Professionals	0.256	0.436	0	1
Skilled Worker	0.503	0.499	0	1
Unskilled Worker	0.240	0.427	0	1
REGIONS				
North	0.106	0.308	0	1
North West	0.089	0.285	0	1
East	0.305	0.460	0	1
South	0.231	0.421	0	1
Inner	0.096	0.294	0	1
Madrid	0.172	0.378	0	1

Notes: Summary statistics are based on weighted values.

Table 3.3: Distribution of Labour Force in Spain by Gender (16 - 64).

Status		Male N	Male %	Female N	Female %	Total N	Total %
Unemployed	(a)	2,892.2	18.23	2,538.1	16.24	5,430.5	17.24
Employed	(b) = (c) + (d) + (e)	10,068.1	63.45	8,202.8	52.50	18,271.5	58.02
Self-employed	(c)	1,962.0	12.37	976.5	6.25	2,938.6	9.33
Wage-earners	(d)	8,102.3	51.06	7,222.6	46.22	15,324.9	48.66
Others	(e)	3.7	0.02	3.7	0.02	7.4	0.02
Total Labour Force	(f) = (a) + (b)	12,960.3	81.68	10,740.9	68.74	23,702.0	75.26
Out of Labour Force	(g)	2,906.5	18.32	4,884.7	31.26	7,791.2	24.74
All	(h) = (f) + (g)	15,866.8	100	15,625.6	100	31,493.2	100

Table 3.4: Average Monthly Earnings of Wage-Earners in Euros, by Education Level and Gender

Education Level / Gender	Male	Female	t-test (M-F)	
NO EDUCATION	1066.4	667.9	-3.49	***
	(65.83)	(87.16)		
PRIMARY	1225.6	823.0	-14.89	***
	(17.76)	(18.44)		
SECONDARY 1	1295.3	894.0	-21.19	***
	(13.24)	(12.06)		
SECONDARY 2	1499.2	1074.9	-17.85	***
	(18.03)	(15.04)		
HIGHER ED.	1970.8	1622.9	-12.94	***
	(22.11)	(16.22)		
ALL	1565.3	1268.0	-20.03	***
	(10.68)	(10.14)		

Notes: Standard errors of the mean are below the coefficients.

Table 3.5: **OLS Regression of the LN_WAGE, by Gender (16-65) and Schooling.**

Variables	Males				Females			
	Years (a)		Levels (b)		Years (c)		Levels (d)	
	Coeff.		Coeff.		Coeff.		Coeff.	
YEARS OF ED.	0.049	***			0.058	***		
	(0.00)				(0.00)			
EXPERIENCE	0.034	***	0.031	***	0.026	***	0.023	***
	(0.00)		(0.00)		(0.00)		(0.00)	
EXPERIENCE^2	-4.73E-04	***	-4.12E-04	***	-3.0E-04	***	-2.4E-04	***
	(0.00)		(0.00)		0.00		(0.00)	
NO EDUC.			-0.107	**			0.062	
			(0.05)				(0.07)	
SECONDARY 1			0.094	***			0.080	
			(0.02)				(0.02)	
SECONDARY 2			0.246	***			0.222	**
			(0.02)				(0.02)	
HIGHER ED.			0.526	***			0.602	***
			(0.02)				(0.07)	
NORTH WEST	-0.110	***	-0.124	***	-0.060	***	-0.076	***
	(0.02)		(0.02)		(0.02)		(0.02)	
NORTH	0.017		-0.001		0.077	***	0.042	**
	(0.02)		(0.02)		(0.02)		(0.02)	
EAST	-0.066	***	-0.076	***	0.003		-0.007	
	(0.02)		(0.02)		(0.02)		(0.02)	
SOUTH	-0.125	***	-0.125	***	-0.01		-0.028	
	(0.02)		(0.02)		(0.02)		(0.02)	
INNER	-0.063	***	-0.077	***	0.017		-0.002	
	(0.02)		(0.02)		(0.02)		(0.02)	
CONSTANT	1.229	***	1.615	***	1.031	***	1.485	***
	(0.03)		(0.02)		(0.03)		(0.03)	
R^2	0.33		0.35		0.34		0.39	
N	4926		4936		4477		4480	

Notes: *, **, * means p-value $<$0.10, p-value $<$0.05 and p-value $<$0.010 respectively. PRIMARY ED. and MADRID are the reference categories for education splines and regions respectively**

Table 3.6: Binary Probit Estimation of Employees (17-65), by Gender

Variables	Males			Females		
	Coefficient		Marg. Effect	Coefficient		Marg. Effect
EXPERIENCE	0.067	***	0.019	0.111	***	0.034
	(0.01)		(0.00)	(0.00)		(0.00)
EXPERIENCE^2	-0.001	***	0.000	-0.002	***	-0.001
	(0.00)		(0.00)	(0.00)		(0.00)
NO EDUC.	-0.397	***	-0.143	-0.375	***	-0.116
	(0.12)		(0.04)	(0.13)		(0.03)
SECONDARY 1	0.390	***	0.135	0.396	***	0.135
	(0.05)		(0.03)	(0.05)		(0.02)
SECONDARY 2	0.685	***	0.224	0.658	***	0.225
	(0.05)		(0.02)	(0.05)		(0.02)
HIGHER ED.	1.063	***	0.314	1.211	***	0.394
	(0.05)		(0.02)	(0.05)		(0.02)
MARRIED	0.272	***	0.081	-0.396	***	-0.121
	(0.05)		(0.01)	(0.04)		(0.01)
INCOME REST	-2.42E-06	***	0.000	-3.56E-06	***	0.000
	(0.00)		(0.00)	(0.00)		(0.00)
DEPENDENT	0.095	***	0.028	0.111	***	0.035
	(0.04)		(0.01)	(0.03)		(0.01)
NORTH WEST	-0.063		-0.018	-0.155	**	-0.048
	(0.07)		(0.01)	(0.06)		(0.01)
NORTH	0.062		0.018	-0.089		-0.028
	(0.07)		(0.01)	(0.06)		(0.01)
EAST	-0.083		-0.024	-0.175	***	-0.054
	(0.06)		(0.01)	(0.06)		(0.01)
SOUTH	-0.285	***	-0.082	-0.309	***	-0.096
	(0.06)		(0.01)	(0.06)		(0.01)
INNER	-0.035		-0.010	-0.166	**	-0.052
	(0.07)		(0.02)	(0.07)		(0.02)
CONSTANT	-0.551	***	(-)	-1.008	***	(-)
	(0.08)			(0.08)		
Log L		-3814.77			-4388.00	
Pseudo-R^2		0.137			0.188	
$\chi^2_{(5)}$ Regions		54.45			36.75	
N		7427			8005	

Notes: *, **, * means p-value <0.10, p-value <0.05 and p-value <0.010 respectively. Dependent Variable: EMPLOYEE. PRIMARY ED. and MADRID are the reference categories for education splines and regions respectively**

Table 3.7: Heckman Sample Selection Model. Wage equations, by Gender (15 - 65)

Variables	Males				Females			
	Years (a)		Levels (b)		Years (c)		Levels (d)	
	Coeff.		Coeff.		Coeff.		Coeff.	
YEARS OF ED.	0.068	***			0.082	***		
	(0.00)				(0.00)			
EXPERIENCE	0.051	***	0.043	***	0.050	***	0.054	***
	(0.00)		(0.00)		(0.01)		(0.01)	
EXPERIENCE^2	-0.001	***	-0.001	***	-0.001	***	-0.001	***
	(0.00)		(0.00)		(0.00)		(0.00)	
MARRIED	0.164	***	0.149	***	0.016		-0.025	
	(0.02)		(0.02)		(0.02)		(0.02)	
NO EDUC.			-0.205	***			-0.151	*
			(0.05)				(0.09)	
SECONDARY 1			0.191	***			0.227	***
			(0.02)				(0.03)	
SECONDARY 2			0.397	***			0.453	***
			(0.03)				(0.04)	
HIGHER ED.			0.735	***			0.965	***
			(0.04)				(0.06)	
NORTH WEST	-0.109	***	-0.128	***	-0.082	***	-0.113	***
	(0.03)		(0.02)		(0.02)		(0.03)	
NORTH	0.042	*	0.015		0.072	***	0.020	
	(0.02)		(0.02)		(0.02)		(0.03)	
EAST	-0.071	***	-0.077	***	-0.019		-0.042	*
	(0.02)		(0.02)		(0.02)		(0.02)	
SOUTH	-0.181	***	-0.168	***	-0.066	**	-0.110	***
	(0.02)		(0.02)		(0.03)		(0.03)	
INNER	-0.053	**	-0.072	***	-0.008		-0.042	
	(0.03)		(0.02)		(0.03)		(0.03)	
CONSTANT	0.543	***	1.093	***	0.259	*	0.691	***
	(0.11)		(0.09)		(0.15)		(0.13)	
LAMBDA	0.466	***	0.411	***	0.394	***	0.493	***
	(0.06)		(0.06)		(0.07)		(0.07)	
N-Uncensored	4611		4620		4126		4129	
Wald Chi	544.56		685.25		675.31		724.82	
P-Value Wald Test	0.000		0.000		0.000		0.000	

Notes: *, **, * means p-value <0.10, p-value <0.05 and p-value <0.010 respectively. Dependent Variable: LN_WAGE. PRIMARY ED. and MADRID are the reference categories for education levels and regions respectively**

Table 3.8: Instrumental Variable Estimation. Dependent Variable: LN_WAGE. Wage-Earners (Males), using Years of Schooling.

Variables	Father's education				Mother's education				Parent Education			
	(a) First Stage		(b) IV		(c) First Stage		(d) IV		(e) First Stage		(f) IV	
YEARS OF ED.			0.071	***			0.072	***			0.072	***
			(0.01)				(0.01)				(0.01)	
EXPERIENCE	-0.120	***	0.038	***	-0.116	***	0.038	***	-0.108	***	0.038	***
	(0.02)		(0.00)		(0.02)		(0.00)		(0.02)		(0.00)	
EXPERIENCE^2	0.000		5.00E-04	***	1.09E-04		5.00E-04	***	1.05E-04		5.00E-04	***
	(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)	
NORTH WEST	-0.740	**	-0.092	***	-0.984	***	-0.099	***	-0.697	**	-0.097	***
	(0.24)		(0.02)		(0.24)		(0.02)		(0.24)		(0.02)	
NORTH	-0.405	*	0.023		-0.538	**	0.022		-0.378	*	0.019	
	(0.22)		(0.02)		(0.22)		(0.02)		(0.22)		(0.02)	
EAST	-1.082	***	-0.049	**	-1.275	***	-0.052	**	-1.050	***	-0.052	**
	(0.21)		(0.02)		(0.21)		(0.02)		(0.21)		(0.02)	
SOUTH	-1.396	***	-0.104	***	-1.602	***	-0.104	***	-1.360	***	-0.106	***
	(0.21)		(0.02)		(0.21)		(0.03)		(0.21)		(0.02)	
INNER	-0.869	***	-0.046	**	-1.137	***	-0.047	*	-0.840	***	-0.049	**
	(0.24)		(0.02)		(0.24)		(0.03)		(0.24)		(0.02)	
FATHER'S HIGHER ED.	3.348	***							2.955	***		
	(0.20)								(0.21)			
MOTHER'S HIGHER ED.					2.956	***			1.487	***		
					(0.27)				(0.29)			
CONSTANT	15.278	***	0.883	***	15.637	***	0.880	***	15.069	***	0.88	***
	(0.29)		(0.10)		(0.29)		(0.14)		(0.29)		(0.09)	
Overidentification p-value			(-)				(-)				0.895	
Endogeneity p-value			0.000				0.007				0.000	
F-Test First Stage			378.00				179.96				214.96	
C-D Wald F			287.00				116.265				154.181	
Partial R^2 of the IV			0.0626				0.026				(-)	
N			4307				4359				4274	

Notes: *, **, * means p-value $<$0.10, p-value $<$0.05 and p-value $<$0.010 respectively. MADRID is the reference category for the regions. FATHER'S ED and MOTHER'S ED are the instrumental variables; they are dummies for higher education of the father and the mother respectively.**

Table 3.9: Instrumental Variable Estimation. Dependent Variable: LN_WAGE. Wage-Earners (Females), using Years of Schooling.

Variables	Father's education				Mother's education				Parent Education			
	(a) First Stage		(b) IV		(c) First Stage		(d) IV		(e) First Stage		(f) IV	
YEARS OF ED.			0.096	***			0.091	***			0.096	***
			(0.01)				(0.01)				(0.01)	
EXPERIENCE	-0.113	***	0.031	***	-0.099	***	0.030	***	-0.105	***	0.031	***
	(0.02)		(0.00)		(0.02)		(0.00)		(0.02)		(0.00)	
EXPERICENCE^2	0.001		0.000	***	0.001		0.000	***	0.001		0.000	***
	(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)	
NORTH WEST	-0.855	***	-0.039		-0.792	**	-0.045	*	-0.840	***	-0.042	*
	(0.23)		(0.02)		(0.23)		(0.02)		(0.23)		(0.02)	
NORTH	-0.347		0.089	***	-0.429	**	0.087	***	-0.365	*	0.090	***
	(0.21)		(0.02)		(0.22)		(0.02)		(0.21)		(0.02)	
EAST	-0.923	***	0.045	**	-1.094	***	0.038		-0.974	***	0.046	**
	(0.20)		(0.02)		(0.21)		(0.02)		(0.20)		(0.02)	
SOUTH	-1.171	***	0.039	*	-1.286	***	0.027		-1.177	***	0.037	
	(0.21)		(0.02)		(0.21)		(0.02)		(0.21)		(0.02)	
INNER	-0.727	**	0.060	**	-0.837	***	0.054	**	-0.726	**	0.059	**
	(0.24)		(0.03)		(0.24)		(0.03)		(0.24)		(0.03)	
FATHER'S HIGHER ED.	2.954	***							2.493	***		
	(0.19)								(0.21)			
MOTHER'S HIGHER ED.					2.996	***			1.585	***		
					(0.27)				(0.29)			
CONSTANT	15.318	***	0.424	***	15.428	***	0.510	***	15.205	***	0.426	***
	0.26		-0.11		0.27		-0.14		0.27		0.10	
Overidentification p-value			(-)				(-)				0.353	
Endogeneity p-value			0.000				0.000				0.000	
F-Test First Stage			452.44				298.41				282.59	
C-D Wald F			250.33				127.26				137.31	
Partial R of the IV			0.0595				0.031					
N			3963				3991				3928	

Notes: *, **, * means p-value <0.10, p-value <0.05 and p-value <0.010 respectively. MADRID is the reference category for the regions. FATHER'S ED and MOTHER'S ED are the instrumental variables.**

Table 3.10: Fuller LIML estimator. Dependent Variable: LN_WAGE. Males

Variables	Father's education		Mother's education		Parent Education	
	(a) First Stage	(b) IV	(c) First Stage	(d) IV	(e) First Stage	(f) IV
YEARS OF ED		0.071 ***		0.071 ***		0.072 ***
		(0.01)		(0.01)		(0.01)
EXPERIENCE	-0.120 ***	0.038 ***	-0.116 ***	0.038 ***	-0.108 ***	0.038 ***
	(0.02)	(0.00)	(0.02)	(0.00)	(0.02)	(0.00)
EXPERIENCE^2	0.000	0.000 ***	0.000	0.000 ***	0.000	0.000 ***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
NORTH WEST	-0.740 ***	-0.092 ***	-0.984 ***	-0.099 ***	-0.697 ***	-0.097 ***
	(0.22)	(0.02)	(0.23)	(0.02)	(0.23)	(0.02)
NORTH	-0.405 **	0.023	-0.538 ***	0.022	-0.378 *	0.019
	(0.20)	(0.02)	(0.21)	(0.02)	(0.20)	(0.02)
EAST	-1.082 ***	-0.049 **	-1.275 ***	-0.052 **	-1.050 ***	-0.052 **
	(0.20)	(0.02)	(0.21)	(0.02)	(0.20)	(0.02)
SOUTH	-1.396 ***	-0.104 ***	-1.602 ***	-0.104 ***	-1.360 ***	-0.106 ***
	(0.21)	(0.02)	(0.21)	(0.03)	(0.21)	(0.02)
INNER	-0.869 ***	-0.047 **	-1.137 ***	-0.047 *	-0.840 ***	-0.049 **
	(0.23)	(0.02)	(0.23)	(0.03)	(0.23)	(0.02)
FATHER'S HIGHER ED.	3.348 ***				2.955 ***	
	(0.17)				(0.19)	
MOTHER'S HIGHER ED.			2.956 ***		1.487 ***	
			(0.22)		(0.24)	
CONSTANT	15.277 ***	0.884 ***	15.637 ***	0.883 ***	15.069 ***	0.879 ***
	(0.27)	(0.10)	(0.28)	(0.16)	(0.28)	(0.10)
F-Test First Stage		380.22		180.78		219.06
N		4307		4359		4274

Notes: *, **, * means p-value <0.10, p-value <0.05 and p-value <0.010 respectively. MADRID is the reference category for the regions. FATHER'S ED and MOTHER'S ED are the instrumental variables.**

Table 3.11: Fuller LIML estimator. Dependent Variable: LN_WAGE. Females

Variables	Father's education		Mother's education		Parent Education	
	(a) First Stage	(b) IV	(c) First Stage	(d) IV	(e) First Stage	(f) IV
YEARS OF ED.		0.096 ***		0.091 ***		0.096 ***
		(0.01)		(0.01)		(0.01)
EXPERIENCE	-0.113 ***	0.031 ***	-0.099 ***	0.030 ***	-0.105 ***	0.031 ***
	(0.02)	(0.00)	(0.02)	(0.00)	(0.02)	(0.00)
EXPERICENCE^2	0.001	0.000 ***	0.001	0.000 ***	0.001	0.000 ***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
NORTH WEST	-0.855 ***	-0.039	-0.792 ***	-0.045 *	-0.840 ***	-0.042 *
	(0.22)	(0.02)	(0.23)	(0.02)	(0.22)	(0.02)
NORTH	-0.347 *	0.089 ***	-0.429 **	0.087 ***	-0.365 *	0.090 ***
	(0.20)	(0.02)	(0.21)	(0.02)	(0.20)	(0.02)
EAST	-0.923 ***	0.044 **	-1.094 ***	0.037	-0.974 ***	0.046 **
	(0.20)	(0.02)	(0.21)	(0.02)	(0.20)	(0.02)
SOUTH	-1.171 ***	0.039	-1.286 ***	0.027	-1.177 ***	0.037
	(0.21)	(0.02)	(0.22)	(0.03)	(0.21)	(0.02)
INNER	-0.727 ***	0.060 **	-0.837 ***	0.054 **	-0.726 ***	0.059 **
	(0.24)	(0.03)	(0.24)	(0.03)	(0.24)	(0.03)
FATHER'S HIGHER ED.	2.954 ***				2.493 ***	
	(0.14)				(0.16)	
MOTHER'S HIGHER ED.			2.996 ***		1.585 ***	
			(0.17)		(0.20)	
CONS	15.319 ***	0.426 ***	15.428 ***	0.514 ***	15.205 ***	0.427 ***
	(0.27)	(0.11)	(0.27)	(0.15)	(0.27)	(0.11)
F-Test First Stage		449.78		302.99		281.81
N		3963		3991		3928

Notes: *, **, * means p-value <0.10, p-value <0.05 and p-value <0.010 respectively. MADRID is the reference category for the regions. FATHER'S ED. and MOTHER'S ED are the instrumental variables.**

Table 3.12: Primary Education VS Secondary First Stage. Dependent Variable: LN_WAGE.

Variables	Mother's Levels of Education		Public Education Expenditure		Both Instruments	
	(a) First Stage	(b) IV	(c) First Stage	(d) IV	(e) First Stage	(f) IV
SECONDARY 1 (D_1)		0.418 ***		0.254 **		0.330 ***
		(0.12)		(0.11)		(0.07)
EXPERIENCE	0.013 ***	0.007 **	0.015 ***	0.013 ***	0.023 ***	0.008 ***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
EXPERICENCE^2	0.000 ***	0.000	0.000 ***	0.000	-0.001 ***	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
NORTH WEST	-0.088 ***	-0.063 **	-0.095 ***	-0.065 **	-0.087 **	-0.070 **
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
NORTH	-0.105 ***	0.113 ***	-0.087 ***	0.098 ***	-0.105 ***	0.104 ***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
EAST	-0.154 ***	0.029	-0.150 ***	0.006	-0.161 ***	0.015
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
SOUTH	-0.156 ***	-0.041	-0.158 ***	-0.062 **	-0.162 ***	-0.056 **
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
INNER	-0.047	-0.012	-0.047	-0.021	-0.047	-0.017
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
EDUCATION EXP.			2.280E-07 ***		2.940E-07 ***	
			(0.00)		(0.00)	
MOTHER'S ED.						
Secondary 1	0.201 ***				0.183 ***	
	(0.03)				(0.03)	
Secondary 2	0.306 ***				0.265 ***	
	(0.06)				(0.06)	
Higher Ed.	0.328 ***				0.269 ***	
	(0.10)				(0.10)	
CONSTANT	0.584 ***	1.473 ***	0.638 ***	1.556 ***	0.320 ***	1.543 ***
	(0.06)	(0.10)	(0.05)	(0.09)	(0.07)	(0.07)
Overidentification p-value						0.125
Endogeneity p-value						0.000
F-Test First Stage						27.974
N		2816		3240		2816

Notes: *, **, * means p-value <0.10, p-value <0.05 and p-value <0.010 respectively. MADRID is the reference category for the regions. Variable instrumented: SECONDARY 1. No Education is the reference category for the mother's level of education.**

Table 3.13: Instrumental Variable Estimation using MOTHER'S ED., including EDUCATION EXP. as an Explanatory Variable. Dependent Variable: LN_WAGE.

Variables	Mother's Levels of Education			
	(a) First Stage		(b) IV	
SECONDARY 1 (D_1)			0.428	***
			(0.13)	
EXPERIENCE	0.023	***	0.005	
	(0.00)		(0.00)	
EXPERICENCE^2	-0.001	***	0.000	
	(0.00)		(0.00)	
NORTH WEST	-0.087	**	-0.063	**
	(0.03)		(0.03)	
NORTH	-0.105	***	0.114	***
	(0.03)		(0.03)	
EAST	-0.161	***	0.031	
	(0.03)		(0.03)	
SOUTH	-0.162	***	-0.039	
	(0.03)		(0.03)	
INNER	-0.047		-0.012	
	(0.03)		(0.03)	
EDUCATION EXP.	2.94E-07	***	-4.970E-08	
	(0.00)		(0.00)	
MOTHER'S ED.				
Secondary 1	0.183	***		
	(0.03)			
Secondary 2	0.265	***		
	(0.06)			
Higher Ed.	0.269	***		
	(0.10)			
CONSTANT	0.320	***	1.513	***
	(0.07)		(0.08)	
N			2816	

Notes: *, **, * means p-value $<$0.10, p-value $<$0.05 and p-value $<$0.010 respectively. MADRID is the reference category for the regions. Variable instrumented: SECONDARY 1. No Education is the reference category for the mother's level of education.**

Table 3.14: Secondary First Stage VS Secondary Second Stage. Dependent Variable: LN_WAGE.

Variables	Mother's Higher Education				Father's Occupation				Both Instruments			
	(a)		(b)		(c)		(d)		(e)		(f)	
	First Stage		IV		First Stage		IV		First Stage		IV	
SECONDARY 2 (D_2)			0.542	***			0.644	***			0.653	***
			(0.16)				(0.10)				(0.10)	
EXPERIENCE	-0.003		0.021	***	-0.003		0.021	***	-0.003		0.022	***
	(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)	
EXPERICENCE^2	0.000		0.000	***	0.000		0.000	**	0.000		0.000	**
	(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)	
NORTH WEST	-0.013		-0.102	***	0.026		-0.124	***	0.021		-0.123	***
	(0.03)		(0.02)		(0.03)		(0.03)		(0.03)		(0.03)	
NORTH	0.003		0.028		0.032		0.008		0.029		0.007	
	(0.03)		(0.02)		(0.03)		(0.03)		(0.03)		(0.03)	
EAST	-0.025		-0.057	**	0.002		-0.074	***	-0.001		-0.072	***
	(0.03)		(0.02)		(0.03)		(0.02)		(0.03)		(0.02)	
SOUTH	-0.110	***	-0.046		-0.082	***	-0.057	**	-0.085	***	-0.053	*
	(0.03)		(0.03)		(0.03)		(0.03)		(0.03)		(0.03)	
INNER	-0.085	**	-0.019		-0.058	*	-0.041		-0.060	*	-0.039	***
	(0.03)		(0.03)		(0.03)		(0.03)		(0.03)		(0.03)	
MOTHER'S HIGHER ED.	0.292	***							0.203	***		
	(0.05)								(0.05)			
FATHER'S OCCUPATION												
Skilled Worker					0.060	***			0.061	***		
					(0.02)				(0.02)			
Managers & Professionals					0.200	***			0.185	***		
					(0.02)				(0.02)			
CONSTANT	0.606	***	1.478	***	0.520	***	1.440	***	0.517	***	1.430	***
	(0.04)		(0.10)		(0.04)		(0.07)		(0.04)		(0.07)	
Overidentification p-value											0.674	
Endogeneity p-value											0.000	
F-Test First Stage											27.284	
N			3911				3584				3541	

Notes: *, **, * means p-value <0.10, p-value <0.05 and p-value <0.010 respectively. MADRID is the reference category for the regions. Variable instrumented: SECONDARY 2. Instrumental Variables: MOTHER'S HIGHER ED. and FATHER'S OCCUPATION. The reference category for the last instrument gathers unskilled workers.**

Table 3.15: Instrumental Variable Estimation (D_2) using MOTHER'S HIGHER ED., including FATHER'S OCCUPATION as an Explanatory Variable. Dependent Variable: LN_WAGE.

