Human Capital Investment and Labour Market Response in Spain

Nuria Mallorqui Ruscalleda
Department of Social Science
Institute of Education, IoE
University College London, UCL

A thesis submitted for the degree of

Doctor of Philosophy (PhD)

February 9th, 2021

To my family

Declarations

I, Nuria Mallorqui Ruscalleda, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Word count (exclusive of appendix and bibliography): 83.526 words

Signature

Nuria Mallorqui Ruscalleda

February 9th, 2021

Acknowledgements

This thesis is the destination of a very long journey. Amazing people have shared parts of this journey with me, and some of them have even stayed by my side during all the progress. To all of them, thank you ever so much for being there and for sharing all your knowledge with me.

I am grateful to Dr. Andrew Jenkins, my main supervisor, for accepting tutoring my thesis. Your guidance, comments, corrections, and recommendations have really improved the final outcome. Thank you for your patience and comprehension and for letting me do it at my own pace. I also want to show my gratitude to professor Francis Green for his comments. They really help me to highlight and improve the thesis.

I am also thankful to the first supervisors I had in London. They guided my first and not very secure steps in the Doctorate in London. Professor Anna Vignoles, your guidance, corrections and advice were crucial for the upgrading. Professor Alfonso Miranda, I will never forget your Stata classes. Without your knowledge, help and suggestions this research would have been impossible. To both of you, thank you for your all your help and patience.

A big thank you to Ms. Catherine Haberfield, your assistance was crucial for the enrolling process, but above all, for finding accommodation at John Adams Hall. Thank you very much for your professionalism and commitment.

A very special thank is for my mentor and friend, Dr. Carmen Sarasúa. I cannot express my gratitude in words, but this thesis would never be finished without your unconditional dedication and encouragement. We have been working together since 2000, and I can just say it is a real pleasure to work beside you. You gave me the opportunity to participate in the academia as a job, being part of projects, journals or teaching. And you have always been there, supporting me, and believing in me. ¡*Muchas gracias! Sin tu permanente ayuda y apoyo esta tesis no existiría.*

I am very grateful to Dr. Roser Nicolau, you were the first university teacher who made me feel confident and who supported me. It was a long time ago, but I still remember your encouragements and help. Thank you ever so much for all your knowledge, time and comprehension. *Moltes gràcies!*

To all the members of the Economic History Unit from