Variables	Mother's Higher Education			
	(a) First Stage		(b) IV	
SECONDARY 2 (D_2)			0.648	**
			(0.26)	
EXPERIENCE	-0.003		0.022	***
	(0.00)		(0.00)	
EXPERIENCE^2	0.000		0.000	**
	(0.00)		(0.00)	
NORTH WEST	0.021		-0.122	***
	(0.03)		(0.03)	
NORTH	0.029		0.007	
	(0.03)		(0.03)	
EAST	-0.001		-0.071	***
	(0.03)		(0.02)	
SOUTH	-0.085	***	-0.054	
	(0.03)		(0.03)	
INNER	-0.060	*	-0.040	
	(0.03)		(0.03)	
MOTHER'S HIGHER ED.	0.203	***		
	(0.05)			
FATHER'S OCCUPATION				
Skilled Worker	0.061	***	-0.012	
	(0.02)		(0.02)	
Managers & Professionals	0.185	***	0.000	
	(0.02)		(0.05)	
CONSTANT	0.517		1.439	***
	(0.04)		(0.14)	
N			3541	

Notes: *, **, * means p-value <0.10, p-value <0.05 and p-value <0.010 respectively. MADRID is the reference category for the regions. Variable instrumented: SECONDARY 2. The reference category for the last instrument gathers unskilled workers.**

Table 3.16: Secondary First Stage VS Secondary Second Stage. Alternative instrumental variables. Dependent Variable: LN_WAGE.

Variables	Financial Difficulty				Financial Situation			
	(a)		(b)		(c)		(d)	
	First Stage		IV		First Stage		IV	
SECONDARY 2 (D_2)			0.713	***			0.799	***
			(0.13)				(0.14)	
EXPERIENCE	-0.004		0.021	***	-0.004		0.022	***
	(0.00)		(0.00)		(0.00)		(0.00)	
EXPERICENCE^2	0.000		0.000	**	0.000		0.000	**
	(0.00)		(0.00)		(0.00)		(0.00)	
NORTH WEST	-0.019		-0.091	***	-0.015		-0.090	***
	(0.03)		(0.03)		(0.03)		(0.03)	
NORTH	-0.016		0.036		-0.009		0.035	
	(0.03)		(0.03)		(0.03)		(0.03)	
EAST	-0.037		-0.047	*	-0.033		-0.045	*
	(0.03)		(0.03)		(0.03)		(0.03)	
SOUTH	-0.102	***	-0.020		-0.103		-0.011	
	(0.03)		(0.03)		(0.03)		(0.03)	
INNER	-0.100	***	0.007		-0.093		0.014	
	(0.03)		(0.03)		(0.03)		(0.03)	
FINANCIAL DIFF.	0.104	***						
	(0.02)							
FINANCIAL SIT.					0.112	***		
					(0.02)			
CONS	0.567	***	1.367	***	0.545	***	1.314	***
	(0.04)		(0.09)		(0.04)		(0.10)	
Overidentification p-value			(-)				(-)	
Endogeneity p-value			0.000				0.000	
F-Test First Stage			41.125				42.216	
N			3963				3969	

Notes: *, **, * means p-value <0.10, p-value <0.05 and p-value <0.010 respectively. MADRID is the reference category for the regions. Variable instrumented: SECONDARY 2. Instrumental Variables: FINANCIAL DIFF. and FINANCIAL SIT.**

Table 3.17: Secondary Second Stage vs Higher Education. Dependent Variable: LN_WAGE.

Variables	Mother's Higher Education		Father's Occupation		Both Instruments	
	(a) First Stage	(b) IV	(c) First Stage	(d) IV	(e) First Stage	(f) IV
HIGHER ED. (D_3)		0.759***		0.916***		0.859***
		(0.11)		(0.09)		(0.08)
EXPERIENCE	-0.005**	0.034***	-0.007***	0.035***	-0.005**	0.035***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
EXPERIENCE^2	0.000	0.000***	0.000	0.000***	0.000	0.000***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
NORTH WEST	-0.049**	-0.089***	-0.049*	-0.068***	-0.047*	-0.079***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
NORTH	0.018	-0.005	0.021	0.002	0.023	-0.006
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
EAST	-0.050**	-0.034*	-0.053**	-0.019	-0.049**	-0.027
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
SOUTH	-0.026	-0.050**	-0.029	-0.039	-0.021	-0.049**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
INNER	-0.010	-0.030	0.001	-0.025	0.007	-0.033
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)
MOTHER'S HIGHER ED.	0.201***				0.151***	
	(0.02)				(0.02)	
FATHER'S OCCUPATION						
Skilled Worker			0.064***		0.061***	
			(0.02)		(0.02)	
Managers & Professionals			0.179***		0.159***	
			(0.02)		(0.02)	
CONSTANT	0.753***	1.392***	0.704***	1.256***	0.680***	1.306***
	0.03	(0.09)	(0.03)	(0.08)	(0.03)	(0.07)
Overidentification p-value						0.261
Endogeneity p-value						0.000
F-Test First Stage						64.926
N		5486		5148		5089

Notes: *, **, * means p-value <0.10, p-value <0.05 and p-value <0.010 respectively. MADRID is the reference category for the regions. Variable instrumented: SECONDARY 2. Instrumental Variables: MOTHER'S HIGHER ED. and FATHER'S OCCUPATION. The reference category for the last instrument gathers unskilled workers**

Table 3.18: Instrumental Variable Estimation (D_3) using MOTHER'S HIGHER ED., including FATHER'S OCCUPATION as an Explanatory Variable. Dependent Variable: LN_WAGE.

Variables	Mother's Higher Education			
	(a) First Stage		(b) IV	
HIGHER ED. (D_3)			0.697	***
			(0.15)	
EXPERIENCE	-0.006	**	0.034	***
	(0.00)		(0.00)	
EXPERIENCE^2	-0.000		-0.000	***
	(0.00)		(0.00)	
NORTH WEST	-0.048	*	-0.086	***
	(0.02)		(0.02)	
NORTH	0.023		-0.001	
	(0.02)		(0.02)	
EAST	-0.048	**	-0.033	
	(0.02)		(0.02)	
SOUTH	-0.021		-0.052	**
	(0.02)		(0.02)	
INNER	-0.007		-0.028	
	(0.03)		(0.02)	
MOTHER'S HIGHER ED.	0.150	***		
	(0.02)			
FATHER'S OCCUPATION				
Skilled Worker			-0.004	
			(0.01)	
Managers & Professionals			0.323	
			(0.03)	
CONSTANT	0.666	***	1.425	***
	(0.04)		(0.11)	
N			5089	

Notes: *, **, * means p-value <0.10, p-value <0.05 and p-value <0.010 respectively. MADRID is the reference category for the regions. Variable instrumented: HIGHER ED. The reference category for the last instrument gathers unskilled workers.**

Table 3.19: OLS with Schooling as consecutive education levels. Dependent Variable: LN_WAGE.

Variables	D_1 Coeff.		D_2 Coeff.		D_3 Coeff.	
EDUCATION (D_i)	0.083	***	0.140	***	0.332	***
	(0.01)		(0.01)		(0.01)	
EXPERIENCE	0.011	***	0.020	***	0.031	***
	(0.00)		(0.00)		(0.00)	
EXPERIENCE^2	0.000		0.000	***	-0.000	***
	(0.00)		(0.00)		(0.00)	
NORTH WEST	-0.090	***	-0.119	***	-0.111	***
	(0.03)		(0.02)		(0.02)	
NORTH	0.081	***	0.014		0.002	
	(0.03)		(0.02)		(0.02)	
EAST	-0.025		-0.078	***	-0.056	***
	(0.02)		(0.02)		(0.02)	
SOUTH	-0.097	***	-0.105	***	-0.068	***
	(0.02)		(0.02)		(0.01)	
INNER	-0.030		-0.078	***	-0.043	**
	(0.03)		(0.02)		(0.02)	
CONS	1.738	***	1.746	***	1.729	***
	(0.04)		(0.03)		(0.02)	
R^2	0.117		0.173		0.273	
N	2816		3541		5089	

Notes: *, **, * means p-value <0.10, p-value <0.05 and p-value <0.010 respectively. MADRID is the reference category for the regions.**

Chapter 4 Returns to Numeracy Skills and Levels of Education in four European Countries: A Comparison of Frequentist and Bayesian Methods

4.1 Introduction

Attracting highly skilled immigrants has become a solution to counterbalance aging populations in the majority of European countries. Policies have been developed to achieve a match between the demand and supply of skills in the labour market. A deep understanding of the skills that immigrants bring from the country of origin and how the reward of these skills takes place would help attain this match between the demand and supply of skills.

To study the working population skills and compare them across countries and between natives and immigrants remains a challenge due to the limited availability of data. Examples of previous papers regarding the returns to skills are Bonikowska et al. (2008) and Hanushek et al. (2015). The former analyse if immigrants receive different returns to skills from similar Canadian natives. The latter study the returns to numeracy skills in OECD countries among several subgroups of the population such as immigrants. The way this chapter fills the gap in the literature is by estimating the returns to numeracy skills for different subgroups: natives and immigrants, immigrants depending on the country of origin and immigrants depending on time spent in the host country. In addition to the immigrants' returns to skills, the returns to levels of education are estimated and a Bayesian methodology is applied. Hoogerheide et al. (2012) study the returns to a year of education using a Bayesian approach, but the novelty of this chapter is that, to the best of my knowledge, it is the first work using a Bayesian approach to estimate the returns to levels of education among immigrants from different countries of origin. Moreover, European data is used rather than data from North America or Australia, which has been extensively used in papers related to immigrants' education and skills.

This chapter analyses if immigrants experience a disadvantage when estimating the returns to education and skills in a cross-sectional context. In the four European countries studied, do immigrants from least developed countries receive lower wage returns to numeracy skills/levels of education than immigrants from developed countries? do immigrants who have spent less time in the host country receive lower returns to numeracy skills/education than those who have spent more years in the host country? The level of development of immigrants' country of origin and time spent in the host country could potentially cause statistically significant differences in the wage returns for these groups.

The Bayesian approach to estimate the returns to levels of education is chosen for its theoretical and practical advantages in comparison with the frequentist approach, broadly employed. Following the Bayesian estimation approach, rather than estimating a 'true fixed parameter' of the rate of return to education, one could treat the returns to education as a random parameter with a probability distribution. Having a probability distribution in the population allows creating a statistical model for that parameter offering possible theoretical values that the parameter of interest/coefficient could take. Bayesian analysis allows uncertainty in the model manifested through the prior distribution of the parameter; also, a posterior distribution of the parameter can be retrieved to do inference about the parameter/coefficient of interest.

Although in general working immigrants decide to migrate to pursue better labour market outcomes than in their home country, immigrants could experience hardship in the labour market upon arrival. This could be caused by obtaining lower returns to their skills than natives. On the one hand, immigrants could be less productive than natives or, on the other employers could manifest a preference for native workers. Factors such as immigrant rights given upon arrival to the host country, visas, working permits or integration policies, could shape the skill mix of immigrants applying.

Education, in the past, was used by researchers as a proxy for workers' skills in the labour market. Nowadays, skill measurements allow differentiating between education and particular cognitive skills that the worker offers the employers. Although educational attainment could be related or linked to skills, the fact that an individual attains formal education does not imply that he will attain skills that will help when trying to enter the labour market. Relating this to immigrants, it remains an issue of debate if more restrictive/selective immigration policies benefit the country in terms of attracting a more educated and skilled pool of workers. Shields and Price (2001) argue that comparable educational attainment and work experience do not automatically guarantee the same chance of employment for natives and for immigrants.

Regarding immigrants' labour outcomes in Europe and according to the OECD(2014), large disparities in employment outcomes arise by country of origin. For example, divergence in employment rates between natives and immigrants are more present among highly educated workers than less educated ones. Country of origin is a relevant variable when analysing differences in labour outcomes between natives and immigrants.

This chapter is structured as follows: Section 2 presents a literature review regarding returns to education and skills and describes the integration approaches followed by the European countries chosen for the analysis; Section 3 explains the data set used for the analysis and the explanatory variables for the earnings equation; Section 4 describes the methodology employed in the estimation, offering detailed information regarding the Bayesian estimation approach; Section 5 reports the results obtained using the different estimation methods, plus a discussion of the econometric techniques; Section 6 concludes the chapter, including limitations and offering some possible extensions.

4.2 Literature Review, Integration Models and Research Question

There is a vast literature regarding immigrant educational mismatch and over-education, but there are fewer papers about the impact that skills have on wages. The literature review is divided into four subsections: education, skills, integration models and research questions.

4.2.1 Education

When immigrants enter the host labour market, an asymmetric information situation arises. According to Chiswick and Miller (2009), employers do not want to take the risk of hiring an immigrant because of the lack of information regarding his/her work experience and because of not knowing how to take account of their foreign qualifications in comparison with a native home-trained or educated individual. This phenomenon contributes to the difficult labour market integration that immigrants suffer. This situation is exacerbated at the moment of arrival to the host country and due to this, job mismatch across immigrants could occur. Immigrants' first jobs do not match their skills and qualifications. Some immigrants stay for a long period in their first job, while others progress, moving to a better job than the starting one. Gaining years of work experience in the host country and studying to obtain extra credentials in the country of destiny helps narrow the initial information gap.

Chiswick (1978) came up with the ability/motivation hypothesis: immigrants from developed countries with low educational attainment succeed less in the labour market than immigrants from less developed countries. This author used data from the 1970 US Census and found that the return to an extra year of schooling for the native born was 7.2% and 5.7% for the immigrants. Using data from the 2000 US Census, Chiswick and Miller (2007) showed that this pattern occurs again three decades later with native returns of 10.6% and immigrant returns of 5.2%. Other studies that support lower returns for immigrants than for natives are Beggs and Chapman (1988), Shields and Price (1998) and Dustmann (1993). Beggs and Chapman (1988) found that differences in wage performance exist across Australian migrants. The authors take into account changes in the labour market quality between immigrant cohorts. Shields and Price (1998) study the English labour market and use the quarterly Labour Force Survey to estimate earnings for native born and foreign born males in England. They found that most immigrant groups have lower returns to schooling than native-born whites. They argue that education obtained in the UK by immigrants is more valuable than education obtained abroad. According to Damas de Matos (2014), data from the PIAAC (Programme for the International Assessment of Adult Competencies) reveals that most immigrants completed their education before moving to the host country.

Chiswick and Miller (2007) offer three possible explanations for why a lower partial effect of schooling among the foreign-born arises: overall, in the educational distribution of immigrants, those at the bottom (i.e. less educated) are, predominantly, the ones who migrate; there is a deficient international transferability of skills and this, contrary to the first explanation, affects highly educated immigrants. Skills are difficult to be transferred due to differences in the language or barriers such as permits to legally reside in a host country. The third and last explanation is a degree of discrimination[1] or favoritism for native workers. These possible explanations/scenarios can vary depending on the level of development of the country of origin.

According to OECD (2004), educational disadvantage suffered by immigrants differs depending on the country in which they studied. As OECD (2004) mentions, countries that are destinations for highly skilled immigrants are places where educational disadvantage is low. In these countries, immigrant composition does not greatly differ to natives' composition. Differences between immigrants and natives' composition could

[1] Immigrants earning less than native workers with the same characteristics.

be a factor to increase immigrant disadvantage, as Castles (1993) argued.

Bratsberg and Terrell (2002) estimate returns to education in the US labour market by country of origin and argue that immigrants from Northern Europe receive high returns, while immigrants from Central America (Costa Rica, Dominican Republic, El Salvador, Guatemala, Honduras, Panama and Trinidad & Tobago) receive lower returns. Supporting Card and Krueger (1994), Bratsberg and Terrell (2002) conclude that decreasing the number of pupils per teacher increases the future wage of the immigrant. Correlation analysis suggests that individuals educated in wealthier nations earn higher returns in the US labour market. The returns to education are higher for immigrants from English-speaking countries in comparison to those with the same characteristics from non-English-speaking countries. If returns to education of immigrants from less developed countries are lower, this could be due to poor school quality in these developing countries. that cause a reduction of transferability of schooling or skills.

Many authors such as Chiswick (1978), Borjas (1991), Portes (1996) and Telles and Ortiz (2008) published papers about the disadvantages that immigrants suffer. For example, Friedberg (2000) argues that the outflow of human capital produced because of emigrating from the country of origin place immigrants in a disadvantaged situation. An example of this occurs when an immigrant is not familiar with the language of the country where he/she is migrating to. Immigrants encounter a situation where they have the opportunity to amplify and enlarge their human capital upon arrival.

4.2.2 Skills

As Green and Riddell (2003) point out, a technological change has recently occurred in developed countries. These new technologies favor more skilled workers over less skilled ones (Katz and Autor (1999)) which is why skills are becoming more important nowadays. Each country, depending on its own labour market, demands certain skills, and immigrants need time to acquire these required skills. If immigrants become highly skilled, they could mitigate shortages of skilled labour in important sectors of the economy. Unfortunately, immigrants often suffer from the imperfect transferability of human capital across countries (Chiswick (1978)).

We need to acknowledge that the empirical evidence regarding immigrant skills is limited. One of the reasons for this is the small number of databases available that measure cognitive skills. Three databases are well known for gathering information regarding skills: International Adult Literacy Survey (IALS), administered between 1994 and 1998, Adult Literacy and Life Skills Survey (ALL), administered between 2003 and 2007, and the Programme for the International Assessment of Adult Competencies (PIAAC) administered in 2012.

There are several strands of the skills literature. Papers such as Antecol et al. (2003), Bleakley and Chin (2004), Barrett (2012) and Berman et al. (2003) focus on language skills. Antecol et al. (2003) use immigrants' data from the 1991 Australian and Canadian Census and data from the US 1990 Census to argue that Australian and Canadian immigrants have higher level of English fluency than US immigrants. Once controlling for language fluency skills, the immigrant-native income differentials are 2.4, 7.5 and 2.7

percent for Australian, Canadian and US immigrants, respectively.

Barrett (2012) concentrates on Australian immigrants to estimate the returns to skills using the Australian IALS dataset. The author found that the average return to an extra year of education was 6.2 percent, of which almost one third could be attributed to cognitive skills. He also focuses on the native-immigrant wage gap and argues that differences in skills accounted for almost half of the negative wage-gap for immigrants with Non-English speaking background. Bleakley and Chin (2004) study the effect of language skills on earnings. Using the 1990 US census data, the estimation sample comprises individuals that migrate to the US as children. Their results suggest that, after dividing the sample according to three different levels of language proficiency, compared to a person that speaks the native language poorly, a person who speaks better earns 33 percent more and a person who speaks even better earns 67 percent more.

Berman et al. (2003) estimate the return to host-language fluency for Russian immigrants in Israel. This return was statistically not different from zero for immigrants working in low-skills occupations, opposite to what was found for immigrants in high-skill occupations.

Another strand of the literature examines the returns to literacy skills. Bonikowska et al. (2008) and Ferrer et al. (2006) find that there are no significant differences between natives and immigrants returns to skills. The first paper, using Canadian data from IALS, find no evidence that immigrants receive a lower return to literacy skills than native-born workers, and conclude that discrimination cannot be used to explain immigrant-native earnings differentials. Along with Bonikowska et al. (2008), Ferrer et al. (2006), using Canadian data also, conclude that the returns to literacy skills for immigrants and natives are very similar. As immigrant returns to skills are not lower than native returns, they argue as well that, in this context, there is no evidence of discrimination defined as equally productive workers getting paid unequally.

Apart from studying Canada, Kahn (2004) also uses the IALS data from New Zealand, Switzerland and the US. He found that in all these four countries, immigrants scored lower than natives in the IALS tests. Although returns to skills were not estimated, the immigrants skills distribution was investigated and evidence showed that this distribution was bimodal with one part of the distribution similar to the natives' distribution of skills and the other part concentrated at low skill levels. Bratsberg et al. (2013) extend the analysis further and use Canada, US and Norway data from ALL to examine the association between literacy proficiency and employment for immigrants and natives. They conclude that for natives, this association is very similar in the four countries and a strong correlation is found between literacy proficiency and employment, while for immigrants there is a large employment penalty in Norway and a low employment penalty in North America associated with low levels of literacy. Clarke and Skuterud (2013) uses as well the ALL database to compare immigrants' test scores obtained in Canada, Australia and US. These authors estimate the returns to literacy skills and found that there is no evidence of immigrant skills under-utilization. In fact, in Canada and Australia the returns to skills are not statistically different from the natives. Rather, US immigrants with a mother tongue that is not English or Spanish obtain a higher return than the rest.

Although a greater number of studies use data from ALL or IALS, Smith and Fernandez (2017) use the PIAAC data and reveal that wage differentials between immigrants and natives disappear once education and skills are taken into account in the US. In Canada, persistence in the wage gap was found in nearly all the occupations after controlling for education and skills, suggesting that the wage premium between natives and better skilled and educated immigrants is not equal.

4.2.3 Integration Models

Depending on the country's labour market characteristics, immigrants could face different situations, some more serious than others, such as having difficulties to be employed or suffering from mismatch/over-education in the labour market. Depending on the different countries of arrival, some cultures will favour immigrants' integration into the society of destination, while in other countries the integration is not smooth. For example, to stereotype immigrants by their country of origin could cause difficulties in the integration process. Not all immigrants are equal and in the process of immigration there are several reasons why individuals choose to migrate from their countries of origin, including factors such as demographic pressure, high unemployment rate or economic inequalities between regions.

An increase in migration in the last decade made host countries review their policies regarding cultural diversity management, and obtaining nationality or citizenship. For example, in Denmark, the government facilitates obtaining permanent residence if the immigrants are integrated in the host country (Mahnke and Nymark (2010)). Denmark and Canada are examples of countries that use the point system as part of their immigration policy to control the the acceptance of immigrants. As Smith and Fernandez (2017) point out, there are countries such as Canada that focus on attracting a pool of highly skilled immigrants, in contrast to the US that prioritise family reunification.

If we review the integration models for the countries studied, following Retortillo et al. (2006), we find that France follows an assimilationist model, while Denmark and the United Kingdom follow a multicultural model. In Spain, there is an absence of a model. These different integration models are reviewed in detail.

The assimilationist model implies that an individual either remains as a foreign citizen or becomes part of the new country's society. In the case of France this model applies and the board for the integration, dissolved in 2012, argued that integration should entail unity and not the creation of minorities (Conseil (1993)). On the other hand, as Retortillo et al. (2006) argue, the multiculturalism model offers the immigrant the option of retaining their particularities in a system that acknowledges the differences between immigrants who are not willing to adopt the customs and traditions of natives.

In 2009, in France, the Immigration and Integration office (OFFII) was created to unite all the actions and activities of the government to receive immigrants and work on their integration. One of the activities was to provide language courses and cultural lessons to let immigrants know more about the Republic. The Government made an effort to fight against discrimination and provide immigrants with the same opportunities as the natives. In the same year, a project was presented to deal with illegal immigrant workers. Regarding assimilation, in 2011, a hardening of the assimilation requirements happened. Since then, immigrants who were candidates for naturalisation have been required to demonstrate and justify their integration in French society. In order to prove this, immigrants must know the language, French culture and French history. What this model looks for is to convert the immigrant into a French citizen. As a consequence, this model is not compatible with the maintenance of the culture of origin and rejects the creation of groups and communities to preserve their culture of origin. With this model, the Government tries to eliminate any differences between immigrants and natives and create an homogeneous society through the immigrants adopting the republican values from the host country. Some of these values are secularism, respect for human rights and pride in being French. This model arises from an attempt to isolate two streams, radical attitudes such as xenophobia or ultra-nationalism of the conservative parties and, on the contrary, attitudes adopted by the radical left-wing parties. Therefore, what the assimilationist model pretends is that the immigrants abandon their own identity, any link to the society and culture of origin, and accept the Republic's principles (Retortillo et al. (2006)).

On the other hand, according to Retortillo et al. (2006), the British and Danish model for integration, multiculturalism, is different from the French model. Multiculturalism is a more relaxed model in terms of how the behaviour of the immigrants should be. The government accepts that immigrants keep their links with their country of origin and preserve their culture and their social networks with other immigrants. This is a very tolerant model in which different ethnic and religious groups and the nationals coexist. Each individual is encouraged to preserve their own identity. This encourages the creation of schools, churches, associations and other groups to connect and socialise with people from the same group. There are no strict norms regarding the way integration should happen, but these groups must respect some rights and obligations to participate in the host country lifestyle.

In Denmark, immigrants must show that they would like to contribute to the development of the country and Danish society. The Danish government has always tried to promote integration. Around 2010, there was a large arrival of immigrants from occidental countries (Mahnke and Nymark (2010)). These immigrants were highly educated, which helped them learn the language of the host country fast, adapt to Danish culture, rapidly find a job, and respect the Danish society. The conservative and liberal parties agreed to fight against the creation of several societies coexisting at the same time and wanted to eliminate places that were concentrations of socially disadvantaged people. Additionally, other measurements such as keeping track of public expenditure devoted to immigration were promoted. Although ministers of integration such as Søren Pind, did not favour multiculturalism as a model for integration, this model is what prevails.

Data published in MigrationWatchUK[2] shows that immigration has remarkably increased in the UK over the last 20 years. The accumulated stock of net migration from 1997 up to the end of September 2016 was 4,341,000 immigrants in total. In particular, according to Vargas-Silva and Markaki (2011) the United Kingdom experienced an immigration growth from 2009 (567,000 immigrants) to 2010 (591,000 immigrants). ONS (2015) shows that the majority of immigrants arriving to the UK in 2010 and 2011 were from India, Pakistan, Poland and Ireland. The Government established limits for the number of immigrants that the country could receive. Highly skilled or highly educated immigrants have more opportunities to get a working visa in the UK than immigrants with low education levels or non skilled immigrants. Regarding immigrants' integration, David Cameron, the British prime minister in 2011, argued that under their multiculturalism model different cultures have separated lives and they remain apart from other communities and apart from the mainstream, which seems not to be considered beneficial for the society in general.

---

[2] https://www.migrationwatchuk.org/statistics-net-migration-statistics

Spain is characterised by the absence of a specific integration model. According to Reher et al. (2011), this country experienced the arrival of a vast number of immigrants since 1998. Before this year, there was more emigration than immigration. While in 1998 immigrants were 3% of the total population, this figure has increased to 14% in 2010. As Reher et al. (2011) argue, 30% (41,370 individuals) of the total foreign-born population arriving to Spain in 2010 came from only three countries: Romania, Morocco and Ecuador.

One of the causes for the lack of a defined integration model such as the ones in France, Denmark or the UK is the large number of immigrants arriving in a short period of time. It can be said that the Spanish integration model has some common characteristics with the aforementioned integration models. Regarding integration, there are organisations supported by the Ministry of Labour and Social Affairs to encourage immigrants who are legally allowed to stay in the country to participate and be integrated in Spanish society.

4.2.4 Research Questions

This chapter presents wage equations controlling for education level and numeracy skills across natives and immigrants, where the last group is separated into two sub-groups depending on the level of development of the country of origin. Two estimation methods, frequentist and a Bayesian, are presented.

As for the research questions addressed, we investigate whether the returns to levels of education and numeracy skills for immigrants and natives are different by introducing interaction variables in the wage equation. We are also interested to see if there are significant differences in the returns for the two immigrant groups.

Friedberg (2000) argues that the schooling attained by immigrants from developed countries might be more valued in the host labour market than the schooling attained by immigrants from developing countries. The same reasoning could be followed to explain the differences in the returns to skills. The author argues that the more similar the origin and destination countries are in characteristics such as economic development, industrial and occupational structures or labour market outcomes, the more likely human capital variables such as education and experience will be highly valued. Borjas (1991) argues that in general, immigrants coming from industralised economies are more skilled and more successful in the host country labour market than those coming from developing countries.

Bertrand and Mullainathan (2004), Jacquemet and Yannelis (2012) and Oreopoulos (2011) present evidence of the important role that country of origin plays for immigrants. Bertrand and Mullainathan (2004) show that in the US labour market, individuals with white-sounding names are more likely to be called back for an interview than those with African-American names. On the same line, Jacquemet and Yannelis (2012) show evidence of ethnic homophily when résumés with Anglo-Saxon names received a higher number of callbacks than the same résumés from either non Anglo-Saxon, Foreign or African-American names. Using Canadian data, Oreopoulos (2011) presents evidence of discrimination when Canadian-born individuals with English-sounding names are more likely to be called back for an interview than foreign-born individuals. Occasionally employers follow stereotypes when selecting employees leading to Discrimination. This can be considered a relevant mechanism to explain poor immigrant outcomes in the labour market.

According to Chiswick (1978), another possible mechanism to explain immigrant lower returns to skills/levels of education lies in the imperfect transferability of human capital across countries. Each host country, depending on its own labour market, demands certain skills. Immigrants need time to acquire those skills. Over time, some immigrants can improve their language proficiency and become more productive in the labour market; however, at the moment of arrival, employers may not want to take the risk to hire them and, unfortunately, immigrants end up in worse job positions than expected and receive lower returns than natives. On this account, it can be expected that human capital transferability will be a positive function of time spent in the host country.

Finally a third possible mechanism that could place immigrants in a disadvantaged position in comparison with natives is the poor school quality in developing countries. Sweetman (2004) argues that a considerable portion of the return to education is associated with educational quality. Given the lower quality of educational systems in developing economies, equivalent qualifications across different countries do not necessarily imply same abilities/competences. A plausible proxy for education quality is the amount of pupils per teacher. Bratsberg and Terrell (2002) show that the less the number of pupils per teacher, the higher the future wage of the immigrant.

4.3 Data Set and Variables

The data used for this chapter comes from the Program for the International Assessment of Adult Competencies (PIAAC). Data was collected by the OECD in 2011-12 and was released in October 2013. This dataset gathers information on adult skills and educational attainment of natives and immigrants in several European countries.

The use of PIAAC has several advantages over the use of other databases that gather information on skills, such as the IALS. For instance, the PIAAC collects more recent information and sample sizes are larger than in IALS. PIAAC gathers nationally representative samples of workers and at the same time facilitates international comparisons since the variables and methodology applied to each country are the same. For the PIAAC, individuals aged 16-65 were interviewed in the official language of the country.

As mentioned before, four countries are taken into account in this chapter: Denmark, France, Spain, and the United Kingdom. These countries are chosen because of two reasons, public data availability and different integration approaches. For some OECD countries such as the US, the country of birth or the education information is not provided in the public files. Socio-economic background variables[3] for natives and immigrants were provided in the public files available from the OECD web database for the four countries selected.

We need to acknowledge that the qualification system differs between countries, but as they are part of the OECD the ISCED system can be used to make the qualifications comparable across countries. The sample is formed by full-time employees defined as those working more than 29 hours per week. Students and self-employed individuals are excluded from the analysis because as Clarke and Skuterud (2013) point out, cognitive skills play a weaker role in determining the labour market outcomes of the

[3] Although there are papers in the literature such as Carnoy (1996) and Psacharopoulos (1993) that argue that ethnicity is an important control variable when estimating the returns to education, this variable is only available in the PIAAC for the US, not for the countries studied here. Language is a variable that might impact on wages (Patrinos et al. (1994), Bleakley and Chin (2004)). It has been tried and appears to be non-statistically significant in the European countries except the UK where it is statistically significant at the 10% level.

self-employed. Only first-generation immigrants (those born in a different country than the country of the survey and whose parents are both from another country than that of the survey) are examined. Second-generation immigrants are considered a different group and are therefore omitted form the estimation sample (rather than coded as natives) to avoid misinterpretation.

Dependent variable

- Log (hourly wages): log gross pay per hour. This variable is created depending on the number of worked hours reported and what the individual earns. Individuals can choose to report their wage per hour, day, week, fortnight, month or year. For all the individuals that have not reported the wage per hour, a transformation is needed. In order to obtain the gross pay per hour, the following factor $f = \frac{40}{\text{reported worked hours per week}}$ is applied. This factor indicates the fraction of worked hours per week assuming that a normal working week has 40 hours. The next step is to divide the gross pay per person by 8/40/80/173/2080 depending if the individual has reported daily, weekly, fortnightly, monthly or yearly gross wages. To proceed, the quantity obtained is multiplied by f. The last step is to calculate the natural logarithm of this quantity to obtain the logarithm of gross pay per hour. The variable is measured in the local currency in each country (Euros in France and Spain, Sterling Pounds in United Kingdom, and Danish Krone in Denmark). For each country, once the hourly wages variable was created, in order to avoid the influence of outliers, the bottom and top 1 percent of the earnings distribution was eliminated from the sample[4].

[4] Regressions were obtained keeping and not keeping the outliers. If outliers are left in the estimation sample

Control Variables:

- Experience and Experience Squared: defined as years of paid work during lifetime.
- Female: takes the value of 1 for women and the reference category is men.
- Education Levels: Medium Level of Education and Tertiary Education. These dummies indicate the corresponding highest educational level attained. The reference category applies to individuals with low educational attainment. The UNESCO's International Standard Classification of Education (ISCED) is used and the PIAAC grouping recommendation of recoding was employed. The initial ISCED levels were reduced into the following three broad categories: Low Education (No formal qualification or below ISCED 1 or ISCED 1, 2 and 3C shorter than two years), Medium Level of Education (ISCED 3C longer than two years or two years, 3A-B and 4) and Tertiary Education (ISCED 5 and 6). The first category corresponds to primary or less than primary education, the second category to secondary school diploma and the third dummy corresponds to Higher/Tertiary Education.
- Numeracy Skills: in PIAAC defined as "the ability to access, use, interpret and communicate mathematical information and ideas, in order to engage in and manage the mathematical demands of a range of situations in adult life" (OECD, 2012b). This variable was measured on a numerical 0-500 point scale and was

we can observe two changes. For Denmark, the coefficient on immigrants from the OECD and with 15 or more years in the country becomes statistically significant at the 10% significance level. For the UK, the interaction coefficient for numeracy skills and immigrants from a Non-OECD country and with less than 15 years in the host country is no longer statistically significant. The rest of the coefficients of interest do not vary in size or sign. Following two papers that use the PIAAC, Hanushek et al. (2015) and Smith and Fernandez (2017), only results not taking into account the outliers are reported to avoid bias in the coefficients caused by the inclusion of very high/low income earners.

standarised for each country. There are a total of 56 items (questions) in the numeracy assessment. According to the OECD, four areas of mathematical content are covered: quantity (e.g. calculate prices or growth rates) and number (e.g. four main operations "$+ - \times \div$"); dimension and shape (e.g. calculate perimeters or areas); pattern, relationships and change; and data and chance (e.g. ideas related to variability, sampling or statistical topics). Four contexts are used to plan the questions: work-related; personal; society and community; education and training. Four cognitive strategies are covered in the test: identify, locate or access; act upon or use; interpret, evaluate/analyse; and communicate. There is a numeracy proficiency scale that divides the total 56 numeracy tasks into five levels of difficulty. More information regarding the numeracy assessment can be found in OECD (2013) with sample questions from the numerical assessment.

- Country of Origin Dummies: two dummy variables reflecting the development of the country of origin. The first dummy takes the value of 1 for an immigrant that belongs to an OECD country and the second dummy variable takes the value of 1 for an immigrant that belongs to a non-OECD country. The reference category corresponds to the natives (those born in the country of the survey)[5].

- Region dummies are included for each of the four European countries where the reference categories are Hovedstaden, Ile De France, Madrid, and London for Denmark, France, Spain and the UK, respectively. These regions are chosen to be reference category because they contain the capital of each country.

---

[5] Since EU migrants have freedom of movement, it would be sensible to consider them as a separate category from other OECD migrants. However, since the immigrants sample size is already small, to split the OECD into two categories is not possible in one of the countries (Spain) because of the lack of observations in the Non EU category; in France there are only 3 immigrants from an Non EU OECD countries and the equivalent figure is 9 for the UK

- Industry Sector dummies: four dummy variables for the following groups: Agriculture, Fishing and Forestry; Industry (Mining, Quarrying, Manufacturing, Electricity, Gas, Water,Waste Management); Construction and Service Low Level (Trade, Repair of Motor Vehicles, Transport, Hostelry, Administration, Home-makers). The reference category is Service High Level (Communications, Finance, Banking, Insurance, Real Estate, Public Sector, Defense, Professionals, Scientific and Technical activities, Education, Nursing, Art and Extra Territorial organisations).

- Economic Sector: Private (reference category), public sector or Non-Profitable organisations /Charities.

- Manages or Supervises: dummy that takes the value of 1 if the individual manages or supervises other employees in the workplace.

- Type of Job contract: five dummy variables representing a fixed-term contract, a temporary employment agency contract, an apprenticeship or other training scheme, no job contract, or other type of contract. The reference category applies to individuals with an indefinite contract.

- Time living in the country of survey: two dummy variables, the first one that takes the value of 1 if the immigrant has been living in the country for less than 15 years, and the second if the immigrant has resided in the country for 15 or more years. This variable is used as a proxy for immigrant adaptation/integration to the host country. The 15 years threshold was chosen because it divides the immigrants (7.9% of the total sample) in two groups (4.4% and 3.5%) of almost the same size. On top of that, this threshold is used to the extent that it guarantees that the permanent residence or nationality could have been obtained in the four European countries.

Table 4.1 shows the descriptive statistics based on the sample used in each corresponding European country. It can be seen that the sample size varies across European countries ranging from Spain with 1695 observations to Denmark with 3107 observations. One of the causes for observing variation in the logarithm of earnings across countries is the use of different currencies. The proportion of female individuals in the four countries is approximately 40%. Differences in years of experience are present because there are differences in the average age of the sample in each country. Age is not used as a variable in the analysis to avoid multicollinearity with experience.

We can see that the percentage of immigrants in all four countries is approximately 10 percent of the total sample (counting together immigrants from the OECD and not from the OECD). Regarding the highest education level attained, we can see differences across countries. Between 42% and 45% of the individuals have attained tertiary education in Denmark, Spain and the UK. This figure is lower in France, with only around 35% of the individuals having attained higher education. Spain has the highest percent of people with lower or no education at all.

Regarding the numeracy skills score, we can see that on average, out of the four European countries studied, Denmark has the highest numeracy score, followed by the UK. Spain has the worst average numeracy score of the four countries. The distribution of the sample across industry sector and economic sector is very similar in the four countries. This is expected because they are four developed countries. In all of them, the majority of individuals work in the service sector and the private sector. If we look at the type of contract variable, we can see that in all the European countries studied, the majority of individuals have an indefinite contract. One possible explanation is that the majority of individuals are in their 40s and they have several years of work experience. Individuals could have spent several years with the same company or in the same job position and the company could reward this commitment with an indefinite contract. The second largest group regarding type of contract are individuals with a fixed-term contract, this type of contract being more common in Spain than in other European countries according to Table 4.1. There are more individuals with temporary contracts in Denmark and France than in Spain and United Kingdom. As we can see, other types of contracts are not very common in the European countries studied.

The last variable that enters the earnings equation is the time (years) spent in the host country. This variable only applies for immigrants. We can observe differences across countries. While in Denmark, Spain and the United Kingdom the majority of immigrants recently arrived in the host country, in France, we observe that the majority of immigrants represented have spent more than 15 years in the host country.

Figure 4.1: Wage Density Plots

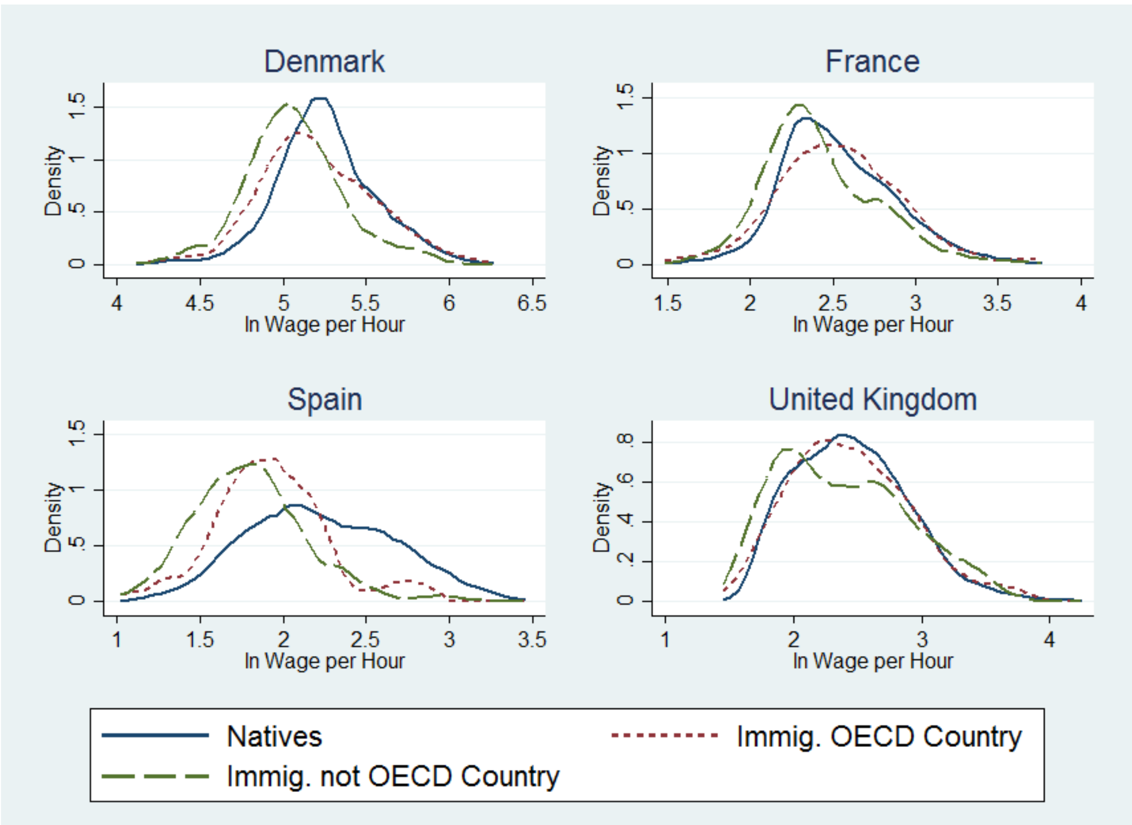
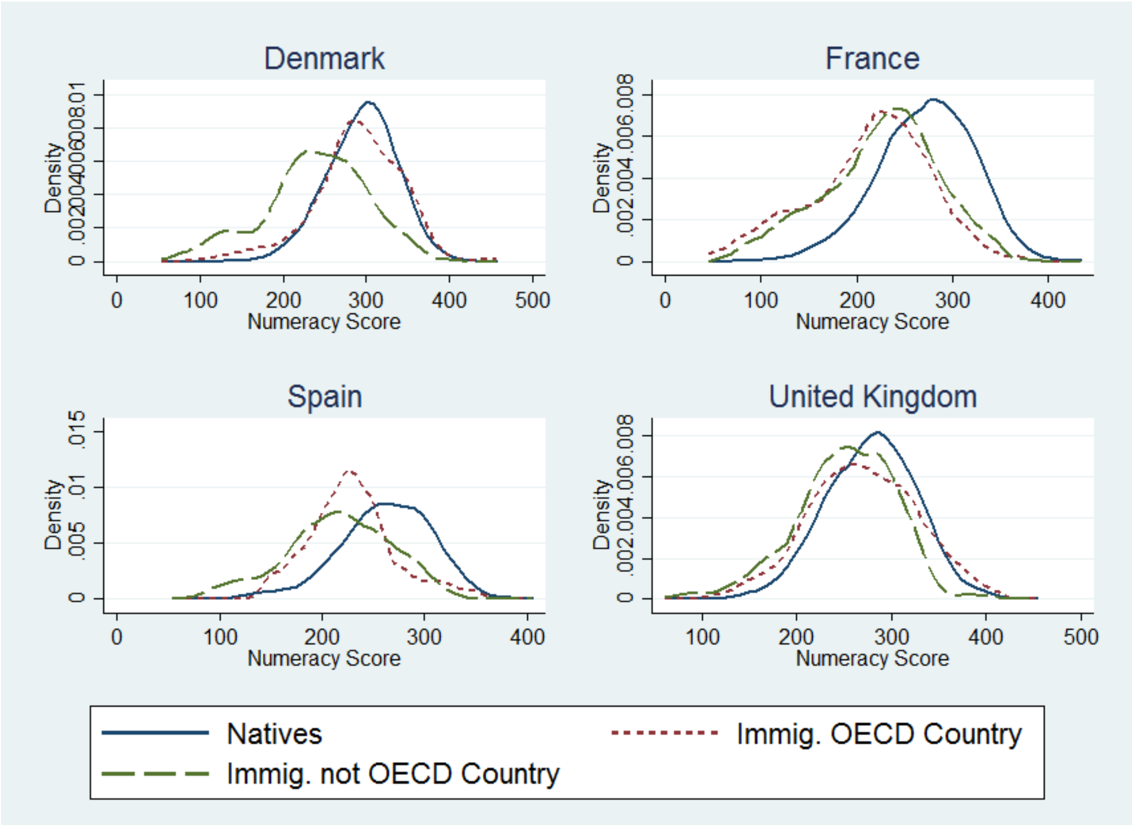


Figure4.1 shows the density plots for the distribution of wages for the three different groups in each country. It can be seen that in Denmark and Spain, the natives' curves are shifted to the right-hand side, therefore, in general, natives earn higher wages in comparison to the immigrants. In Denmark, if the logarithm wage per hour is above 5.5, there are no differences between immigrants from OECD countries and natives earning above that wage, while there are fewer immigrants coming from developing countries in that high wage band. In Denmark, although the curve for natives is slightly displaced to the right, the shape of the wage distributions is similar. On the other hand, in France, we can see that the shape of the curve for natives resembles the curve corresponding to immigrants from developing countries. The curve representing wages of immigrants coming from a OECD country is more humpback and symmetrical than the other two curves. This indicates that there is more income equality for immigrants from OECD countries. In Spain, we can see that the wage distribution for the two immigrant groups are more similar to each other than the natives' curve. The natives' curve is displaced to the right, indicating that the majority of natives earn higher wages than the immigrants in this country. In the United Kingdom, the curves are very similar in shape, but it can be seen that the natives' curve resembles the immigrant curve from OECD countries. In this country, the displacement of the curve occurs for immigrants from developing countries. It can be seen that the displacement of the natives' curve with respect to the immigrant curves is minimal in comparison to other countries such as Spain.

Regarding the numeracy skills density plot, we can observe a pattern. The natives' curves are always shifted to the right of the immigrants' curves. In general, immigrants score fewer points in numeracy than the natives. In France and the United Kingdom, there is not much of a difference between the two immigrant groups' curves, while in Denmark

Figure 4.2: Numeracy Skills Score Density Plots



and Spain, immigrants from developing countries score on average less points and this explains why the curve is displaced to the left.

We now examine the characteristics in the four countries analysed across three groups: natives, immigrants from an OECD country, and immigrants from a non-OECD country; some patterns arise and some differences are highlighted. Regarding average wages and average numeracy scores, Denmark, Spain, and the UK share common patterns not observed in France. Tables 4.2, 4.4 and 4.5 show that natives, on average, earn the highest salary of the three groups. Several studies such as, Chiswick (1978), Borjas (1995) and Aydemir and Skuterud (2005) observe that immigrants earn lower wages than the natives with equal characteristics such as education and experience. This could be due to several reasons such as natives not having to face cultural adaption to the host country or perfect knowledge of language.

It can be seen that immigrants from countries belonging to the OECD have lower salaries than the natives, but these differences between average salaries are not very large. Immigrants from developing countries have a lower average salary in the four European countries, but it can be seen that the wage gap is narrow as well. As an exception, Table 4.3 shows that on average, immigrants from a country of the OECD have a slightly higher wage per hour than a native in France.

Regarding average numeracy skills, the same pattern arises: natives score the highest out of the three groups in the four countries. Immigrants from less developed countries have on average the lowest average numeracy score in all the European countries except France. Surprisingly, in France, immigrants from developed countries have the highest percentage of individuals attaining low education level as the highest level of education achieved.

Regarding the highest level of education attained, in Denmark and the UK the modal group across the three categories is tertiary education. These two countries have more restrictive policies regarding the level of education of the individuals applying for a working visa, and as the descriptive statistics show, these countries attract highly skilled immigrants. Regarding immigrants' attainment, Borjas (1991) argues that differences in the foreign-born average educational attainment can be mostly explained by different national-origin mix of immigrants admitted into a country. If the point system is used, an increase in the average educational level of foreign-born could be obtained by altering the different countries of origin of the immigrants admitted into the country.

Both immigrant groups have higher average numeracy scores in Denmark and the UK than in France and Spain. In France and Spain, the modal group of immigrants, irrespective of whether they are from a developed or developing country, have attained a lower level of education or no form of education. Comparing all natives, in Denmark, Spain, and the UK, more than 40% have attained higher education, while in France, the modal group of natives have only attained a medium education level. To sum up, in general and taking into account average wages and average numeracy skills, immigrants from developed and developing countries are worse off than natives.

4.4 Methodology

First, the frequentist approach is introduced and Ordinary Least Squares used as the estimation method. Following Koop (2003), we have a $n \times 1$ vector with the dependent variable, $y_i = log(\text{gross pay per hour})$ for observations $i = 1...n$. We have a size $n \times k$ matrix of explanatory variables, X. The first column of this matrix is a vector of 1s to allow the model to have an intercept. The classic OLS equation is

$$y_i = \beta_1 + \beta_2 x_{i1} + \beta_3 x_{i3} + ... + \beta_k x_{ik} + \varepsilon_i \tag{4.1}$$

Using the following elements:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{bmatrix} \qquad X_{m \times k} = \begin{pmatrix} 1 & x_{12} & \cdots & x_{1k} \\ 1 & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n2} & \cdots & x_{nk} \end{pmatrix} \qquad \beta_{1 \times k} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ \beta_k \end{bmatrix} \qquad \varepsilon_{1 \times n} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ \varepsilon_n \end{bmatrix}$$

we can write equation 4.1 in matrix form:

$$y = X\beta + \varepsilon$$

where some assumptions need to be met: ε is a multivariate Normally distributed error term with mean 0_n and variance $h^{-1}I_n$[6]. X is independent of ε, $var(\varepsilon) = h^{-1}$ and $cov(\varepsilon_i, \varepsilon_j) = 0$ for $i \neq j$. The **likelihood** function[7], $p(y|\beta,h)$ is multivariate Normally distributed and can be written as

$$p(y|\beta,h) = \frac{h^{\frac{n}{2}}}{(2\pi)^{\frac{n}{2}}}\left\{\exp\left[-\frac{h}{2}(y-X\beta)'(y-X\beta)\right]\right\} \tag{4.2}$$

where, in the OLS framework v, β_{OLS} and s^2 are the degrees of freedom, estimators for β and standard error, respectively; the three mentioned parameters are defined in the following three equations

$$v = n - k \tag{4.3}$$

$$\beta_{OLS} = (X'X)^{-1}X'y \tag{4.4}$$

$$s^2 = \frac{(y - X\beta_{OLS})'(y - X\beta_{OLS})}{v} \tag{4.5}$$

Equation 4.6 defines the sum of squared errors of prediction (SSE) in the OLS framework. The SSE is later used in the Bayesian approach to enter the likelihood function.

$$SSE = (y - X\beta_{OLS})'(y - X\beta_{OLS}) \tag{4.6}$$

---

[6] To simplify the mathematical calculations we tend to use the error precision, $h = \sigma^{-2}$, instead of the variance. This will simplify the equations used for the Bayesian estimation method.

[7] If one of the assumptions of the model is that the error terms are normally distributed this implies that the likelihood is normally distributed as well.

Bayesian

Using Bayes' rule of probability, if A and B are two random variables and $P(\cdot|\cdot)$ and $P(\cdot)$ are the conditional and marginal probabilities, respectively. We can write

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{4.7}$$

In the frequentist approach, we are interested in the effect that certain explanatory variables, X, have on a dependent variable, y. As mentioned before, the latter is a vector of data and β contains the parameters of interest. The data will be used to learn about the parameters. If we replace A and B in equation 4.7 by β and y, we can write

$$P(\beta|y) = \frac{P(y|\beta)P(\beta)}{P(y)} \tag{4.8}$$

where β is going to be a random variable with a probability distribution, $P(\beta|y)$ is the posterior density, $P(y|\beta)$ is the likelihood function described before and $P(\beta)$ the prior density. As we are not interested in the prior distribution of the data, $P(y)$, because of its independence from β, we can write equation 4.8 as

$$P(\beta|y) \propto P(y|\beta)P(\beta) \tag{4.9}$$

The fundamental interest relies is in the posterior distribution that combines the prior information (what the researcher believes about the initial values of the parameters before seeing the data) with the information given by the data. It can be seen that the prior distribution is independent from the data. In order to obtain the posterior distribution, we need to choose a prior distribution and multiply this by the likelihood distribution. For analytical convenience and as a starting point, a conjugate prior distribution[8] is chosen. For our convenience and following Koop (2003), we choose the following distributions for the prior parameters[9]:

$$\beta|h \sim N(\beta_0, h^{-1}V_0) \qquad h \sim G(s_0^{-2}, v_0) \tag{4.10}$$

and their probability density functions equal to

$$P(\beta|h) = \frac{h^{\frac{k}{2}}}{(2\pi)^{\frac{k}{2}}}\exp\bigg\{-\frac{h}{2}(\beta-\beta_0)'|V_0|^{-1}(\beta-\beta_0)\bigg\} \tag{4.11}$$

$$P(h) = c_G^{-1}h^{\frac{v_0-2}{2}}\exp\bigg(-\frac{hv_0s_0^2}{2}\bigg) \quad \text{where} \quad c_G^{-1} = \bigg(\frac{2s_0^{-2}}{v_0}\bigg)^{\frac{v_0}{2}}\Gamma\bigg(\frac{v_0}{2}\bigg) \tag{4.12}$$

respectively. Using the fact that $P(\beta,h) = P(\beta|h)P(h)$ and rearranging some terms we can write the joint **prior** density distribution as

$$P(\beta,h) \propto h^{\frac{k}{2}}\exp\bigg\{-\frac{h}{2}(\beta-\beta_0)'|V_0|^{-1}(\beta-\beta_0)\bigg\}h^{\frac{v_0-2}{2}}exp\bigg(-\frac{hv_0s_0^2}{2}\bigg) \tag{4.13}$$

The joint prior follows a normal-gamma distribution $P(\beta,h) \sim NG(\beta_0, V_0, s_0^{-2}, v_0)$, where β_0 contains all the prior averages of the k coefficients and V_0 is the prior covariance matrix.

---

[8] A conjugate prior distribution once multiplied by the likelihood distribution gives a posterior distribution from a known analytical form because it belongs to the same class of distributions as the likelihood.

[9] Parameters with a 0 subscript refer to prior hyperparameters and parameters with a 1 subscript refer to posterior parameters.

Using equation 4.6 and equation[10] 4.14,

$$(y-X\beta)'(y-X\beta) = SSE + (\beta-\beta_{OLS})'X'X(\beta-\beta_{OLS}) \tag{4.14}$$

the likelihood exponent from equation 4.2 can be rearranged as

$$p(y|\beta,h) = \frac{h^{\frac{n}{2}}}{(2\pi)^{\frac{n}{2}}}\left\{\exp\left[-\frac{h}{2}\left\{SSE+(\beta-\beta_{OLS})'X'X(\beta-\beta_{OLS})\right\}\right]\right\} \tag{4.15}$$

Now we have all the elements to obtain the posterior distribution. Adapting equation 4.9, we can derive the posterior distribution for β and h, multiplying the likelihood, equation 4.15, by the joint prior distribution, equation 4.13, to get

$$\begin{aligned}
P(\beta,h|y) \propto & P(y|\beta,h)\times P(\beta,h)\\
& \propto h^{\frac{n}{2}}\left\{\exp\left[-\frac{h}{2}\left\{SSE+(\beta-\beta_{OLS})'X'X(\beta-\beta_{OLS})\right\}\right]\right\}\\
& h^{\frac{k}{2}}\exp\left\{-\frac{h}{2}(\beta-\beta_0)'|V_0|^{-1}(\beta-\beta_0)\right\}h^{\frac{v_0-2}{2}}exp\left(-\frac{hv_0s_0^2}{2}\right)\\
& = h^{\frac{v_1+k-2}{2}}\left\{\exp\left[-\frac{h}{2}\left\{v_0s_0^2+SSE+(\beta-\beta_{OLS})'X'X(\beta-\beta_{OLS})+\right.\right.\right.\\
& \left.(\beta-\beta_0)'|V_0|^{-1}(\beta-\beta_0)\right\}
\end{aligned} \tag{4.16}$$

where $v_1 = v_0 + n$.

The term inside the exponent, $(\beta-\beta_{OLS})'X'X(\beta-\beta_{OLS})+(\beta-\beta_0)'|V_0|^{-1}(\beta-\beta_0)$ can

[10] If the term $(y-X\beta)'(y-X\beta)$ is expanded and we add and subtract β_{OLS}, it can be shown that that equals $[(y-X(\beta_{OLS}+\beta-\beta OLS))'(y-X(\beta_{OLS}+\beta-\beta OLS)) = [(y-X\beta_{OLS})+X(\beta-\beta_{OLS})]'[(y-X\beta_{OLS})+X(\beta-\beta_{OLS})] = (y-X\beta_{OLS})'(y-X\beta_{OLS})+(\beta-\beta_{OLS})'X'X(\beta-\beta_{OLS})+2(\beta-\beta_{OLS})'X'(y-X\beta_{OLS})$. This last term is canceled out because $X'(y-X\beta_{OLS}) = X'e = 0$.
Therefore $(y-X\beta)'(y-X\beta) = SSE+(\beta-\beta_{OLS})'X'X(\beta-\beta_{OLS})$.

be expressed as

$$= (\beta_{OLS} - \beta_0)'X'XV_1V_0^{-1}(\beta_{OLS} - \beta_0) + (\beta - \beta_1)'V_1^{-1}(\beta - \beta_1) \tag{4.17}$$

where

$$V_1 = (V_0^{-1} + X'X)^{-1} \tag{4.18}$$

$$\beta_1 = V_1(V_0^{-1}\beta_0 + X'X\beta_{OLS}) \tag{4.19}$$

The only part of equation 4.17 containing β can be used to write:

$$p(\beta|y,h) \propto \exp\left[-\frac{h}{2}(\beta - \beta_1)'V_1^{-1}(\beta - \beta_1)\right] \tag{4.20}$$

Therefore, as this is the kernel of a Normal distribution, we can say that $p(\beta|y,h)$ is Normally distributed

$$p(\beta|y,h) \sim N(\beta_1, h^{-1}V_1^{-1}) \tag{4.21}$$

To obtain the remaining parameters for the posterior distribution $P(\beta,h|y)$, we need to integrate

$$p(h|y) = \int p(\beta,h|y)d\beta = \int p(h|y)p(\beta|y,h)d\beta. \tag{4.22}$$

Probability density functions must integrate to 1. If we integrate out the part $P(\beta|y,h)$, we can write the following proportional term[11]:

$$p(h|y) \propto h^{\frac{v_1-2}{2}} \exp\left[v_0 s_0^2 + SSE + (\beta_{OLS}-\beta_0)' X'XV_1V_0^{-1}(\beta_{OLS}-\beta_0)\right\} \tag{4.23}$$

$$= h^{\frac{v_1-2}{2}} \exp\left[v_0 s_0^2 + vs^2 + (\beta_{OLS}-\beta_0)'[(X'X)^{-1}+V_0]^{-1}(\beta_{OLS}-\beta_0)\right\} \tag{4.24}$$

$$= h^{\frac{v_1-2}{2}} \exp\left[-\frac{h}{2} v_1 s_1^2\right] \tag{4.25}$$

Equation 4.25 shows the kernel of the Gamma distribution with the following parameters:

$$p(h|y) \sim G(s_1^{-2}, v_1) \tag{4.26}$$

Knowing that multiplying a Normal density function by a Gamma distribution gives a Normal-Gamma distribution, as in equation 4.13, then it is straightforward to see that if we multiply the density function $(\beta|y,h)$ by $(h|y)$, the posterior distribution $(\beta,h|y)$ is Normal-Gamma with the following parameters:

$$\beta, h|y \sim NG(\beta_1, V_1, s_1^{-2}, v_1) \tag{4.27}$$

β is the parameter of interest. In order to make inferences about this parameter, we must obtain its marginal posterior distribution, $P(\beta|y)$:

$$P(\beta|y) = \int P(\beta,h|y)dh \tag{4.28}$$

Using Theorem $B.15$[12] from Koop (2003), it can be shown that the posterior distribution

---

[11] We use the proportional sign $\propto$ to group certain terms to form a kernel of a known distribution

[12] If $\theta = (Y', H)' \sim NG(\mu, \Sigma, m, v)$ then the marginal for Y is given by $Y \sim t(\mu, m^{-1}\Sigma, v)$

is a Student t with the following parameters

$$\beta|y \sim t(\beta_1, s_1^2 V_1, v_1) \tag{4.29}$$

With the respective mean, $E(\beta|y) = \beta_1$ (only exists if $v_1 > 1$) and variance $V(\beta|y) = \frac{v_1 s_1^2}{v_1 - 2} V_1$ (only exists if $v_1 > 2$). Initial values should be given for the prior hyperparameters $\beta_0, V_0, s_0^{-2}, v_0$. Once these are chosen, prior sensitivity analysis can be carried out. If v_0 is set to a value much smaller than n, and V_0 is set to a large value (depending on the determinant because of dealing with matrix form), we will have a relatively non-informative prior. This means that we have high prior uncertainty and the prior parameters will not have much of an effect in the posterior distribution; the data will mostly influence the posterior distribution. The purely non-informative prior is if $v_0 = 0$ and $V_1^{-1} = aI_k$ where a is a scalar with a value close to zero. If these values are chosen, the posterior parameters will equal the OLS values:

$$V_1 = (X'X)^{-1} \tag{4.30}$$

$$\beta_1 = \beta_{OLS} \tag{4.31}$$

$$v_1 = n \tag{4.32}$$

$$v_1 s_1^2 = vs^2 \tag{4.33}$$

In this case, the non-informative prior is improper (the density function does not integrate to one) and the notation for this prior is:

$$p(\beta, h) = h^{\frac{k-2}{2}} \tag{4.34}$$

Highest Posterior Density Intervals (HPDI)

These intervals are analogous to confidence intervals in the frequentist approach. We can use them to test point estimates hypotheses regarding the parameters of interest, β_j. If $[\beta_{min}, \beta_{max}]$ is an interval for β_j, this means that $\beta_{min} \leq \beta_j \leq \beta_{max}$. By definition, $[\beta_{min}, \beta_{max}]$ is a $100(1-\alpha)\%$ credible interval if

$$P(\beta_{min} \leq \beta_j \leq \beta_{max}|y) = \int_{\beta_{min}}^{\beta_{max}} P(\beta_j|y)d\beta_j = 1-\alpha \tag{4.35}$$

where α is the chosen confidence level. For each chosen confidence level, there are several (infinite) possible credible intervals, but the HPDI is defined as the narrower credible interval out of all the credible interval sets for a given α. In order to calculate these intervals, percentiles of the posterior distribution are used. Following equation 4.29 the Student T distribution is used instead of the Normal distribution. It is useful to report the posterior averages of the parameters of interest, $E(\beta_j|y)$ plus a 95 % HPDI of β_j to see if the variable X_j attached to β_j should be included in the model or not. If the HPDI includes zero, we can conclude $\beta_j = 0$.

Monte Carlo Integration Method

In order to infer the parameter of interest, the Monte Carlo integration method, explained below, can be used to obtain posterior estimates: If $\beta^{(r)}$ for $r = 1, ..., R$ is a random sample from $P(\beta|y)$, if $f(\cdot)$ and

$$\hat{f}_R = \frac{1}{R}\sum_{r=1}^{R} f(\beta^{(r)}) \tag{4.36}$$

then $\hat{f}_R \to E(f(\beta)|y)$ as $R \to \infty$. This integration method can be used to obtain point estimates such as the posterior mean $E(\beta|y)$. The procedure has the following steps:

- *First Step*: A random draw, $\beta^{(r)}$ is taken from the posterior for β, therefore from a t distribution as shown in equation 4.29.
- *Second Step*: Obtain $f(\beta^{(R)})$ and save the result.
- *Third Step*: Repeat the first and second steps R times.
- *Fourth Step*: Calculate the average of the R draws $f(\beta^{(1)}),...,f(\beta^{(R)})$.

Bear in mind that this procedure is always going to give us an approximation since we can not set $R = \infty$. Using this procedure, a 95% confidence interval for $E(\beta|y)$ can be obtained and will take the following form:

$$\hat{f}_R \pm 1.96 \frac{V[f(\beta)|y)]}{\sqrt{R}} \tag{4.37}$$

where $\frac{V[f(\beta)|y)]}{\sqrt{R}}$ is the numerical standard error which measures the accuracy of the approximation.

Bayesian Model Averaging (BMA)

In order to use this econometric technique, the Markov Chain Monte Carlo Model Composition (MC^3) is used for posterior computation. If there are $m = 1,2,...M$ potential models, each model contains a different set of parameters β_m, and the following probabilities, prior $P(\beta_m|M_m)$, likelihood $P(y|\beta_m, M_m)$ and posterior $P(\beta_m|y, M_m)$. These probabilities can be obtained for the $m = 1,2,...M$ models. Following equation 4.9, the posterior model probability, $P(M_m|y)$, can be obtained as well for each model. This probability can be defined as being proportional to the prior model probability, $P(M_m)$ times the marginal likelihood, $P(y|M_m)$. According to Koop (2016) it can be interpreted as "we are only $P(M_m|y)$ sure that M_m generated the data".

If we are interested in a particular set of parameters δ, we can summarise the information regarding these parameters of interest calculating the considered posterior probability using the following equation

$$P(\delta|y) = \sum_{m=1}^{M} P(\delta|y, M_m) P(M_m|y) \tag{4.38}$$

and calculating the posterior mean using

$$E(\delta|y) = \sum_{m=1}^{M} E(\delta|y, M_m) P(M_m|y) \tag{4.39}$$

Results should be obtained for all possible models and then average these results (as shown in equations 4.38 and 4.39). Posterior model probabilities act as weights. Computational methods help us obtain the posterior model probability and the marginal likelihood. M, the total number of possible models, depends on the number of explanatory variables chosen for the analysis. As the latter could be large, M could be large as well, and this could slow down the process of obtaining probabilities. In order to solve this, the algorithm MC^3 is used to take draws from the models' posterior distributions in order not to visit all the possible models, speeding the process.

Posterior probabilities give us the probability of inclusion of each independent variable. This probability is calculated as the proportion of models visited by the algorithm containing that specific variable. This probability can be used as a diagnosis tool to see if the variable has an effect on the dependent variable.

4.5 Results

4.5.1 Frequentist Analysis: Ordinary Least Squares

Tables 4.6, 4.7, 4.8 and 4.9 show the ordinary least squares estimates for Denmark, France, Spain and the UK. Each of these four tables have two columns. In the first column, the numeracy skills variable is included in the regression model, while in the second column this variable is omitted.

As expected in a wage equation, the results show that in each European country, holding other factors constant, one year of experience impacts earnings positively, but at a decreasing rate. In France and the United Kingdom, we observe that, controlling for skills, the returns to a year of experience are larger (2.1 and 2.6 percentage points, respectively) than in Denmark and Spain (1.3 and 1.4 percentage points, respectively). The results are in accordance to what Barrett (2012) found, controlling for skills has a small effect on the return to experience.

With respect to the male-female earnings gap, we can see that women in these four European countries earn less than men. Regarding this gap, in Denmark and France, women earn around 8% less than men, while in Spain and the UK, the gap is slightly larger, 13% and 11%, respectively if we control for numeracy skills. Papers controlling for skills such as Bonikowska et al. (2008) and Hanushek et al. (2015) also find that female wages are lower than men's.

If the numeracy skills variable is not included in the wage equation, the male-female earnings gap becomes larger in all the four countries. In Denmark and France, the coefficient becomes around 1.8 more negative, while in Spain and the UK the difference is even larger (more negative): 2.9 percentage points in the latter and 3.7 in the former.

Turning the attention to the coefficients regarding the highest level of education attained (medium education and tertiary education), it can be observed that, in the four countries, all of the coefficients are positive and statistically significant at the 1% significance level. The coefficients referring to education levels indicate the difference in the return (positive increase) that each native individual receives in comparison with a native with low level education or no education. It can be seen that the largest return comparing medium education level against low education or no education is found in Spain, where there is a significant 9.4% increase in earnings for those who have attained mid education level. The smallest return to a medium education level for natives is found in Denmark (5.2%), while these figures are slightly higher in France (6.8%) and in the UK (7.6%).

If native individuals with higher education are compared to native individuals with low education level or no education we can see that there is more homogeneity in the coefficient sizes across the four countries oscillating around 29%, with the exception of Denmark, where this figure is again slightly lower, 23%[13]. These coefficients are lower than the ones found for Canada in Bonikowska et al. (2008) with a magnitude of around 70% for tertiary education and similar but still lower than the 35% return found for Canada as well in Ferrer et al. (2006).

The education coefficients' sign is in accordance to our expectations and the Human Capital Theory. The more years of schooling completed, the higher the earnings

[13] As the literature shows that there are different returns for males and females, an interaction term (gender with education and numeracy skills) has been included to control for these differences, but results are found to be insignificant.

received. However, one could think that these returns to tertiary education are lower than expected. The reason for this could be that nowadays, more individuals have access to higher education than before. If more people have attained higher education, this qualification loses its own differentiating value as Carabaña (1996) pointed out for the Spanish case.

Once we explore the variation of the natives' education coefficients when controlling/not controlling for skills, a pattern in the four countries arises. If numeracy skills are not present in the wage equation, the returns to education are higher than if the numeracy skills variable is included. This difference is larger in the tertiary education coefficients than in the medium education ones. The smallest difference, 2.7 percentage points, corresponds to Denmark for the medium education coefficient, while in France this difference rises to 4 percentage points and 5.8 percentage points in the UK. Regarding tertiary education in Denmark and Spain, a difference of around 7 percentage points arises, while in France and the UK the difference is even larger, 10 percentage points.

For natives, the numeracy skills absorb part of the returns to levels of education, and independently of having attained certain level of education numeracy skills impact positively on earnings. The returns to education decline when controlling for skills and a significant part of these returns arises from how skills are rewarded in the labour market, at least in the four European countries analysed. When estimating earnings equations, Ferrer et al. (2006), Barrett (2012), Green and Riddell (2003) and Bonikowska et al. (2008) found this same pattern as well. Green and Riddell (2003) used years of education rather than attained levels and found a reduction in the education coefficient of 6-10 percentage points when skills are present in the earnings equation. Ferrer et al. (2006) found that this difference is 2 percentage points for holding a high school certificate and 21 percentage points for holding a university degree in Canada. Returns to credentials were estimated in Barrett (2012) where the decrease found in the coefficient was 2 percentage points for post-secondary diploma, 4 percentage points for a Bachelor degree and non existent for holding a Post-Graduate certificate in Australia.

Focusing now on the return to numeracy skills for natives, we can see that the coefficient attached to this variable has a positive sign in the four European countries, suggesting that the higher numeracy skills score an individual has, the higher his/her wage. While in the United Kingdom a one standard deviation increase in the numeracy skills score is associated with an increase of around 11% in hourly earnings, this figure is lower in the remaining European countries. Denmark has the smallest return, 7%, to numeracy skills out of the four countries analysed, while a one standard deviation increase in the numeracy skills produces an increase in earnings of 8% and 9% in France and Spain, respectively. The numeracy skills coefficients are very similar to those obtained in Hanushek et al. (2015) which are 7%,8%, 9% and 11% for Denmark, France, Spain and the UK, respectively when controlling for educational attainment (levels) as we do in the present analysis.

For the interaction terms referring to Country of Origin and Time in the country, we can see that, in general and when controlling for numeracy skills, these variables are not statistically significant. Only the coefficient for recent immigrants from developing countries is statistically significant in Denmark and Spain. These coefficients indicate that immigrants not from the OECD with less than 15 years in the host country have 17.2% and 8.8 % lower earnings than natives in Denmark and Spain, respectively. If we explore how these coefficients change in size when including/not including the numeracy skills variables we can see that, when controlling for skills, the coefficient becomes more negative in both countries, specifically 2 percentage points more negative in Denmark and 4 percentage points more negative in Spain.

Once we look at the interaction terms for immigrants between time spent in the country, country of origin and level of education, in general, only a few of them are statistically significant in Denmark and Spain. None of these variables are significant in France or the UK indicating that there are no difference in returns to levels of education between natives and immigrants no matter the level of education attained, country of origin or time spent in the country.

Looking at Denmark, the difference in returns to a medium education level between immigrants from the OECD who have spent more than 15 years in the host country and natives is negative, -21 percentage points, while this difference in slopes for recent immigrants not from the OECD and natives is positive, 11 percentage points, at the 10% significance level. In Spain, we find that the returns observed for recent immigrants from developed countries who have attained a medium education level are 22 percentage points lower than mid educated natives at the 5% significance level. The returns to a mid education level of recent immigrants from developing countries are approximately 13 percentage points lower than natives with the same level of education at the 5% significance level. The negative coefficients on these interactions could reflect discrimination, disadvantages that immigrants suffer in the host country, different educational distributions or other factors.

In we look at the tertiary education interaction variables it can be seen that in Denmark, depending on the immigrant's country and length of stay the coefficients have different sign, while in Spain the sign is always negative although not all the coefficients are statistically significant. The returns to a tertiary education level for established immigrants from the OECD are 18 percentage points lower than natives who have attained this education level (significant at the 5% level). Tertiary educated immigrants who spent less than 15 years but are from developing countries, in Denmark surprisingly have 10.8 percentage points higher returns than tertiary educated natives at the 10% significance level, while we observe that the return to higher education is 8.5 percentage points lower than natives for those immigrants from developing countries who spent more than 15 years in the host country. In Spain, in comparison to tertiary educated natives, recent immigrants and established immigrants from developing countries have 16.6% and 106% lower returns, respectively.

It can be seen that regarding these coefficients, none of the interaction terms have a statistically significant effect in France and the UK. One possible explanation is the assimilation scheme followed by France to try to eliminate the differences between immigrants and natives and promote immigrants' integration into French society and the point system used in the UK to admit highly skilled and educated immigrants. Aydemir (2012) argues the point system (used to gain a Visa) generates a more skilled immigrant pool and prevents immigrants from taking up a low-skilled occupation. Related to not finding statistically significant differences between immigrants and natives returns' in these countries, Clark and Lindley (2009) argue that immigrants acquire skills specific to the host country and this allows them to succeed in and improve their labour market performance in comparison to natives.

The only countries that have statistically significant effects when interacting numeracy skills with time spent in the country and immigrant group are Denmark and the UK. In Denmark, a one standard deviation increase in the numeracy skills score of established immigrants from developed countries is associated with around 7 percentage points higher return than natives at the 10% significance level, while for recent immigrants from developing countries a one standard deviation increase in the numeracy score provides them with about 4 percentage points less return than natives at the 5% level. In the UK recent immigrants from developing countries' returns to numeracy skills are around 7 percentage points higher than natives' returns to skills. None of the immigrants' numeracy skills coefficients are significant in France and Spain. This is in accordance to Bonikowska et al. (2008) who rejects the hypothesis that immigrants receive lower returns to skills than natives in Canada, while the positive difference in returns to skills favouring the immigrants found in the UK could arise from the point system admission policy.

If we look at the change in the size or sign of the coefficients regarding the interactions no pattern is found in Denmark. In the latter, the coefficient indicating recent immigrants from developing countries becomes less negative when controlling for skills, from -19 percentage points to -17 percentage points, while if we compare returns for established immigrants from the OECD when controlling for numeracy skills the coefficient becomes more negative, from -14 percentage points to -21 percentage points. In Spain it can be seen that although for the natives' education dummies controlling for skills decrease the returns, for immigrants the opposite occurs. In Spain, the statistically significant returns to education decrease around 4 percentage points when controlling for skills.

Although Tables 4.2, 4.3, 4.4 and 4.5 show that across all the countries immigrants have on average lower numeracy skills score than natives, these differences in returns are not statistically significant in France and Spain. After looking at the results obtained, it can be said that in the UK the education level attained could be more of a channel for differences in earnings between immigrants and natives than numeracy skills. Differences in the returns to education between natives and immigrants are only observed in Denmark and Spain.

One possible explanation for these results is immigrant self-selection. On the one hand, immigrants determine whether and where to migrate, but at the same time, host countries follow quotas to admit a certain number of immigrants and with a particular set of skills. As Damas de Matos (2014) argues, immigrants could be positively selected relative to their compatriots who stayed in the country of origin. In this case, the pool of immigrants that arrive to the host country will not be a representative sample of all the potential immigrants from a country. Unfortunately, this issue cannot be addressed in this context. According to this argument, our estimates could be biased upwards because of having immigrants living in the host countries representing the top part of the educational and skills distributions.

In the next section we use Bayesian methods which allow us to include prior valuable information regarding the coefficients of interest. It will be seen that the results from the Bayesian analysis reinforce the conclusions drawn from the frequentist analysis.

4.5.2 Bayesian Analysis

Prior Elicitation

In order to perform the Bayesian analysis we need to elicit the prior information that will complement the information given by the data. We could perform objective Bayesian analysis by setting a non-informative prior where all the beta averages (β_0) are equal to

0. If these priors are chosen, as shown in the methodology section, the posterior betas will give practically equal coefficients as the betas obtained by OLS. By doing this, the data mainly define the posterior parameters and the estimates become only a function of the likelihood. Alternatively, an informative prior could be used. The betas of interest are the coefficients attached to the following variables: Medium Level of Education, Tertiary Education and the four interaction terms created by multiplying these two variables with the variables indicating the country of origin group[14]. The priors attached to these variables could take several possible values, but, in order to be informative, they must all differ from 0.

Two papers, Bratsberg and Terrell (2002) and Hanushek et al. (2015), are used in conjunction to set the prior parameters (betas). The rate of return to a year of education by country of birth in the US 1990 Census is used to set the prior averages for the interaction terms. Bratsberg and Terrell (2002) give the return to a year of schooling that an immigrant from a particular country of origin receives in the US labour market. The second paper, Hanushek et al. (2015), is used for the return to a specific level of education for the coefficients that refer to native returns to schooling. These two papers present the return to schooling per extra year of education. In order to make the frequentist and the Bayesian results comparable in this chapter, we need to convert yearly returns into returns to levels of education. This is done by multiplying the returns to a year of education by the factor Δ explained below. To calculate the prior values for the coefficients that refer to the highest education level attained by the natives ($\beta_0^{Mid.Ed.}$

[14] The numeracy skills variable is not included in the Bayesian analysis because one of the papers (Bratsberg and Terrell (2002)) used to build the priors for the interactions does not offer estimates controlling for this variable. As shown earlier in the chapter, when the skills variable is included in the regressions, it is likely that the magnitude of the education coefficients is somewhat larger

Average years of education per highest level of education attained

	Denmark	France	Spain	UK	USA
Low Education (reference)	9	6	8	11	9
Mid. Education	12	12	12	12	12
Tertiary Education	16	16	16	16	16
Difference in Years Low to Mid. Ed. (Δ_1)	3	6	4	1	3
Difference in Years Low to Tertiary Ed. (Δ_2)	7	10	8	5	5

Source of Data: PIAAC

and $\beta_0^{Tert.Ed.}$), recall that the reference category for these coefficients is natives that have attained a Low Education level as their highest level of education. The latter will be the comparison group.

The table above presents the average years of education for each level of education attained per country. The prior averages are calculated using two quantities; the rate of return per year of education and, the last rows (Δ) from the table above. An example of the procedure to calculate the prior averages is given for Denmark. According to Hanushek et al. (2015), the return to a year of education in Denmark is 5.5%. For the beta corresponding to the Middle Level of Education category, we know that in Denmark, the difference in the average number of years between attaining a Lower Education level (the reference category for this group) and Middle Education Level is $\Delta_1 = 3$ so we can calculate the prior using the following formula:

$$\beta_0^{Mid.Ed.} = \Delta_1 \times 0.055 = 3 \times 0.055 = 0.165$$

The same procedure can be followed to calculate the coefficient for Tertiary Education

$$\beta_0^{Tert.Ed.} = \Delta_2 \times 0.055 = 7 \times 0.055 = 0.385$$

To calculate the prior values for the interaction terms that correspond to the levels of education ($Mid\ \ Ed. \times OECD$, $Tertiary\ \ Ed. \times OECD$, $Mid\ \ Ed. \times Non-OECD$ and $Tertiary\ \ Ed. \times Non-OECD$), the returns to a year of education from Bratsberg and Terrell (2002) are used. This paper presents a table with the returns to a year of education that immigrants from a particular country of origin obtain in the host country (the paper uses US as the host country). As we have divided the immigrants into two groups depending on the level of development of their country of origin, we can separate the returns by countries and create two categories, OECD countries and Non-OECD countries. Once this is done we will have a list of returns to a year of schooling that immigrants from the OECD receive in US and we will have a list of returns that immigrants from a Non-OECD country receive in the US. If we average these returns, we obtain that the return for a year of education for immigrants from an OECD country is 5.9 % and the same figure corresponding to immigrants from Non-OECD countries is 4.1%.

We need to bear in mind that using this method has a main caveat. The return for an OECD immigrant could be higher than for a native because we are are comparing, for example, natives in Denmark to immigrants in the US. The following assumption is made to calculate the prior parameters: immigrants from a particular country of origin receive the same return in Denmark (France, Spain or the UK) as in the US. If papers that study the returns for immigrants in these four countries had been used instead, two potential problems would arise: different regression models are used across papers specific to a host country and papers rarely divide the immigrants into the immigrant group that we use (OECD and Non-OECD countries). The ideal scenario to set the priors will be to find four papers (or one containing all the information) that study the returns to education for immigrants divided by their country of origin into OECD and Non-OECD countries in Denmark, France, Spain and the UK as host countries and using the same specification model, i.e controlling for the same variables in the earnings equation. In the absence of this, Bratsberg and Terrell (2002) is used to calculate the priors in the most accurate and fair way.

Now that we know the average return to a year of education for immigrants from the OECD and for immigrants not from the OECD (5.9% and 4.1%, respectively) we can calculate the coefficient for the interaction variables. For Mid Ed. $\times$ OECD, we will use the same procedure as before, but using the difference in average years of education across levels attained. In this case, we take into account that the reference category will be individuals who have Low Level of Education and are born in the country of the survey. These calculations are illustrated using Denmark:

$$\begin{aligned}\beta_0^{MidEd.\times OECD} =& \Delta_1 \times (RtE_{OECDcountries} - RtE_{CountryofSurvey(Denmark)}) \\ =& 3 \times (0.059 - 0.055) = 0.012\end{aligned}$$

$$\begin{aligned}\beta_0^{MidEd.\times Non-OECD} =& \Delta_1 \times (RtE_{Non-OECDcountries} - RtE_{CountryofSurvey(Denmark)}) \\ =& 3 \times (0.041 - 0.055) = -0.042\end{aligned}$$

$Mid\ \ Ed. \times OECD$ is the difference in returns to a mid level education between immigrant from the OECD and natives in Denmark. Since 5.9% is used as the return for immigrants and 5.5% is used for the return for natives (i.e. a higher return for immigrants), then this means the prior for this coefficient should be positive. If we calculate the prior for $MidEd. \times Non - OECD$, then as we observe that the natives have a higher return than the immigrants from Non-OECD countries, the sign of this coefficient must be negative.

The prior standard deviations would reflect the level of uncertainty about that parameter. For prior elicitation of the standard deviations and following Koop (2003)'s methodology, we first need to choose a range where the dependent variable, logarithm of wage per hour (y), could lie. In each of the countries, we have a different currency. Income per hour is chosen to range between a maximum of the equivalent to £200 per hour and the minimum wage per hour from each country in 2011. Following Koop (2003), 1 % of the range is chosen to calculate h in each country. We choose $\frac{v_0}{N} \approx 0.01$, meaning that the prior information regarding h should be around 1% of the weight as the information given by the data. These calculations are needed to calculate the prior variances of the betas of interest.

By choosing the maximum and minimum value that each coefficient could take, we can create a measure of dispersion for each coefficient and then we divide these prior variances by the constant $\frac{v_0 s_0^2}{v_0 - 2}$. [15] (which will depend on s_0^2 that at the same time depends on h) and obtain the prior standard deviations $S.D(\beta)$. This procedure is needed to scale the prior variance. The marginal distribution of β is a t distribution and if $\beta \sim t(\beta_0, s_0^2 V_0, v_0)$ then

[15] Results were obtained for different values of s_0^2 and v_0 keeping $Var(\beta)$ constant. If s_0^2 is set to a particular value and different values of v_0 are tried, the results remain. If v_0 is set to a particular value and different values of s_0^2 are tried, this could alter the size of the posterior coefficients. This occurs because of the formula used to scale the variance, $Var(\beta) = \frac{v_0 s_0^2}{v_0 - 2} V_0$. As we multiply by v_0 and divide by $v_0 - 2$ this does not make a big effect in the variances, while as s_0^2 is present only in the numerator of the fraction, it allows the posterior betas to change in size for different values of s_0^2 due to a change in the variance.

the prior variance is equal to $Var(\beta) = \frac{v_0 s_0^2}{v_0 - 2} V_0$.

To choose the maximum value that each coefficient can take, we can use the OLS coefficients (Table 4.10) as a guide to set the lower and upper limits for the returns to levels of education and the interaction variable coefficients. The coefficient on the Middle Level of Education variable is going to range from 0 to 0.35, therefore this coefficient's variance will equal $Var(\beta_0^{MidEd.}) = 0.07^2$ and a 95% prior probability[16] to the range [0.07,0.35] is attached. The coefficient on Tertiary Education will range between [0.36,0.70], therefore if the same procedure is used, $Var(\beta_0^{TertEd.}) = 0.085^2$. The range chosen for the interaction variables is [-0.3,0.11] therefore $Var(\beta_0^{Mid.Ed. \times OECD}) = Var(\beta_0^{TertEd. \times OECD}) = Var(\beta_0^{Mid.Ed. \times Non-OECD} = Var(\beta_0^{TertEd. \times Non-OECD}) = 0.05^2$. These ranges are wide enough to acknowledge that there is uncertainty about the coefficient prior values.

Bayesian Posterior Results

In order to make a fair comparison between both estimation approaches we will use the results in Table 4.10 as the baseline. This table contains the OLS wage equations for each country but not controlling for time in the country[17] and numeracy skills. Tables 4.11, 4.12, 4.13, and 4.14 show the Bayesian estimation results. The first two columns of these tables show the prior values given to each beta related to returns to education. In order to see how prior parameters affect the posterior results we use two priors, one based

[16] As we are dealing with a bell symmetric distribution we can apply the conecpt that 95% of the observations will lie within two standard deviations from the mean.

[17] Time in the country was not included in the Bayesian analysis because no papers were found to elicit the priors regarding the interaction coefficients.

in information from Denmark and one based in information from United States[18]. The origin of these prior values was explained in the prior elicitation section. The next three columns show the posterior means when using a non-informative prior and using these two informative priors. The posterior betas will be a compromise of the prior betas and the observed data. If a non-informative prior is used the results are very similar to the OLS results, while if an informative prior is used, the posterior betas are slightly different in size and sign from the OLS. The last column shows the upper and lower limits of the Highest Posterior Density Intervals when the informative prior based in Denmark is used. If this interval contains 0, then it could be inferred that the explanatory variable does not affect the dependent variable (logarithm of wage per hour). Differences comparing the frequentist to the Bayesian approach are next highlighted.

Table 4.11 contains information for Denmark. We can see that the Bayesian analysis supports the statistical relevance of the set of variables related to natives' education levels. The HPDI attached to these variables do not contain the value zero. This is in accordance with the frequentist results, where all these variables have positive and statistically significant coefficients. Regarding the size of the effect for these variables, it can be seen that the three coefficients have a larger effect (i.e. higher returns) in the Bayesian framework. The latter suggests that for the natives in Denmark the returns to a medium education level in comparison with natives with low education or no education are around 10 percentage points larger, while the returns to tertiary education for natives in comparison with the reference group are 32 percentage points larger in the Bayesian analysis using the priors based on Denmark. The interaction terms are found to be

[18] When the priors are based in Denmark the returns to education are calculated using returns from Denmark and average years of education for each education level in Denmark while if the prior is based in the US, the priors are calculated using returns in the US and average years of education in the US.

insignificant in both the OLS and the Bayesian results. However, the Bayesian approach estimates slightly larger coefficients for all the coefficients.

Table 4.12 presents the results for France. For the first two variables regarding Medium and Tertiary education for the natives, as in Denmark, the Bayesian results coincides with the frequentist results (Table 4.10), and all coefficients are statistically different from zero. Using an informative prior based in France suggests 11% larger returns to mid education level in comparison with the natives with low level of education or no education, and 39% larger returns for natives who attained tertiary education. The size of these coefficients is barely the same than the corresponding ones in the frequentist approach for this country.

In France, the education level and interactions coefficients (column 5) have the same sign in both estimation approaches, the frequentist and the Bayesian. The Bayesian results suggests that immigrants from developed and developing countries that have attained a middle education or tertiary education level have lower returns than the reference group, although all of the interaction coefficients are not statistically different from zero. Again, this coincides with the frequentist approach. The same pattern as in Denamark is observed regarding the size of the coefficients. All the interaction coefficients but the one on Tertiary Ed. $\times$ OECD become less negative when using the Bayesian estimation method.

Posterior results from Spain in Table 4.13 coincide with the positive signs of the coefficients regarding the natives' education levels. Column 5 shows that the coefficients are all statistically different from zero in both estimation approaches. As observed in Denmark and France, the Bayesian results for the Spanish case suggest slightly higher returns to education for the natives than the returns given in the frequentist approach.

It can be seen that Spain is the only country with a discrepancy regarding the statistical significance of the interaction terms regarding the two immigrant groups. Column 5 from Table 4.13 shows that the posterior Bayesian results using informative prior supports the idea of the two first interaction coefficients, Mid. Ed. $\times$ OECD and Tertiary Ed. $\times$ OECD, being equal to zero. This is different to what the OLS estimation suggests from Table 4.10 column 3, where all the coefficients referring to immigrants' education are statistically significant but the one indicating Tertiary Ed. $\times$ OECD. Both estimation approaches support that the interaction variables corresponding to immigrants from developing countries are negative and statistically significant in both approaches, suggesting that immigrants from developing countries have lower returns to education than natives in Spain. In this country, there is no difference in the returns to education between natives and immigrants from developed countries. Regarding the size of the interaction coefficients in Spain it can be seen that the Bayesian approach suggests slightly less negative coefficients for interactions corresponding to immigrants from developed countries and more negative coefficients for immigrants from developing countries.

Table 4.14 presents the results for the UK. We can see that the coefficients that represent the returns to levels of education between the natives that have attained low level of education or no education and the natives with higher levels are statistically different from zero in both approaches. Again, the Bayesian approach, column 5 from Table 4.14, suggests that for this country the natives' education coefficients are slightly larger than in the frequentist estimation. Both education coefficients have a positive sign, as observed in the other three countries.

Turning our attention to the interaction coefficients' size and sign between the two estimation approaches, no pattern is observed. The Bayesian approach suggests a more negative coefficient for immigrants from developed countries that attained a medium education level and an opposite sign for immigrants from developed countries that attained a tertiary education. For the interactions indicating returns to education of immigrants from developing countries it can be observed that the Medium education level coefficient in the Bayesian approach has a different sign, suggesting that immigrants from developing countries have lower returns than natives with a medium level of education. A positive difference in returns is observed when comparing immigrants from developed countries and natives that attained tertiary education, although the Bayesian approach suggests a slightly smaller coefficient. Regarding the statistical significance of the interaction coefficients, none of them are statistically significant in any of the estimation approaches.

The influence that the prior parameters have on the posterior ones is next reviewed to check the Bayesian estimation stability. In order to do this, the size of the coefficients in columns 4 and 5 from Tables 4.11-4.14 is compared. It can be seen that a pattern arises. The posterior coefficients when using the US's prior are more negative that when using the informative prior from each individual country. There is only an exception of this occurring in the UK, where the UK and the US priors lead to opposite signs in the coefficient for immigrants from developing countries that attained Tertiary Ed. The variation in the size of the coefficients ranges between 2% (Mid. Ed. $\times$ Non-OECD and Tertiary Ed.$\times$ Non-OECD in Spain) and 14% (Tertiary $\times$ OECD).

The influence of the prior on the posterior curve is inspected graphically in Figures 4.3 - 4.6. The figures could help us to check if changing the prior parameters leads to significant changes in the posterior curve once the prior information is combined with the likelihood.

These figures contain two plots for each coefficient, one contains the curve using the prior based on the respective European country data and the other plot contains the prior with US information. The posterior curves are represented with a solid line, the prior curves with a dashed line and the likelihood curves with a dotted line. β_1 and β_2 correspond to the natives' education dummies, Medium Education and Tertiary Education, while β_3, β_4, β_5 and β_6 correspond to the interactions from Tables 4.11-4.14.

A pattern that all these four countries share is the stability for the coefficients on natives' education (β_1 and β_2) no matter what prior information is used. The posterior curves for these coefficients remain closer to the likelihood than to the prior curves, therefore the prior does not drive the posterior results. Figure 4.3 shows that in Denmark, β_3 and β_4 are highly affected by the prior chosen because the posterior curve does not remain centered when another prior is used. In fact, β_3 largely resembles the US's prior. β_5 and β_6 posterior curves are slightly less influenced by the prior parameters chosen because when US's prior is used, the posterior curves do not move away from the data.

Figure 4.4 for France shows the same pattern as in Denmark; β_1 and β_2, the posterior curves corresponding to native education levels are not influenced by the prior parameters. Again, β_3 and β_4 are more affected than β_5 and β_6 by the choice of prior. When using the US's prior, the β_3 posterior curve resembles to the US's prior curve as happened in Denmark.

In Spain, Figure 4.5, the natives' education coefficients, β_1 and β_2 are not affected by the prior choice, while the rest of the coefficients are. The β_3 and β_4 posterior curves move closer to the US's prior curve than the Spanish prior. Although the β_5 and β_6 posterior curves change when using US priors, the difference is small.

Some differences in the posterior curves could be observed in the UK when using different prior parameters. β_1 and β_2 remain almost unchanged when the US's prior is used, while the β_3 and β_4 posterior curves are slightly displaced when using the US's prior. Nevertheless, β_5 is more affected when using the US's prior and the posterior curve is displaced to the left, bringing the posterior curve closer to the US prior curve. There is not much difference in the posterior curve for β_6 when using the two prior curves.

Figure 4.3: Bayesian Density Plots. Denmark

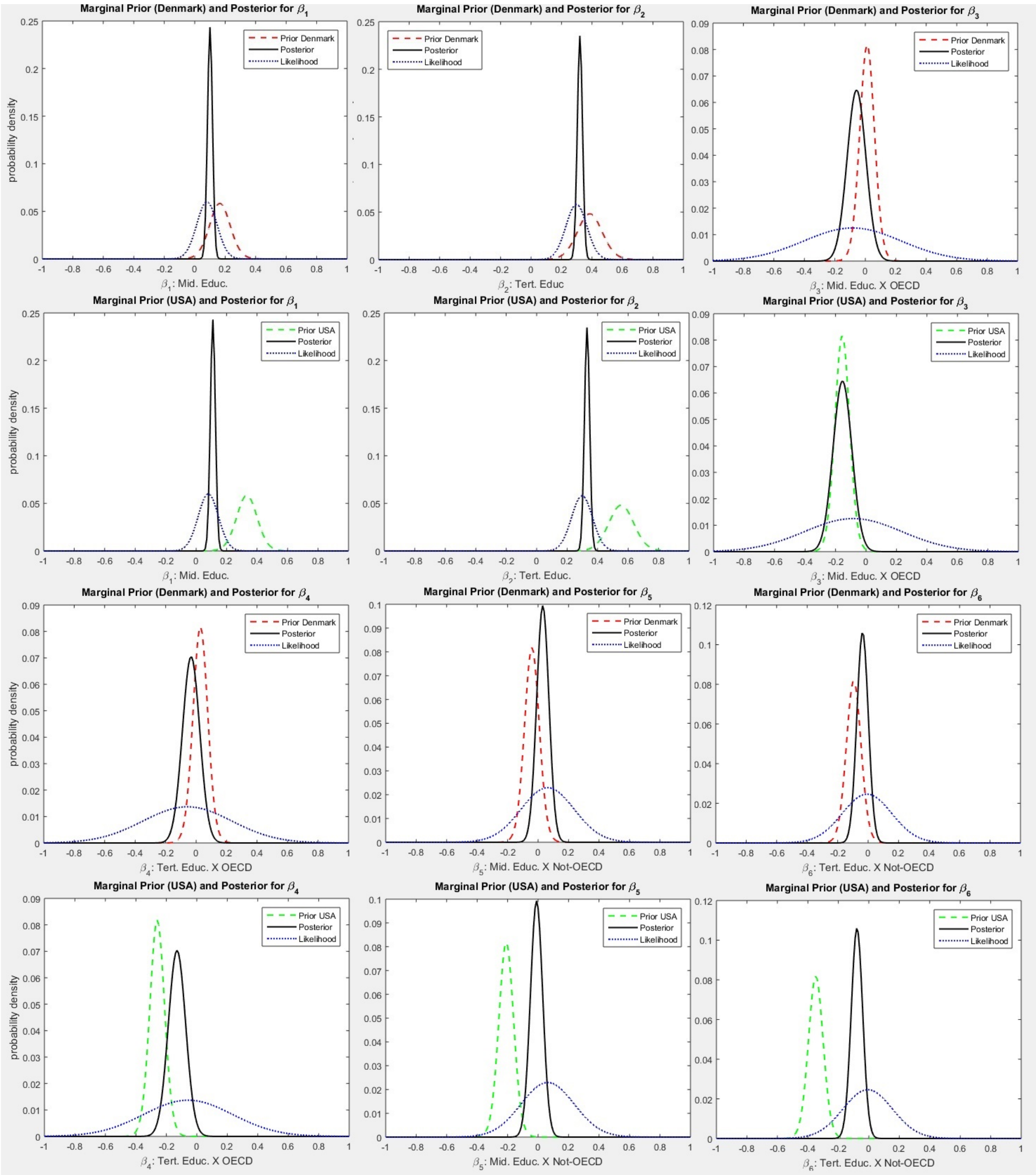
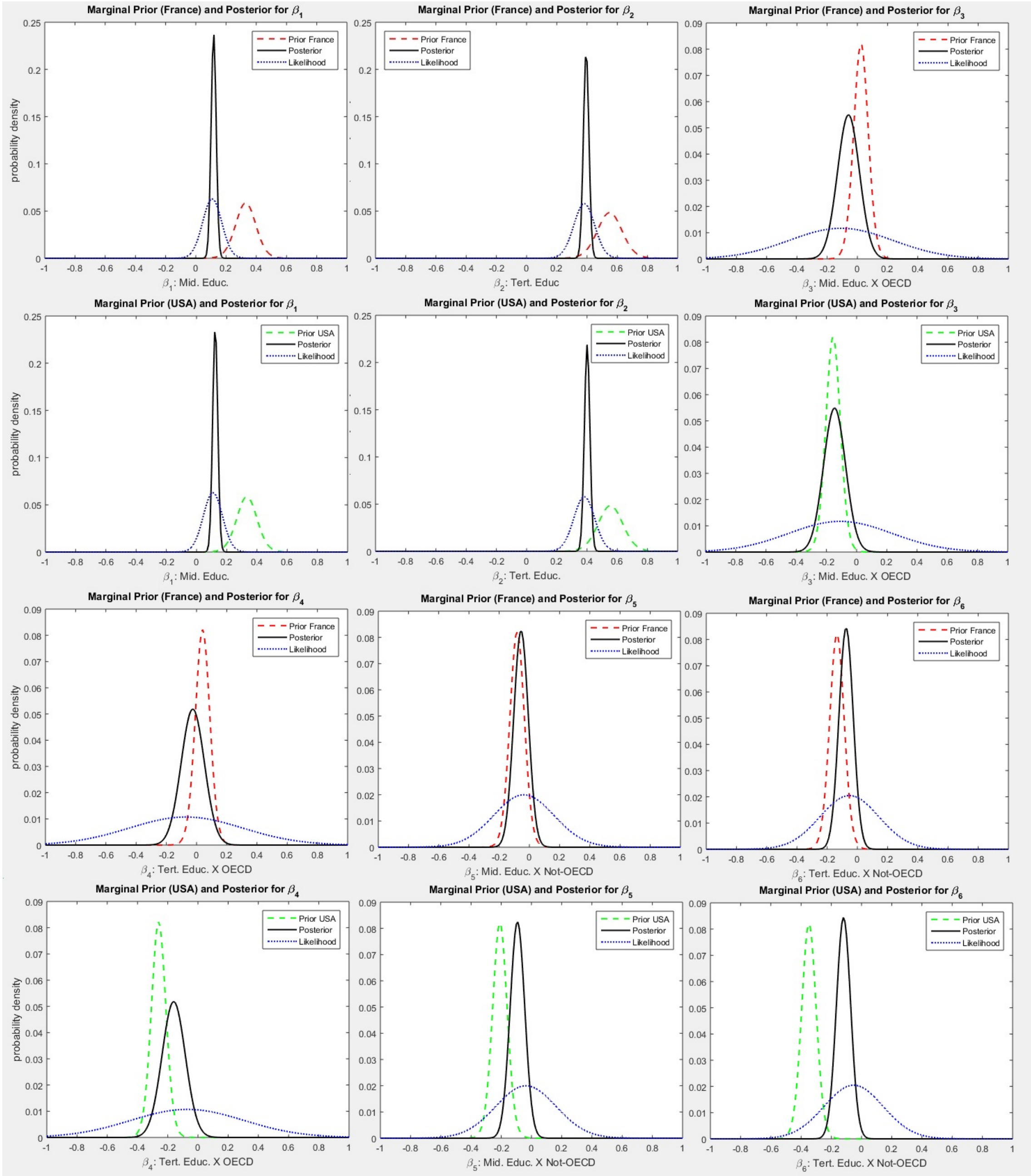
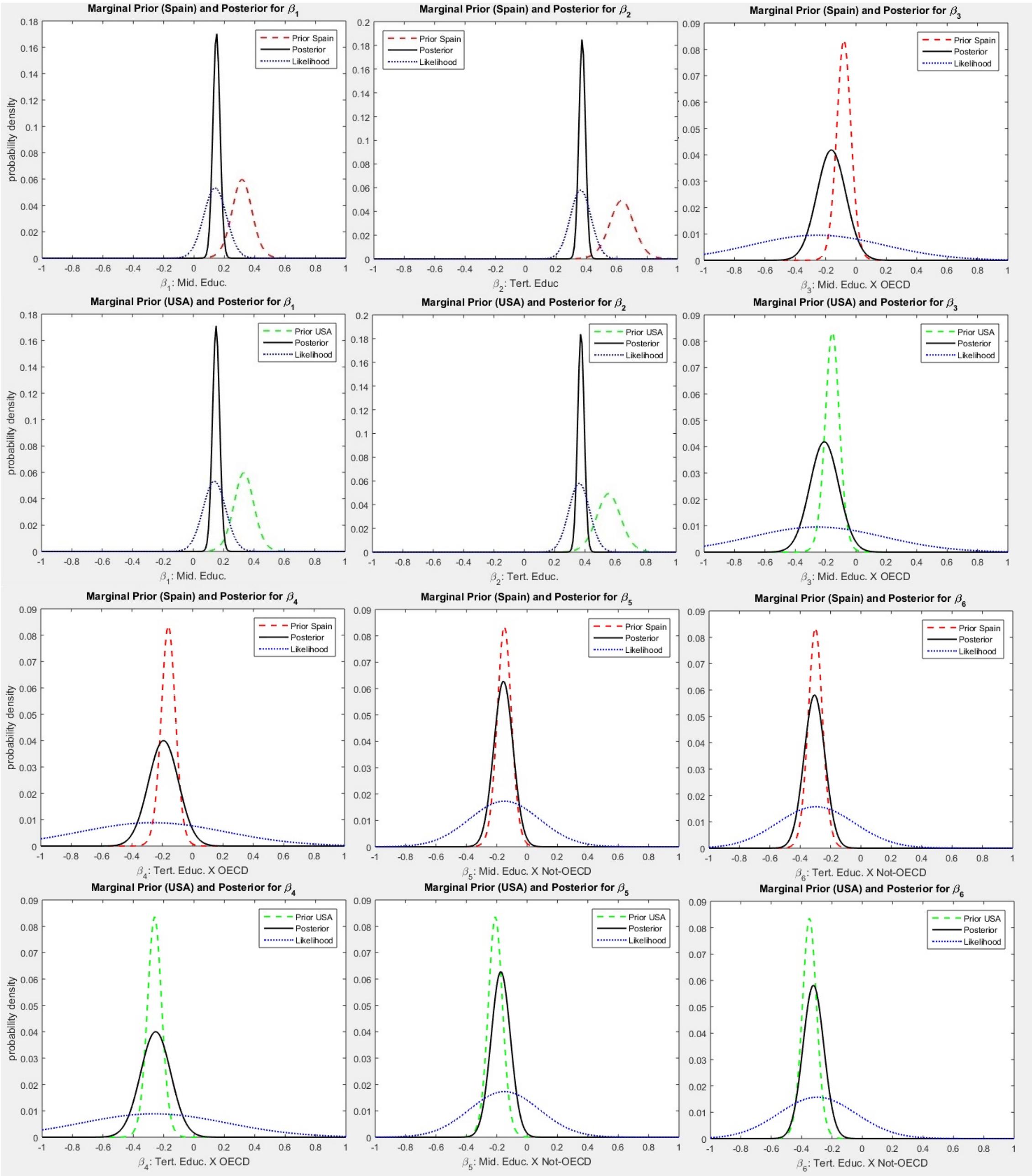


Figure 4.4: Bayesian Density Plots. France



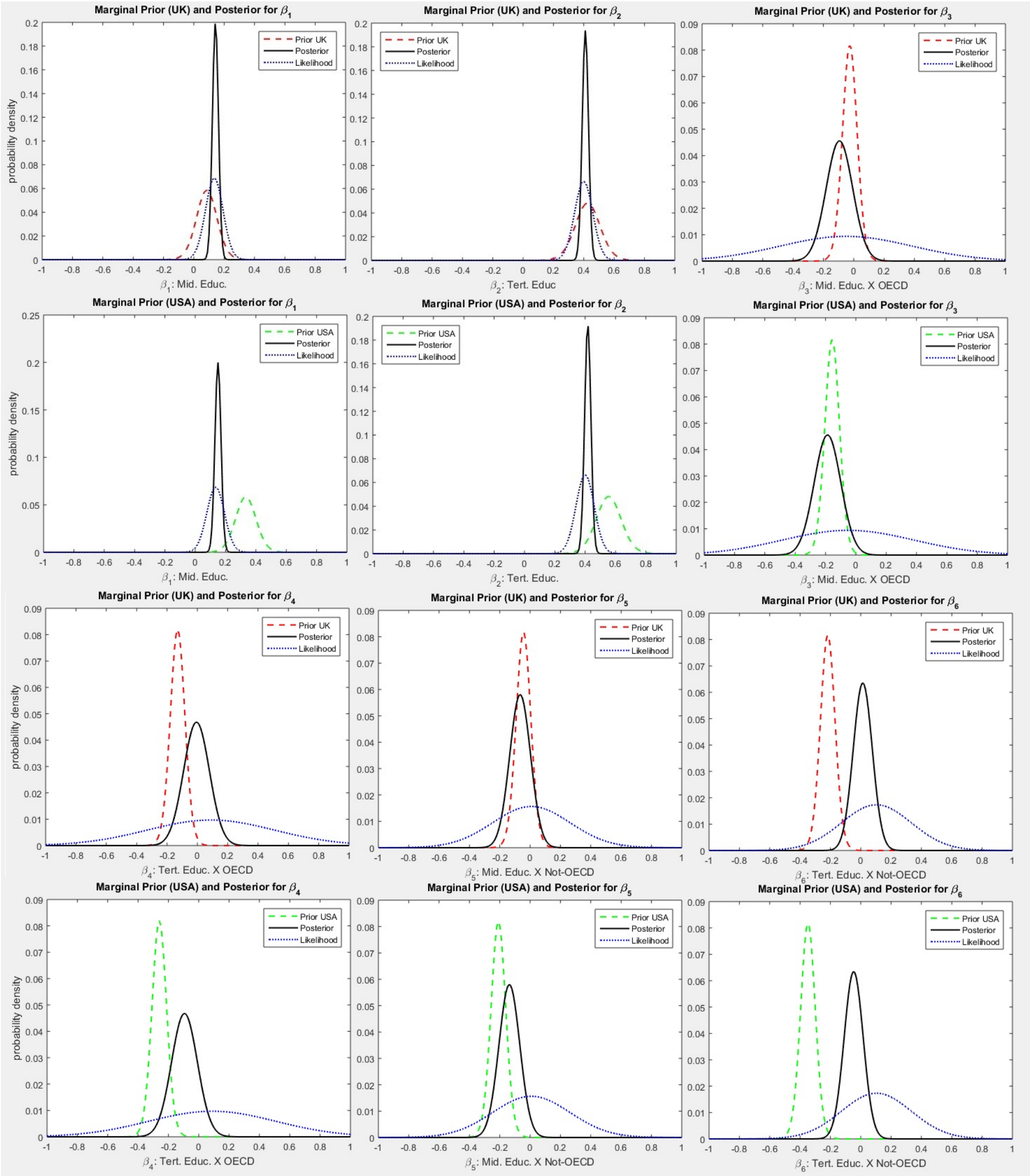
To summarise, we can say that for the four European countries, the Bayesian and frequentist approaches support the positive returns to levels of education for the natives. Regarding the interaction term variables, both estimations methods suggest there is no difference in returns between immigrants and natives in all the countries studied apart from Spain. Although the difference in returns to education between immigrants and natives is not statistically significant for any group of immigrants, the Bayesian posterior

Figure 4.5: Bayesian Density Plots. Spain



results (graphs shown in Figures 4.3 - 4.6 and coefficients in Tables 4.11-4.14) are sensitive in size to the choice of prior, therefore, a definitive coefficient for the β's cannot be established using the Bayesian estimation approach.

Figure 4.6: Bayesian Density Plots. UK



Bayesian Model Averaging (BMA)

As there are a few discrepancies between the statistical relevance of the variables using the frequentist approach and the Bayesian approach, the analysis presented is complemented by a Bayesian model averaging estimation. In the economics of education literature, researchers are interested in finding what variables affect wages, so using the BMA it can be seen which explanatory variables in the earnings equation present a high posterior probability, and therefore are relevant to explain the dependent variable.

Table 4.15 presents the explanatory variables of interest. In each country, we have a different number of explanatory variables because of differences in the region dummies included in each country-specific regression. The same control variables in 4.15 are used as in the frequentist approach, 4.10. Columns 1, 4, 7 and 10 from Table 4.15 indicate the probability that the explanatory variable should be included in the model. This probability is calculated as a percentage of models visited by the MC^3 algorithm that contain the specific control variable out of the total number of models visited. The presented results are obtained by setting the number of draws to 10^6 and discarding the first 10^5 to achieve accuracy. These first draws are discarded because they could be highly influenced by the starting values of the chain and therefore not provide accurate information regarding the target distribution. The other columns contain the posterior mean and posterior standard deviation for each coefficient in each European country. These posterior measurements are averaged across the visited models. We can see that some posterior standard deviations are larger than the posterior means, expressing some degree of uncertainty.

Focusing on the posterior probabilities and checking common patterns in the four countries, we can see that the first set of variables (Experience, Experience^2 and Female) reported in Table 4.15 should be included in the wage equation because of having a probability of 1 in all the cases, except the Experience^2 term in Spain which gives a probability larger than 50%. Although this probability is not 1, the variable should still be included.[19]

[19] We have chosen 50% as the threshold for the probability of inclusion because is the middle value of the possible values a probability could take. There is no rule of thumb established for the probability to accept

Regarding the next set of variables, the Country of Origin, the probabilities of inclusion are smaller than 50% in all the European countries except for immigrants from developed and developing countries in Denmark and immigrants from developing countries in Spain. Although the descriptive statistics show that immigrants have lower wages in all the countries, this Bayesian estimation method suggests that immigrants from developed or developing countries do not seem to have different wages than natives in the analysed countries apart from Denmark holding other factors constant.

For the education variables that refer to the natives' returns to further levels of education in comparison with natives that attained low level of education or no education, we can see that the posterior probabilities of inclusion of these variables are higher than 98% in all four countries, which suggests that they are important variables to enter the wage equation. This is in accordance with the Human Capital theory and to the frequentist results presented in Table 4.10.

Lastly, a strong pattern could be seen regarding the interaction terms and the probabilities of inclusion. The only country that presents probabilities of inclusion higher than 50% is Spain. This occurs for the variable indicating differences in the returns between immigrants from developing countries and natives. The probability of inclusion for Mid. Ed. $\times$ Non-OECD and Tertiary Ed. $\times$ Non-OECD are 52.6% and 93.6%, respectively. These probabilities are high compared with the rest of the interaction coefficients in Spain and in other countries. For the rest of the countries, the BMA results suggest that there are no differences with the natives no matter the immigrant group and the level of education attained. This is in line with the frequentist

that the variables should be included. Because of this absence, a probability larger than 50% is chosen as a threshold to accept the inclusion of a variable in the model.

approach, where all the interactions (but in Spain) were not different from zero.

4.5.3 Discussion

The Bayesian approach allows us to take into account parameter and model uncertainty and to include prior information. The drawback of the frequentist approach is that there is a chance that the model selected is not the best to explain the dependent variable and model uncertainty is not taken into account. However, the Bayesian approach has a drawback as well; in this context, numeracy skills and time since arrival for immigrants cannot be included in the Bayesian analysis due to lack of relevant information for the priors. For the education variables, Bayesian statistics allow us to incorporate prior knowledge into the analysis in comparison with the frequentist analysis, which relies only on the information extracted from the data. In this chapter, we aimed to use the Bayesian approach to complement the frequentist estimation.

Bayesian statistics offer theoretical and practical advantages. Regarding the theoretical advantages, intuition plays a larger role in Bayesian statistics because, when working under this framework, the key focus is on predicting how good a model is rather than the significance testing carried out in frequentist statistics. Bayesian statistics allows for uncertainty to be incorporated in the model or even ignorance by setting non-informative priors. The posterior results may lead to convergence with the priors, or contrarily, diverge and lead to different conclusions.

A practical advantage of the Bayesian analysis in this context is that it could be applied even on a small sample of immigrants. The only thing we need to bear in mind is that the smaller the number of observations we have, the larger the influence the prior will have on the posterior parameters. In our analysis, the total estimation samples are large and this suggest the data will dominate the prior. As a consequence, the posterior for the natives' education coefficients is highly influenced by the data in all the countries, but not the coefficients regarding the difference in returns to education between natives and immigrants. Despite the large sample size in each country, the number of immigrants is relatively small.

In Bayesian statistics, it is assumed that every single parameter (β) has its own statistical distribution, while in the frequentist approach, a true fixed parameter is commonly assumed. If we acknowledge the existence of uncertainty in the parameters estimated, we can retrieve the posterior distribution of the parameters and obtain the likely values that these parameters take in each country. Because of heterogeneity in the sample of each country, the coefficients obtained are not of equal sizes across countries.

The Bayesian analysis also offers some other advantages over the frequentist method. It is well known that the returns to education are very sensitive to the estimation sample and the model functional form chosen for the analysis. Different functional forms will yield different coefficients. The Bayesian Average Modeling provides a specific probability indicating how important the variable is in the model once hundreds of models are visited by the algorithm, while when we run the simple OLS we assume a true model and an optimal selection of explanatory variables. One could run several robustness checks after running the frequentist analysis, but the BMA can complement the latter analysis in this context.

Although the majority of the interactions referring to the difference in returns to levels of education between natives and immigrants in the four countries are not statistically significant in the frequesntist and the Bayesian approach, it can be seen a pattern of a negative sign in the coefficients suggesting that this difference is negative in more than half of the coefficients. This indicates that even in the frequentist approach, controlling for skills does not remove the negative sign and other factors such as discrimination could drive these differences. Being these differences negative could place immigrants in a disadvantaged position compared to the natives. Coefficients for Denmark, France and Spain have the majority of negative signs while this pattern is not extensively found in the UK.

The availability of previous research regarding the parameters of interest allowed creating meaningful prior values for these parameters. We need to acknowledge that one of the main critiques of Bayesian statistics is subjectivity in the prior elicitation process. The prior parameters can be thought of as conveniently chosen. This problem was overcome by being transparent when explaining all the prior elicitation and offering reasons why the prior parameters are taking certain values, which means that the results can be replicated.

4.6 Conclusions

An application of Koop (2003)'s Bayesian methodology to estimate the returns to levels of education between natives and immigrants in four European countries was presented in this chapter. This analysis was performed to compare two approaches, the frequentist and the Bayesian, and see if they yield the same results.

Returns to levels of education and numeracy skills were estimated for natives and immigrants from developed and developing countries. Two frequentist models are presented. The first model controls for numeracy skills and the highest level of education attained and the reduced model only controls for the highest level of education attained. It was found that when controlling for numeracy skills the returns to levels of education decrease. Therefore, as expected from the existing literature, numeracy skills absorb part of the returns to levels of education.

Controlling for numeracy skills in the wage equations allows us to see if differences between immigrants' and natives' returns to levels of education disappear or they remain in which case these differences could be due to other factors. Regarding the countries analysed, the findings obtained in this chapter indicate that the returns to education and skills are different in each European country. Thus, in response to the research questions, it can be said that, when controlling for skills, the frequentist approach suggests that only in Denmark and Spain there are differences regarding natives with a low level of education and immigrants who attained higher education. Results presented both for France, a country that holds the assimilationist integration approach, and for the UK, a country with a multicultural pattern, suggest that there are no differences between natives' and immigrants' returns to levels of education.

In the UK, although natives and immigrants do not have different returns to levels of education, immigrants from developing countries have different (larger) returns to skills than natives. Similar to the UK, Denmark exhibits differences in the returns to skills. However, it is not clear which one of the mechanisms presented (discrimination, quality of schooling received or imperfect transferability of human capital) is contributing to a larger extent to explain the difference in returns. Unfortunately, the data used do not allow us to identify which particular factors are behind these differences.

When not controlling for numeracy skills and time spent in the country, the Bayesian estimation suggests that there are no differences between natives and immigrants' returns to education in Denmark, France and the UK. The Bayesian Model Averaging results support the frequentist estimates. In relation to the research questions presented, immigrants from developing countries are worse off than immigrants from developed ones but seem to be treated fairly in these countries.

Linking the results to the integration models we can see that Spain is the only country presenting differences in the returns to education between natives and immigrants when not controlling for numeracy skills and time spent in the country; at the same time, it is the only country out of the four analysed that does not have a clear integration approach. It would be tempting to conclude this chapter proposing a causal relationship for the observed results for Spain (lack of integration model leads to lower returns to education among immigrants). It should be noted, however, that both the observational design used here and the few variables available to build the model do not facilitate a causal interpretation of the results. Further research and better data would be needed to test such a causal relationship.

Important points and limitations are next offered. Immigrants' sample sizes in comparison with that of natives is relatively small: in each European country, less than 10% percent of the total sample are immigrants, which is representative of the real proportion of immigrants in the countries studied. The total sample size allows us to run both analyses, the frequentist and the Bayesian although they do not lead to the exact same results. The Bayesian analysis allows us to introduce prior information. This fact may be considered an advantage in comparison to the frequentist but, at the same time cause the Bayesian analysis not to give stable coefficients for the difference in the returns in this context.

In the earnings equations we control for categories of years since arrival in the host country. The analysis would be more robust if we could run different models separating the immigrants by years since arrival, but the small sample size does not allow for this. According to the pertinent literature, numeracy skills and returns to level of education would be expected to narrow between the natives and immigrants with more years in the country. Because of the small sample size, we need to assume equal returns to education for men and women too.

Another limitation is that we cannot differentiate between the reasons why the individual has decided to migrate: work, family, or humanitarian reasons. This could have an effect on the returns and wages. According to Damas de Matos (2014), in Europe, a large part of migration is due to work reasons.

Although there are some limitations mentioned above, this chapter contributes to the literature with the methodology used and the comparison of the two econometric approaches presented. There is no work published comparing the frequentist and the Bayesian approach and there is complete transparency in the subjective part of the Bayesian analysis when choosing the prior parameters. We cannot conclude that one approach is better than the other to estimate the returns in this context, but instead we can say that they complement each other.

Other than numeracy skills, such as non-cognitive skills (e.g. motivation or perseverance) could influence earnings and differ between natives and immigrants. Unfortunately, the availability of these measurements is very limited, and this hypothesis cannot be tested with the data used. The acquisition of these types of skills could also differ within cultural environments and could be related as well with the decision to migrate. The quality (measured by skills) of different cohorts of immigrants that move to the United States has been analysed in Borjas (1991). Individuals who have recently migrated, have in general attained higher levels of education than longstanding immigrants. This is related to the fact that, in the last decades, and in a majority of countries, educational attainment has risen. Related to this, Friedberg (2000) comments on the issue of potential "cohort quality". Unfortunately, in a single cross-section (which is used in this chapter), it is not possible to distinguish age and cohort effects since two or more cross-sections are needed to do so.

A possible extension of the analysis presented is to look at other countries with a greater sample of immigrants such as Canada or the United States or use the new data that the OECD has released for nine additional countries that conducted the survey in 2014-15, build new prior parameters and apply the Bayesian approach.

Chapter 4 Tables

Table 4.1: Descriptive Statistics

	Denmark		France		Spain		United Kingdom	
	(1)		(2)		(3)		(4)	
Log(Hourly Wages)	5.216	(0.295)	2.503	(0.333)	2.194	(0.450)	2.448	(0.456)
Experience	23.367	(11.551)	18.961	(11.587)	18.621	(10.847)	20.242	(11.864)
Experience^2	679.435	(572.736)	493.738	(493.412)	464.370	(483.525)	550.492	(542.730)
Female	0.423	(0.494)	0.410	(0.491)	0.406	(0.491)	0.387	(0.487)
Country of Origin								
OECD Country	0.024	(0.154)	0.021	(0.145)	0.018	(0.133)	0.019	(0.136)
Non OECD Country	0.055	(0.228)	0.073	(0.260)	0.066	(0.250)	0.101	(0.301)
Native	0.920	(0.270)	0.905	(0.292)	0.914	(0.279)	0.879	(0.325)
Education Levels								
Low Education Level	0.140	(0.347)	0.171	(0.377)	0.337	(0.472)	0.172	(0.378)
Medium Education Level	0.402	(0.490)	0.478	(0.499)	0.234	(0.423)	0.382	(0.486)
Tertiary Education	0.456	(0.498)	0.349	(0.477)	0.427	(0.494)	0.445	(0.497)
Numeracy Skills Score	290.315	(46.256)	265.116	(52.811)	259.512	(46.866)	275.029	(51.946)
Manage/Supervise	0.248	(0.432)	0.359	(0.479)	0.315	(0.464)	0.429	(0.495)
Industry Sector								
Agric., Fishing, Forest., etc.	0.013	(0.116)	0.012	(0.112)	0.034	(0.183)	0.002	(0.052)
Industry	0.197	(0.398)	0.200	(0.400)	0.152	(0.359)	0.185	(0.388)
Construction	0.069	(0.254)	0.084	(0.277)	0.077	(0.267)	0.055	(0.228)
Service Low	0.217	(0.412)	0.263	(0.440)	0.282	(0.450)	0.265	(0.441)
Service High	0.501	(0.500)	0.438	(0.496)	0.452	(0.497)	0.491	(0.500)
Economic Sector								
Public	0.340	(0.473)	0.249	(0.432)	0.284	(0.451)	0.267	(0.442)
Private	0.641	(0.479)	0.730	(0.443)	0.704	(0.456)	0.708	(0.454)
Non-Profit Org.	0.017	(0.132)	0.019	(0.139)	0.010	(0.102)	0.023	(0.152)
Type of Job Contract								
Indefinite	0.905	(0.292)	0.878	(0.327)	0.798	(0.401)	0.851	(0.355)
Fixed Term	0.072	(0.259)	0.074	(0.263)	0.146	(0.353)	0.094	(0.292)
Temporary	0.003	(0.059)	0.032	(0.178)	0.009	(0.097)	0.020	(0.142)
Apprenticeship	0.006	(0.079)	0.002	(0.052)	0.005	(0.074)	0.0008996	(0.029)
No contract	0.008	(0.092)	0.004	(0.067)	0.012	(0.111)	0.030	(0.171)
Other Contract	0.003	(0.058)	0.007	(0.083)	0.027	(0.163)	0.002	(0.049)
Time in the Country (only for immigrants)								
Less than 15 years	0.044	(0.206)	0.037	(0.190)	0.072	(0.260)	0.085	(0.279)
15 or more Years	0.035	(0.183)	0.057	(0.232)	0.012	(0.109)	0.035	(0.185)
Observations	3107		2676		1695		2971	

Standard Deviations are in brackets. Adjusted for population weights

Table 4.2: Denmark

	All	Natives	OECD Country	Non OECD Country
	(1)	(2)	(3)	(4)
Wage per Hour (kr)	192.607	194.297	192.470	164.393
Numeracy Skills Score	290.315	293.548	288.690	236.949
Low Education Level	14.05	13.83	8.68	20.23
Medium Education Level	40.25	41.27	22.96	30.96
Tertiary Education	45.70	44.91	68.36	48.81
Unwegithed Observations	3107	2614	166	327

Table 4.3: France

	All	Natives	OECD Country	Non OECD Country
	(1)	(2)	(3)	(4)
Wage per Hour (euro)	12.975	13.075	13.226	11.661
Numeracy Skills Score	265.116	269.641	213.606	224.286
Low Education Level	17.18	14.55	57.16	37.98
Medium Education Level	47.83	49.71	27.02	30.63
Tertiary Education	35.00	35.74	15.83	31.39
Unwegithed Observations	2676	2452	53	171

Table 4.4: Spain

	All	Natives	OECD Country	Non OECD Country
	(1)	(2)	(3)	(4)
Wage per Hour (€)	9.953	10.267	7.324	6.372
Numeracy Skills Score	259.512	263.100	227.956	218.996
Low Education Level	33.73	32.68	47.96	44.22
Medium Education Level	23.48	22.51	32.85	34.15
Tertiary Education	42.79	44.80	19.19	21.63
Unwegithed Observations	1695	1549	31	115

Table 4.5: UK

	All	Natives	OECD Country	Non OECD Country
	(1)	(2)	(3)	(4)
Wage per Hour (£)	12.955	13.056	12.532	12.162
Numeracy Skills Score	275.029	279.108	267.312	241.025
Low Education Level	17.26	17.74	10.24	14.42
Medium Education Level	38.22	39.64	44.67	24.61
Tertiary Education	44.52	42.61	45.09	60.97
Unwegithed Observations	2971	2701	65	205

Table 4.6: Ordinary Least Squares. Denmark

	With Num. Skills (1)		Without Num. Skills (2)	
Experience	0.014	***	0.015	***
	(0.0015)		(0.001)	
Experience^2	-0.000199	***	-0.000227	***
	(0.000030)		(0.00003)	
Female	-0.075	***	-0.093	***
	(0.0092)		(0.0093)	
Education Levels				
Medium Education Level	0.052	***	0.079	***
	(0.016)		(0.016)	
Tertiary Education	0.236	***	0.298	***
	(0.018)		(0.017)	
Numeracy Skills Score	0.070	***		
	(0.0063)			
Country of Origin times Time in the Country				
$<$ 15 Years$\times$OECD	0.026		-0.023	
	(0.079)		(0.073)	
$\geq$ 15 Years$\times$OECD	0.067		-0.038	
	(0.058)		(0.051)	
$<$ 15 Years$\times$ Non-OECD	-0.172	***	-0.195	***
	(0.058)		(0.055)	
$\geq$ 15 Years$\times$Non-OECD	-0.015		-0.105	***
	(0.041)		(0.038)	
Education Levels Interactions				
Mid. Ed. $<$ 15 Years$\times$OECD	-0.0617		-0.025	
	(0.092)		(0.085)	
Mid. Ed. $\geq$ 15 Years$\times$OECD	-0.214	***	-0.145	*
	(0.077)		(0.082)	
Mid. Ed. $<$ 15 Years$\times$ Non-OECD	0.111	*	0.079	
	(0.063)		(0.065)	
Mid. Ed. $\geq$ 15 Years$\times$Non-OECD	0.054		0.0694	
	(0.048)		(0.050)	
Tert. Ed. $<$ 15 Years$\times$OECD	-0.085		-0.0593	
	(0.110)		(0.084)	
Tert. Ed.$\geq$ 15 Years$\times$OECD	-0.181	**	-0.0537	
	(0.087)		(0.067)	
Tert. Ed.$<$ 15 Years$\times$ Non-OECD	0.108	*	0.0677	
	(0.064)		(0.063)	
Tert. Ed.$\geq$ 15 Years$\times$Non-OECD	-0.085	*	-0.0628	
	(0.049)		(0.049)	
Numeracy Skills Interactions				
Num. Skills.$<$ 15 Years$\times$OECD	-0.015			
	(0.036)			
Num. Skills. $\geq$ 15 Years$\times$OECD	0.075	*		
	(0.042)			
Num. Skills. $<$ 15 Years$\times$ Non-OECD	-0.045	**		
	(0.020)			
Num. Skills. $\geq$ 15 Years$\times$Non-OECD	0.016			
	(0.020)			
Observations	3107		3107	

Controls: Manage/Supervise, industry sector, economic sector and type of job contract. Standard errors are below the coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.7: Ordinary Least Squares. France

	With Num. Skills (1)		Without Num. Skills (2)	
Experience	0.021	***	0.021	***
	(0.001)		(0.001)	
Experience^2	-0.000257	***	-0.000271	***
	(0.00004)		(0.000041)	
Female	-0.0807	***	-0.098	***
	(0.011)		(0.012)	
Education Levels				
Medium Education Level	0.068	***	0.108	***
	(0.017)		(0.018)	
Tertiary Education	0.282	***	0.382	***
	(0.022)		(0.020)	
Numeracy Skills Score	0.080	***		
	(0.007)			
Country of Origin times Time in the Country				
$<$ 15 Years$\times$OECD	0.131		0.044	
	(0.081)		(0.050)	
$\geq$ 15 Years$\times$OECD	0.012		0.026	
	(0.160)		(0.110)	
$<$ 15 Years$\times$ Non-OECD	-0.0585		-0.082	**
	(0.054)		(0.038)	
$\geq$ 15 Years$\times$Non-OECD	0.020		-0.060	
	(0.068)		(0.057)	
Education Levels Interactions				
Mid. Ed. $<$ 15 Years$\times$OECD	-0.001		-0.013	
	(0.180)		(0.120)	
Mid. Ed. $\geq$ 15 Years$\times$OECD	-0.164		-0.127	
	(0.110)		(0.110)	
Mid. Ed. $<$ 15 Years$\times$ Non-OECD	-0.066		-0.083	
	(0.100)		(0.100)	
Mid. Ed. $\geq$ 15 Years$\times$Non-OECD	-0.003		-0.011	
	(0.055)		(0.052)	
Tert. Ed. $<$ 15 Years$\times$OECD	-0.218		-0.293	**
	(0.170)		(0.150)	
Tert. Ed.$\geq$ 15 Years$\times$OECD	-0.005		0.080	
	(0.170)		(0.130)	
Tert. Ed.$<$ 15 Years$\times$ Non-OECD	-0.096		-0.072	
	(0.089)		(0.080)	
Tert. Ed.$\geq$ 15 Years$\times$Non-OECD	-0.010		-0.035	
	(0.083)		(0.073)	
Numeracy Skills Interactions				
Num. Skills.$<$ 15 Years$\times$OECD	-0.065			
	(0.063)			
Num. Skills. $\geq$ 15 Years$\times$OECD	0.026			
	(0.048)			
Num. Skills. $<$ 15 Years$\times$ Non-OECD	0.009			
	(0.04)			
Num. Skills. $\geq$ 15 Years$\times$Non-OECD	-0.026			
	(0.032)			
Observations	2676		2677	

Controls: Manage/Supervise, industry sector, economic sector and type of job contract. Standard errors are below the coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.8: Ordinary Least Squares. Spain

	With Num. Skills (1)		Without Num. Skills (2)	
Experience	0.013	***	0.014	***
	(0.002)		(0.002)	
Experience^2	-0.00015	***	-0.00021	***
	(0.00005)		(0.00005)	
Female	-0.132	***	-0.161	***
	(0.017)		(0.017)	
Education Levels				
Medium Education Level	0.094	***	0.138	***
	(0.024)		(0.024)	
Tertiary Education	0.291	***	0.361	***
	(0.024)		(0.023)	
Numeracy Skills Score	0.093	***		
	(0.012)			
Country of Origin times Time in the Country				
$<$ 15 Years$\times$OECD	0.069		0.052	
	(0.075)		(0.067)	
$\geq$ 15 Years$\times$OECD	0.088		0.078	
	(0.180)		(0.130)	
$<$ 15 Years$\times$ Non-OECD	-0.088	*	-0.122	***
	(0.050)		(0.046)	
$\geq$ 15 Years$\times$Non-OECD	0.021		-0.008	
	(0.140)		(0.110)	
Education Levels Interactions				
Mid. Ed. $<$ 15 Years$\times$OECD	-0.225	***	-0.288	***
	(0.087)		(0.091)	
Mid. Ed. $\geq$ 15 Years$\times$OECD	-0.003		-0.025	
	(0.230)		(0.130)	
Mid. Ed. $<$ 15 Years$\times$ Non-OECD	-0.134	**	-0.155	**
	(0.067)		(0.067)	
Mid. Ed. $\geq$ 15 Years$\times$Non-OECD	-0.143		-0.146	
	(0.170)		(0.14)	
Tert. Ed. $<$ 15 Years$\times$OECD	-0.195		-0.224	
	(0.240)		(0.190)	
Tert. Ed.$\geq$ 15 Years$\times$OECD	-0.411		-0.459	***
	(0.250)		(0.130)	
Tert. Ed.$<$ 15 Years$\times$ Non-OECD	-0.166	**	-0.215	***
	(0.079)		(0.079)	
Tert. Ed.$\geq$ 15 Years$\times$Non-OECD	-1.062	***	-1.101	***
	(0.250)		(0.260)	
Numeracy Skills Interactions				
Num. Skills.$<$ 15 Years$\times$OECD	-0.012			
	(0.087)			
Num. Skills. $\geq$ 15 Years$\times$OECD	-0.075			
	(0.190)			
Num. Skills. $<$ 15 Years$\times$ Non-OECD	-0.036			
	(0.030)			
Num. Skills. $\geq$ 15 Years$\times$Non-OECD	0.001			
	(0.080)			
Observations	1695		1695	

Controls: Manage/Supervise, industry sector, economic sector and type of job contract. Standard errors are below the coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.9: Ordinary Least Squares. United Kingdom

	With Num. Skills (1)	Without Num. Skills (2)
Experience	0.026***	0.027***
	(0.002)	(0.002)
Experience^2	-0.0003***	-0.00042***
	(0.00004)	(0.00004)
Female	-0.107***	-0.144***
	(0.014)	(0.014)
Education Levels		
Medium Education Level	0.076***	0.134***
	(0.017)	(0.018)
Tertiary Education	0.298***	0.399***
	(0.020)	(0.020)
Numeracy Skills Score	0.113***	
	(0.007)	
Country of Origin times Time in the Country		
$<$ 15 Years×OECD	0.070	-0.069
	(0.140)	(0.089)
$\geq$ 15 Years×OECD	0.075	0.004
	(0.100)	(0.100)
$<$ 15 Years× Non-OECD	0.050	-0.149**
	(0.089)	(0.073)
$\geq$ 15 Years×Non-OECD	-0.003	-0.148
	(0.130)	(0.120)
Education Levels Interactions		
Mid. Ed. $<$ 15 Years×OECD	-0.204	-0.071
	(0.150)	(0.120)
Mid. Ed. $\geq$ 15 Years×OECD	-0.070	0.005
	(0.140)	(0.140)
Mid. Ed. $<$ 15 Years× Non-OECD	-0.079	-0.005
	(0.099)	(0.095)
Mid. Ed. $\geq$ 15 Years×Non-OECD	0.040	0.059
	(0.140)	(0.160)
Tert. Ed. $<$ 15 Years×OECD	0.015	0.131
	(0.170)	(0.120)
Tert. Ed.$\geq$ 15 Years×OECD	-0.023	0.057
	(0.150)	(0.130)
Tert. Ed.$<$ 15 Years× Non-OECD	-0.052	0.095
	(0.097)	(0.085)
Tert. Ed.$\geq$ 15 Years×Non-OECD	0.019	0.135
	(0.140)	(0.140)
Numeracy Skills Interactions		
Num. Skills.$<$ 15 Years×OECD	0.026	
	(0.058)	
Num. Skills. $\geq$ 15 Years×OECD	-0.002	
	(0.065)	
Num. Skills. $<$ 15 Years× Non-OECD	0.073**	
	(0.031)	
Num. Skills. $\geq$ 15 Years×Non-OECD	0.095	
	(0.059)	
Observations	2972	2972

Controls: Manage/Supervise, industry sector, economic sector and type of job contract. Standard errors are below the coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.10: Ordinary Least Squares.

	Denmark (1)		France (2)		Spain (3)		UK (4)	
Expericence	0.015	***	0.0213	***	0.014	***	0.0273	***
	(0.0015)		(0.001)		(0.002)		(0.002)	
Experience^2	-0.00023	***	-0.000273	***	-0.000205	***	-0.000430	***
	(0.00003)		(4.11e-05)		(5.72e-05)		(4.83e-05)	
Female	-0.093	***	-0.0979	***	-0.160	***	-0.144	***
	(0.009)		(0.011)		(0.017)		(0.014)	
Education Levels								
Mid Level	0.079	***	0.108	***	0.138	***	0.135	***
	(0.016)		(0.017)		(0.023)		(0.017)	
Tertiary Ed	0.298	***	0.382	***	0.361	***	0.399	***
	(0.017)		(0.020)		(0.022)		(0.019)	
Country of Origin								
OECD country	-0.030		0.039		0.056		-0.0272	
	(0.046)		(0.048)		(0.060)		(0.070)	
Not-OECD country	-0.137	***	-0.072	**	-0.105	**	-0.148	**
	(0.033)		(0.034)		(0.043)		(0.065)	
Interactions								
Mid Educ X OECD	-0.084		-0.108		-0.255	***	-0.051	
	(0.061)		(0.099)		(0.088)		(0.094)	
Tertiary X OECD	-0.057		-0.070		-0.263		0.088	
	(0.055)		(0.111)		(0.161)		(0.094)	
Mid Educ X Not-OECD	0.062		-0.036		-0.150	**	0.008	
	(0.041)		(0.050)		(0.063)		(0.084)	
Tertiary X Not-OECD	-0.007		-0.052		-0.296	***	0.103	
	(0.040)		(0.054)		(0.087)		(0.075)	
Observations	3107		2677		1695		2972	

Controls: Manage/Supervise, industry sector, economic sector and type of job contract. Standard errors are below the coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.11: Denmark. Prior and Posterior Means for β

	Priors		Posterior			HPDI
	Informative USA	Informative Denmark	Using Noninformative Prior	Using Informative Prior USA	Using Informative Prior Denmark	95% Credible Interval using Denmark's prior
	(1)	(2)	(3)	(4)	(5)	(6)
Education Levels						
Medium Education Level	0.333	0.165	0.079	0.112	0.100	[0.068,0.132]
	(0.070)	(0.070)	(0.015)	(0.016)	(0.016)	
Tertiary Education	0.555	0.385	0.297	0.330	0.321	[0.288,0.354]
	(0.085)	(0.085)	(0.016)	(0.017)	(0.017)	
Interactions						
Mid Ed. × OECD	-0.156	0.012	-0.084	-0.155	-0.057	[-0.178,0.064]
	(0.050)	(0.050)	(0.076)	(0.062)	(0.062)	
Tertiary Ed. × OECD	-0.260	0.028	-0.057	-0.130	-0.033	[-0.144,0.078]
	(0.050)	(0.050)	(0.069)	(0.057)	(0.057)	
Mid Ed. × Non-OECD	-0.209	-0.042	0.062	-0.010	0.031	[-0.048,0.110]
	(0.050)	(0.050)	(0.041)	(0.040)	(0.040)	
Tertiary Ed. × Non-OECD	-0.348	-0.094	-0.007	-0.078	-0.036	[-0.110,0.037]
	(0.050)	(0.050)	(0.038)	(0.038)	(0.038)	

Standard deviations in parenthesis. The control variables are the same used in the OLS estimation excluding numeracy skills and time in the country.

Table 4.12: France. Prior and Posterior Means for β

	Priors		Posterior			HPDI
	Informative USA	Informative France	Using Noninformative Prior	Using Informative Prior USA	Using Informative Prior France	95% Credible Interval using France's prior
	(1)	(2)	(3)	(4)	(5)	(6)
Education Levels						
Medium Education Level	0.333	0.330	0.108	0.125	0.117	[0.085,0.150]
	(0.070)	(0.070)	(0.016)	(0.017)	(0.017)	
Tertiary Education	0.555	0.550	0.381	0.403	0.394	[0.359,0.430]
	(0.085)	(0.085)	(0.018)	(0.018)	(0.018)	
Interactions						
Mid Ed. × OECD	-0.156	0.024	-0.107	-0.146	-0.059	[-0.201,0.084]
	(0.050)	(0.050)	(0.091)	(0.073)	(0.073)	
Tertiary Ed. × OECD	-0.260	0.040	-0.070	-0.161	-0.025	[-0.176,0.126]
	(0.050)	(0.050)	(0.100)	(0.077)	(0.077)	
Mid Ed. × Non-OECD	-0.209	-0.081	-0.036	-0.093	-0.055	[-0.150,0.040]
	(0.050)	(0.050)	(0.053)	(0.048)	(0.048)	
Tertiary Ed. × Non-OECD	-0.348	-0.135	-0.052	-0.119	-0.074	[-0.167,0.018]
	(0.050)	(0.050)	(0.052)	(0.047)	(0.047)	

Standard deviations in parenthesis. The control variables are the same used in the OLS estimation excluding numeracy skills and time in the country.

Table 4.13: Spain. Prior and Posterior Means for β

	Priors		Posterior			HPDI
	Informative USA	Informative Spain	Using Noninformative Prior	Using Informative Prior USA	Using Informative Prior Spain	95% Credible Interval using Spain's prior
	(1)	(2)	(3)	(4)	(5)	(6)
Education Levels						
Medium Education Level	0.333	0.316	0.138	0.149	0.146	[0.100,0.191]
	(0.070)	(0.070)	(0.023)	(0.023)	(0.023)	
Tertiary Education	0.555	0.632	0.361	0.373	0.371	[0.329,0.413]
	(0.085)	(0.085)	(0.021)	(0.021)	(0.021)	
Interactions						
Mid Ed. × OECD	-0.156	-0.080	-0.255	-0.209	-0.163	[-0.349,0.024]
	(0.050)	(0.050)	(0.139)	(0.095)	(0.095)	
Tertiary Ed. × OECD	-0.260	-0.160	-0.263	-0.253	-0.192	[-0.387,0.004]
	(0.050)	(0.050)	(0.152)	(0.100)	(0.100)	
Mid Ed. × Non-OECD	-0.209	-0.150	-0.149	-0.173	-0.156	[-0.281,-0.032]
	(0.050)	(0.050)	(0.073)	(0.064)	(0.064)	
Tertiary Ed. × Non-OECD	-0.348	-0.300	-0.296	-0.322	-0.306	[-0.440,-0.171]
	(0.050)	(0.050)	(0.080)	(0.069)	(0.069)	

Standard deviations in parenthesis. The control variables are the same used in the OLS estimation excluding numeracy skills and time in the country.

Table 4.14: United Kingdom. Prior and Posterior Means for β

	Priors		Posterior			HPDI
	Informative USA	Informative UK	Using Noninformative Prior	Using Informative Prior USA	Using Informative Prior UK	95% Credible Interval using UK's prior
	(1)	(2)	(3)	(4)	(5)	(6)
Education Levels						
Medium Education Level	0.333	0.085	0.151	0.149	0.143	[0.104,0.182]
	(0.070)	(0.070)	(0.020)	(0.020)	(0.020)	
Tertiary Education	0.555	0.425	0.399	0.420	0.412	[0.372,0.452]
	(0.085)	(0.085)	(0.020)	(0.021)	(0.021)	
Interactions						
Mid Ed. × OECD	-0.156	-0.026	-0.051	-0.187	-0.094	[-0.266,0.077]
	(0.050)	(0.050)	(0.151)	(0.088)	(0.087)	
Tertiary Ed. × OECD	-0.260	-0.130	0.088	-0.094	-0.006	[-0.173,0.161]
	(0.050)	(0.050)	(0.146)	(0.085)	(0.085)	
Mid Ed. × Non-OECD	-0.209	-0.043	0.008	-0.137	-0.066	[-0.200,0.069]
	(0.050)	(0.050)	(0.089)	(0.069)	(0.069)	
Tertiary Ed. × Non-OECD	-0.348	-0.217	0.103	-0.047	0.013	[-0.110,0.136]
	(0.050)	(0.050)	(0.080)	(0.063)	(0.063)	

Standard deviations in parenthesis. The control variables are the same used in the OLS except numeracy skills and time in the country.

Table 4.15: Bayesian Model Averaging Results

	Denmark			France			Spain			UK		
	BMA Post. Prob. (1)	Posterior Mean (2)	Posterior St. Dev. (3)	BMA Post. Prob (4)	Posterior Mean (5)	Posterior St. Dev. (6)	BMA Post. Prob. (7)	Posterior Mean (8)	Posterior St. Dev. (9)	BMA Post. Prob (10)	Posterior Mean (11)	Posterior St. Dev. (12)
Experience	1.000	0.015	0.0017	1.000	0.021	0.001	1.000	0.014	0.003	1.000	0.027	0.002
Experience^2	1.000	-0.0002	0.0000	1.000	-0.0003	0.000	0.943	-0.0002	0.0001	1.000	-0.0004	0.0000
Female	1.000	-0.093	0.010	1.000	-0.096	0.011	1.000	-0.157	0.017	1.000	-0.137	0.014
Country of Origin												
OECD Country	0.946	-0.082	0.028	0.021	0.0004	0.006	0.036	-0.002	0.015	0.018	0.0002	0.006
Non OECD Country	1.000	-0.123	0.020	0.410	-0.025	0.033	0.546	-0.081	0.082	0.384	-0.031	0.046
Education Levels												
Medium Education Level	0.999	0.082	0.015	1.000	0.101	0.016	1.000	0.137	0.025	1.000	0.135	0.020
Tertiary Education	1.000	0.298	0.016	1.000	0.385	0.017	1.000	0.365	0.023	1.000	0.413	0.020
Interactions												
Mid Ed. × OECD	0.054	-0.004	0.023	0.020	-0.0008	0.012	0.083	-0.015	0.062	0.027	-0.001	0.016
Tertiary Ed. × OECD	0.065	-0.004	0.020	0.023	-0.001	0.015	0.066	-0.012	0.059	0.029	0.001	0.015
Mid Ed. × Non-OECD	0.183	0.012	0.030	0.035	-0.001	0.011	0.526	-0.119	0.123	0.251	-0.033	0.063
Tertiary Ed. × Non-OECD	0.056	-0.002	0.012	0.161	-0.013	0.035	0.936	-0.310	0.127	0.064	0.004	0.027

Chapter 5 Conclusions

This thesis focuses on three relevant aspects of the Economics of Education literature: (1) the impact of tuition fees cap reform on degree choices in the UK; (2) the estimates of the returns to education in Spain by means of instrumental variables; and (3) the differential returns to education among natives and immigrants in several developed societies. The relevance of these three topics is proved by the fact that one of the most prestigious Handbooks of the discipline addresses them as specific chapters (Altonji et al. (2016), Heckman et al. (2006) and Smith (2006)). The evaluation of the higher education reform is an important topic considering the overall recent upward trend to increase university tuition fees in many developed countries. Returns to education has been a much studied issue since at least the 60s, and each year several papers are published containing additional methodological contributions, revisiting the already published work, and reassessing previous conclusions. The implementation of an instrumental variables approach to educational returns in Spain, something never done before using the Living Conditions Survey with Spanish data, seems to be a suitable strategy to effectively accomplish these tasks. Looking at the differences in human capital returns between immigrants and natives is a timely purpose considering the ongoing globalization processes: in an ever increasing globalized world, migratory movements across the entire planet and migrant populations in developed countries are expected to grow at an unprecedented rate. Thus it is interesting to evaluate and measure differences between both immigrants and natives that arise as a consequence of these migratory movements and compare their respective performance in the labour market related to their educational attainment.

Three different datasets were used for different contexts in order to tackle the three substantive challenges raised in this thesis: Higher Education Statistics Agency (HESA) students dataset in the UK for chapter 2; data from the Spanish Living Conditions Survey for chapter 3; and data from the OECD Programme for the International Assessment of Adult Competencies (PIAAC) survey for chapter 4. From a methodological point of view, non-linear difference-in-differences, two stages least squares, and frequentist plus Bayesian statistics are the econometric tools chosen to analyse the respective data.

5.1 Summary of the Thesis

The first contribution presented in chapter 2 estimates the effect of the 2012 tuition fees reform in the UK on degree choices by students. Chapter 3 presents the returns to levels of education in Spain supporting the validity of socioeconomic background characteristics as instrumental variables to take into account the potential endogeneity bias on the returns to education coefficients. The third contribution contained in Chapter 4 estimates returns to levels of education and numeracy skills comparing natives to immigrants in four different European countries. All the chapters are based on publically available micro-level data and on my own exploitation of the respective microdata bases.

Chapter 2 investigates the effect of the UK 2012 Tuition Fees cap reform on degree choices made by students. Two cross sections of students were observed to estimate that effect: the first one consists of students not affected by the reform; the second one is composed by those affected. A multinomial logistic regression was chosen to model the degree choice made by the student. To evaluate the effect of the reform, the difference-in-differences methodology was applied. Students domiciled in Scotland were chosen as the non-affected (e.g., control) group; students residing in England were taken as the treatment group. The reported results show that the affected group, i.e. English domiciled students, takes more into account, after the educational reform, the future employment prospects when choosing a degree subject. As a result, these treated students were around 2 percentage points less likely to study a degree in Arts & Humanities after the reform, while approximately 3 percentage points more likely to study a degree in Health & Life Science. As expected, it was shown that students with unemployed parents, with less economic resources available, are more affected by the reform than students with employed parents.

After analysing Spanish data, the second contribution contained in chapter 3 supports the idea of using socioeconomic background variables to correct for the endogeneity bias present when estimating the returns to education. Results show that, other things being equal, male returns to education are lower than female; the more education an individual attains, the higher his/her chances of participating in the labour market in comparison to individuals with no education. Interestingly, but in accordance with most results found in the relevant literature, the human capital of the mother has a larger impact on the education of children than of the father. In addition, results show that in Spain mother's education and public expenditure act as valid instruments to correct the endogeneity bias at the lower education levels while mother's higher education attainment and father's occupation are preferred to instrument offspring's higher education levels.

The contribution in chapter 4 presents the comparison of two statistically different approaches –frequentist and Bayesian– to the question of returns to education. The substantive goal here is comparing the relevance of acquired human capital for natives and immigrants in developed societies. Returns to levels of education and numeracy skills are reported for natives and immigrants in four European countries with relatively large fractions of migrant populations: Denmark, France, Spain and the UK. Differences in human capital returns between immigrants from different countries of origin and natives are investigated. Firstly, significant differences in the returns to education are found for the four observed countries. Secondly, in line with the pertinent literature, our results indicate that taking into account the numeracy skills variables makes the returns to levels of education decrease for all the groups and subgroups. However, no consistent pattern was found in the coefficients regarding the differences in the returns to levels of education between natives and immigrants in the four analysed countries. In Denmark, differences in the returns between immigrants from Non-OECD countries that have been residing in the host country for 15 or more years and natives are still present after controlling for numeracy skills. Similarly to what happens in Denmark, this effect is also observed in Spain, although it impacts on more groups of immigrants –in fact, all the immigrants from Non-OECD countries but the ones who spent 15 or more years in the destination country. Contrary to Denmark and Spain, in France, after controlling for numeracy skills, the differences in the returns to education for immigrants from an OECD country who spent less than 15 years in the country and natives disappear. On the other hand, the UK seems to exhibit a pattern different from the other three observed countries: significant differences between immigrants and natives educational returns were not found in the UK. In those cases such as Denmark or Spain where differences do exist after controlling for numeracy skills, it could be thought that discrimination or other unobserved factors produce these differences in returns to education. If this interpretation is correct, these factors would place immigrants in a worse economic position than natives with the same educational attainment.

5.2 Future Research

Chapter 2 includes a novel combination of econometric techniques that can be applied to analyse further tuition fees increases or either any reform or change as long as a sound counter-factual (a reasonable control group) is available and a multiple choice framework is given. Therefore, the content of this chapter could be a guide for further attempts of jointly using a non-linear model (multinomial logit) and an evaluation econometric technique (difference-in-differences). Whenever data are available and suit the context, other potential counterfactuals could be estimated, based on students not affected by the reform in the exercise of this chapter to perform a similar analysis. At any rate, investigating students' degree choices after the reform that increased tuition fees is relevant because it allows assessing the consequences of a specific higher education policy. In particular this sort of reform could directly affect the occupational structure of the labour market, either making it more or less concentrated upon some occupations.

Further research regarding chapter 3 could be developed if the next data release of the Living Conditions Survey in Spain is used. This next release contains a special questionnaire module aimed at collecting information about intergenerational transmission of socioeconomic status and retrospective living conditions background. This module is carried out each 6 years and, since the last one corresponds to the year

2011, a new one is about to be published in the near future. In the new dataset two cross sections would be available instead of one to further estimate changes in the educational returns and to re-test the validity of the instruments used here. New evidence is always welcome to verify previous hypotheses and to strengthen (or otherwise discard) our former interpretations.

Finally, chapter 4 could be expanded with the purpose of finding another suitable set of prior parameters for Bayesian statistical analysis. If new data are produced and analysed, and new papers are published estimating the returns to education by country of birth for immigrant populations, the prior parameters used here could be refined and updated and new posterior parameters retrieved. In this way the stability of the educational returns could be verified with more certainty than in the present attempt. Likewise, if further data become available from more countries in the OECD, the analysis presented here could be performed on these new national cases to check differences in the returns to education obtained between immigrants and natives. Additionally, with more data available it would be interesting to test several plausible explanations why differences in the returns to certain levels of education exist, ceteris paribus, between immigrants and natives.

Despite all the possible limitations of this work, some important contributions to methodology and knowledge of economics of education were presented.

Reference List

Abadie, A. (1997). Changes in spanish labor income structure during the 1980s: A quantile regression approach. *Investigaciones Economicas*, 21(2):253–272.

Alba-Ramirez, A. (1993). Mismatch in the spanish labor market: Overeducation? *Journal of Human Resources*, 28(2):259–78.

Altonji, J., Arcidiacono, P., and Maurel, A. (2016). The analysis of field choice in college and graduate school. *Handbook of the Economics of Education*, 5:305 – 396.

Angrist, J. and Krueger, A. (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics*, 106(4):979–1014.

Angrist, J. and Krueger, A. (1992). Estimating the payoff to schooling using the vietnam-era draft lottery. Technical Report 4067, National Bureau of Economic Research, Inc.

Antecol, H., Cobb-Clark, D. A., and Trejo, S. J. (2003). Immigration policy and the skills of immigrants to australia, canada, and the united states. *The Journal of Human Resources*, 38(1):192–218.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1):343–375.

Arrazola, M. and De Hevia, J. (2006). Gender differentials in returns to education in Spain. *Education Economics*, 14(4):469–486.

Arrazola, M. and De Hevia, J. (2008). Three measures of returns to education: An illustration for the case of Spain. *Economics of Education Review*, 27(3):266–275.

Aslam, M., Bari, F., and Kingdon, G. (2012). Returns to schooling, ability and cognitive skills in Pakistan. *Education Economics*, 20(2):139–173.

Aydemir, A. (2012). Skill based immigrant selection and labor market outcomes by visa category. Discussion Papers 6433, Institute for the Study of Labor (IZA).

Aydemir, A. and Skuterud, M. (2005). Explaining the deteriorating entry earnings of Canada's immigrant cohorts, 1966–2000. *Canadian Journal of Economics/Revue canadienne d'économique*, 38(2):641–672.

Barceinas, F. (2001). *Capital humano y rendimientos de la educación en México*. PhD thesis, Universitat Autònoma de Barcelona.

Barceinas, F. (2003). Endogeneidad y rendimientos de la educacin. *Estudios Econmicos*, 18(1 (35)):79–131.

Barr, N. (2004). Higher education funding. *Oxford Review of Economic Policy*, 20(2):264–283.

Barrett, G. F. (2012). The return to cognitive skills in the australian labour market. *Economic Record*, 88(280):1–17.

Becker, G. S. (1964). *Human Capital Theory*. Columbia University Press, New York.

Becker, G. S. (2009). *A Treatise on the Family*. Harvard:University Press.

Becker, G. S., Murphy, K. M., and Tamura, R. (1990). Human capital, fertility, and economic growth. *Journal of Political Economy*, 98(5, Part 2):S12–S37.

Beffy, M., Fougere, D., and Maurel, A. (2012). Choosing the field of study in postsecondary education: Do expected earnings matter? *Review of Economics and Statistics*, 94(1):334–347.

Beggs, J. J. and Chapman, B. J. (1988). Immigrant wage adjustment in australia: Cross section and time-series estimates. *Economic Record*, 64(3):161–167.

Berger, M. C. (1988). Predicted future earnings and choice of college major. *Industrial & Labor Relations Review*, 41(3):418–429.

Berman, E., Lang, K., and Siniver, E. (2003). Language-skill complementarity: returns to immigrant language acquisition. *Labour Economics*, 10(3):265–290.

Bertrand, M. and Mullainathan, S. (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *The American Economic Review*, 94(4):991–1013.

Biosca, O., Lenton, P., and Mosley, P. (2011). Microfinance non-financial services: A key for poverty alleviation? Lessons from Mexico. Technical Report SERP 2011021.

Blackburn, M. L. and Neumark, D. (1995). Are OLS estimates of the return to schooling biased downward? another look. *The Review of Economics and Statistics*, 77(2):217–230.

Blanchflower, D. G. and Oswald, A. J. (2004). Well-being over time in britain and the USA. *Journal of Public Economics*, 88(7):1359–1386.

Blau, F. D. and Kahn, L. M. (2000). Gender differences in pay. *Journal of Economic Perspectives*, 14(4):75–100.

Bleakley, H. and Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and Statistics*, 86(2):481–496.

Bonikowska, A., Riddell, W. C., and Green, D. A. (2008). *Literacy and the Labour Market: Cognitive Skills and Immigrant Earnings*. Statistics Canada Ottawa.

Borjas, G. (1991). Immigration policy, national origin, and immigrant skills: A comparison of canada and the united states. *National Bureau of Economic Research*, 3691.

Borjas, G. (1995). Assimilation and changes in cohort quality revisited: what happened to immigrant earnings in the 1980s? *Journal of Labor Economics*, 13(2):201–245.

Bound, J., Jaeger, D. A., and Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90(430):443–450.

Bowles, S. and Gintis, H. (1976). *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. New York: Basic Books.

Bratsberg, B., Hægeland, T., and Raaum, O. (2013). Immigrant skills and employment. cross-country evidence from the adult literacy and life skills survey. Technical Report 730.

Bratsberg, B. and Terrell, D. (2002). School quality and returns to education of U.S. immigrants. *Economic Inquiry*, 40(2):177–198.

Britton, J., Dearden, L., Shephard, N., and Vignoles, A. (2016). How English domiciled graduate earnings vary with gender, institution attended, subject and socio-economic background. IFS Working Papers W16/06, Institute for Fiscal Studies.

Brunello, G. and Miniaci, R. (1999). The economic returns to schooling for italian men. an evaluation based on instrumental variables. *Labour Economics*, 6(4):509–519.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge: University Press.

Carabaña, J. (1996). ¿ Se devaluaron los títulos? *Reis*, (75):173–213.

Card, D. (1999). The causal effect of education on earnings. *Handbook of Labor Economics*, 3:1801–1863.

Card, D. and Krueger, A. B. (1994). Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania. *The American Economic Review*, 84(4):772–793.

Carnoy, M. (1996). Race, gender, and the role of education in earnings inequality: An introduction. *Economics of Education Review*, 15(3):207–212.

Castles, S. (1993). *The Age of Migration : International Population Movements in the Modern World*. Macmillan.

Chevalier, A. (2011). Subject choice and earnings of uk graduates. *Economics of Education Review*, 30(6):1187–1201.

Chiswick, B. R. (1978). The Effect of Americanization on the Earnings of Foreign-born Men. *Journal of Political Economy*, 86(5):897–921.

Chiswick, B. R. and Miller, P. W. (2007). The International Transferability of Immigrants Human Capital Skills. IZA Discussion Papers 2670, Institute for the Study of Labor (IZA).

Chiswick, B. R. and Miller, P. W. (2009). Educational Mismatch: Are High-Skilled Immigrants Really Working at High-Skilled Jobs and the Price They Pay If They Aren't? IZA Discussion Papers 4280, Institute for the Study of Labor (IZA).

Chowdry, H., Crawford, C., Dearden, L., Goodman, A., and Vignoles, A. (2013). Widening participation in higher education: analysis using linked administrative data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(2):431–457.

Clark, K. and Lindley, J. (2009). Immigrant assimilation pre and post labour market entry: evidence from the uk labour force survey. *Journal of Population Economics*, 22(1):175–198.

Clarke, A. and Skuterud, M. (2013). Why do immigrant workers in australia perform better than those in canada? is it the immigrants or their labour markets? *Canadian Journal of Economics/Revue canadienne d'conomique*, 46(4):1431–1462.

Conseil, H. (1993). Lintégration à la française. haut conseil à lintégration. *Paris, Editions*, 10:18.

Corugedo, I., Pérez, E. G., and Pages, J. M. (1992). Educación y rentas. una aplicación a la enseñanza media en españa: Una nota. *Investigaciones Económicas*, 16(2):299–304.

Crawford, C. and Jin, W. (2014). Payback time? student debt and loan repayments: What will the 2012 reforms mean for graduates. *Institute for Fiscal Studies IFS Report*, 93.

Damas de Matos, A. (2014). Migrants' Skills: Use, Mismatch and Labour Market outcomes–A First Exploration of the International Survey of Adult Skills (PIAAC). *Matching Economic Migration with Labour Market Needs*, pages 249–312.

Davidson, R. (1993). *Estimation and Inference in Econometrics*. Oxford University Press.

Davies, E. and Lea, S. E. (1995). Student attitudes to student debt. *Journal of Economic Psychology*, 16(4):663–679.

Dearden, L., Fitzsimons, E., and Wyness, G. (2011). The impact of tuition fees and support on university participation in the UK. IFS Working Papers W11/17, Institute for Fiscal Studies.

Dearden, L., Machin, S., and Reed, H. (1997). Intergenerational mobility in britain. *The Economic journal*, 107(440):47–66.

Denison, E. F. (1962). *The sources of economic growth in the United States and the alternatives before us*. Committee for Economic Development.

Dickson, M. (2009). The Causal Effect of Education on Wages Revisited. IZA Discussion Papers 4419, Institute for the Study of Labor (IZA).

Dustmann, C. (1993). Earnings adjustment of temporary migrants. *Journal of Population Economics*, 6(2):153–168.

Feinstein, L. (2002). *Quantitative Estimates of the Social Benefits of Learning, 2: Health (Depression and Obesity). Wider Benefits of Learning Research Report.* London: Centre for Research on the Wider Benefits of Learning, Institute of Education,.

Ferrada, L. M. and Zarzosa, P. (2010). Diferencias regionales en la participación laboral femenina en chile. *Cuadernos de Economía*, 47(136):249–272.

Ferrer, A., Green, D., and Riddell, W. (2006). The effect of literacy on immigrant earnings. *Journal of Human Resources*, 41(2):380–410.

Freeman, R. B. (1996). Why do so many young american men commit crimes and what might we do about it? *Journal of Economic Perspectives*, 10(1):25–42.

Friedberg, R. M. (2000). You can't take it with you? immigrant assimilation and the portability of human capital. *Journal of Labor Economics*, 18(2):221–251.

García-Prieto, C., Martín Román, A., and Pérez Domínguez, C. (2005). Actual and potential returns to schooling in spain. *Education Economics*, 13(4):387–407.

Grao, J. and Ipiña, A. (1996). *Economía de la educación: Temas de Estudio e Investigación*. Departamento de Educación, Universidades e Investigación, Gobierno Vasco.

Green, D. A. and Riddell, W. C. (2003). Literacy and earnings: an investigation of the interaction of cognitive and unobserved skills in earnings generation. *Labour Economics*, 10(2):165–184.

Greene, W. H. (2003). *Econometric Analysis*. Pearson Education India.

Griliches, Z. (1977). Estimating the returns to schooling: some econometric problems. *Econometrica*, 45(1):1–22.

Hainmueller, J. and Xu, Y. (2013). Ebalance: a stata package for entropy balancing. *Journal of Statistical Software*, 54(1):1–18.

Hanushek, E., Schwerdt, G., Wiederhold, S., and Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, 73:103–130.

Harmon, C., Oosterbeek, H., and Walker, I. (2003). The returns to education: Microeconomics. *Journal of Economic Surveys*, 17(2):115–156.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.

Heckman, J., Lochner, L. J., and Todd, P. E. (2006). Earnings functions, rates of return and treatment effects: The mincer equation and beyond. *Handbook of the Economics of Education*, 1:307–458.

Heckman, J. and Polachek, S. (1974). Empirical evidence on the functional form of the earnings-schooling relationship. *Journal of the American Statistical Association*, 69(346):350–354.

HEFCE (2012). Recurrent grants and student number controls for 2012-13. Higher Education Funding Council for England. `http://www.hefce.ac.uk`.

HEFCE (2013). Higher Education in England: Impact of the 2012 reforms. Higher Education Funding Council for England. `http://www.hefce.ac.uk`.

Hernández, L. and Serrano, L. (2013). Efectos económicos de la educación en España: una aproximacón con datos. *PIAAC*, 2:66–88.

Hidalgo, M. A. (2010). A demand-supply analysis of the spanish education wage premium. *Revista de Economía Aplicada*, 18(54):57–78.

Hoogerheide, L., Block, J. H., and Thurik, R. (2012). Family background variables as instruments for education in income regressions: A bayesian analysis. *Economics of Education Review*, 31(5):515–523.

Hoogerheide, L. and Dijk, H. K. (2006). A reconsideration of the Angrist-Krueger analysis on returns to education. Technical Report EI 2006-15, Econometric Institute Research Papers.

Jacquemet, N. and Yannelis, C. (2012). Indiscriminate discrimination: a correspondence test for ethnic homophily in the chicago labor market. *Labour Economics*, 19(6):824–832.

Jimeno, J. F. and Fuente, A. (2005). The private and fiscal returns to schooling and the effect of public policies on private incentives to invest in education: a general framework and some results for the EU. *Documentos de trabajo del Banco de Espaa*, (9):1–82.

Kahn, L. M. (2004). Immigration, skills and the labor market: international evidence. *Journal of Population Economics*, 17(3):501–534.

Katz, L. F. and Autor, D. H. (1999). Changes in the wage structure and earnings inequality. *Handbook of Labor Economics*, 3:1463 – 1555.

Kirkeboen, L., Leuven, E., and Mogstad, M. (2016). Field of study, earnings, and self-selection. *Quarterly Journal Of Economics*, 131(3):1057–1111.

Koop, G. (2003). *Bayesian Econometrics*. Chichester: Wiley.

Koop, G. (2016). Bayesian methods for fat data. Strathclyde Business School. University Handout. `http://personal.strath.ac.uk/gary.koop/sgpe_bayesian.html`.

Lassibille, G. and Gomez, L. N. (1998). The evolution of returns to education in Spain 1980–1991. *Education Economics*, 6(1):3–9.

Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends(R) in Econometrics*, 4(3):165–224.

Leppel, K., Williams, M., and Waldauer, C. (2001). The impact of parental occupation and socioeconomic status on choice of college major. *Journal of Family and Economic Issues*, 22(4):373–394.

Levin, J. and Plug, E. J. (1999). Instrumenting education and the returns to schooling in the netherlands. *Labour Economics*, 6(4):521–534.

Lindley, J. and McIntosh, S. (2015). Growth in within graduate wage inequality: the role of subjects, cognitive skill dispersion and occupational concentration. *Labour Economics*, 37:101 – 111.

Machin, S. (2014). Developments in economics of education research. *Labour Economics*, 30:13–19.

Machin, S., Wyness, G., and McNally, S. (2013). Education in a devolved Scotland: a quantitative analysis. Technical Report 30, The London School of Economics and Political Science, Center of Economic Performance.

Mahnke, H. and Nymark, J. (2010). The expat study 2010. http://www.oxfordresearch.dk/media/47318/451901-THEEXPATSTUDY2010.pdf.

Maringe, F. (2006). University and course choice: Implications for positioning, recruitment and marketing. *International Journal of Educational Management*, 20(6):466–479.

MECD (2011). Datos y cifras del sistema universitario español: Curso 2010-2011. Catálago de publicaciones del ministerio. http://www.mecd.gob.es.

MECD (2012). Datos y cifras del sistema universitario español: Curso 2011-2012. Catálago de publicaciones del ministerio. http://www.mecd.gob.es.

Mincer, J. (1974). *Schooling, Experience, and Earnings*. Human behavior and social institutions. National Bureau of Economic Research; distributed by Columbia University Press.

Mosca, I. and Wright, R. E. (2010). National and international graduate migration flows. *Population Trends*, 141(1):36–53.

Murray, M. P. (2006). The bad, the weak, and the ugly: Avoiding the pitfalls of instrumental variables estimation. *Journal of Economic Perspectives*, 20(4):111–132.

OECD (2013). The survey of adult skills: Readers companion. http://dx.doi.org/10.1787/9789264204027-en.

Oliver, J., Raymond, J.-L., Roig, J.-L., and Barceinas, F. (1999). Returns to human capital in spain: A survey of the evidence. *Returns to Human Capital in Europe, Helsinki, The Research Institute of the Finnish Economy*.

ONS (2015). Population of the UK by country of birth and nationality: 2015. Office for National Statistics. https://www.ons.gov.uk/.

Oreopoulos, P. (2011). Why do skilled immigrants struggle in the labor market? a field experiment with thirteen thousand resumes. *American Economic Journal: Economic Policy*, 3(4):148–171.

Pastor, Raymond, R. and Serrano, L. (2008). ¿Es rentable invertir en capital humano en Españia? Análisis regional e internacional. *Investigaciones de Economía de la Educación*, 3:163–172.

Patrinos, H., Velez, E., and Psacharopoulos, G. (1994). Language, education, and earnings in Asuncion, Paraguay. *Journal of Developing Areas*, 29(1):57–68.

Pineda Herrero, P. (2000). Economía de la educación: una disciplina pedagógica en pleno desarrollo. *Teoría de la Educación*, 12:143–158.

Pino, F. J. B. (2014). Rendimiento privado de la educación y prima salarial en España. Un meta-análisis. *Investigaciones de Economía de la Educación*, 9:923–938.

Pons, E. and Gonzalo, M. T. (2002). Returns to schooling in Spain: how reliable are instrumental variable estimates? *Labour*, 16(4):747–770.

Portes, A. (1996). *Immigrant America : A Portrait*. University of California Press, Berkeley, California.

Psacharopoulos, G. (1993). Ethnicity, education, and earnings in Bolivia and Guatemala. *Comparative Education Review*, 37(1):9–20.

Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World Development*, 22(9):1325–1343.

Puebla, M. M.-C. and Velasco, M. S. (2006). La demanda de educación superior: un análisis microeconómico con datos de corte transversal. *Revista de Educación*, (339):637–660.

Puhani, P. (2000). The heckman correction for sample selection and its critique. *Journal of Economic Surveys*, 14(1):53–68.

Puhani, P. A. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear “difference-in-differences” models. *Economics Letters*, 115(1):85–87.

Purcell et al., K. (2008). Applying for higher education – the diversity of career choices, plans and expectations. findings from the first futuretrack survey of the ‘class of 2006’ applicants for higher education. *Higher Education Career Services Unit (HECSU), Universities and Colleges Admissions Service (UCAS) and IER*.

Ramos, M. I. A. and García-Crespo, D. (2003). Desajuste educativo y salarios en España. In *Economía de la educación AEDE XII: 18 y 19 septiembre*, pages 187–201. Universidad Carlos III.

Raymond, J. L. (2011). *¿Es Rentable Educarse? Marco Conceptual y Principales Experiencias en los Contextos Español, Europeo y en Países Emergentes*. Estudios de la Fundacin de las Cajas de Ahorros / Serie economa y sociedad ; 53. Madrid : FUNCAS. Enth. 4 Beitr.

Reher, D., Requena, M., and Sanz, A. (2011). ¿ España en la encrucijada? consideraciones sobre el cambio de ciclo migratorio. *Revista Internacional de Sociología*, 69(M1):9–44.

Requena, M. and Bernardi, F. (2005). *Tres Décadas de Cambio Social en España*. Madrid: Alianza.

Retortillo, Á., Ovejero, A., Cruz, F., Arias, B., and Lucas, S. (2006). Inmigración y modelos de integración: entre la asimilación y el multiculturalismo. *Revista Universitaria Deficiencias del Trabajo*, 7:123–139.

Rodríguez, M. A. D. and Albert, C. (2004). Salud, salarios y educación. *Hacienda pública española*, (169):11–34.

Sa, F. (2014). The effect of tuition fees on university applications and attendance: Evidence from the UK. IZA Discussion Papers 8364, Institute for the Study of Labor (IZA).

Salas, M. (2002). Cuatro décadas en economía de la educación. *Revista de Educación*, (328):427–449.

Schleicher, A. (2009). Securing quality and equity in education: Lessons from PISA. *Prospects*, 39(3):251–263.

Schultz, T. (1961). Investment in human capital. *The American Economic Review*, 51(1):1–17.

Shields, M. A. and Price, S. (2001). Language fluency and immigrant employment prospects: evidence from britain's ethnic minorities. *Applied Economics Letters*, 8(11):741–745.

Shields, M. A. and Price, S. W. (1998). The earnings of male immigrants in England: evidence from the quarterly LFS. *Applied Economics*, 30(9):1157–1168.

Simpson, H. and de Sheridan, A. (2014). Reforms in higher education = higher quality provision and better-informed choice? *CMPO Bulletin, Research in Public Policy*.

Smith, J. P. (2006). Immigrants and their schooling. *Handbook of the Economics of Education*, 1:155–187.

Smith, W. C. and Fernandez, F. (2017). Education, skills, and wage gaps in Canada and the United States. *International Migration*, 55(3):57–73.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374.

Staiger, D. and Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3):557–586.

Stock, J. H., Wright, J. H., and Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4):518–529.

Stock, J. H. and Yogo, M. (2005). *Testing for Weak Instruments in Linear IV Regression*, pages 80–108. Cambridge University Press.

Sweetman, A. (2004). Immigrant source country educational quality and canadian labour market outcomes. Statistics Canada, Analytical Studies Research Paper No 234.

Telles, E. M. and Ortiz, V. (2008). *Generations of Exclusion: Mexican-Americans, Assimilation, and Race*. New York:Russell Sage Foundation.

Thurow, L. C. (1975). *Generating Inequality: Mechanisms Of Distribution In The US*. New York: Basic Books.

Trostel, P., Walker, I., and Woolley, P. (2002). Estimates of the economic return to schooling for 28 countries. *Labour Economics*, 9(1):1–16.

Uusitalo, R., Honkapohja, S., and Koskela, E. (1999). *Essays in Economics of Education*. Department of Economics, University of Helsinki.

Valbuena, J. (2012). A longitudinal perspective on higher education participation in the UK. Studies in Economics 1215, School of Economics, University of Kent.

Vargas-Silva, C. and Markaki, Y. (2011). Long-term international migration flows to and from the uk. In *Migration Observatory, Oxford: University of Oxford*.

Vila, L.-E. and Mora, J.-G. (1998). Changing returns to education in spain during the 1980s. *Economics of Education Review*, 17(2):173–178.

Walker, I. and Zhu, Y. (2011). Differences by degree: Evidence of the net financial rates of return to undergraduate study for england and wales. *Economics of Education Review*, 30(6):1177–1186.

Walker, I. and Zhu, Y. (2013). The impact of university degrees on the lifecycle of earnings: some further analysis. *Department for Business, Innovation and Skills. Research Paper*, (No 112).

Wilson, R. A. and Homenidou, K. (2011). Working futures 2010-2020: main report. Technical Report 41, UKCES.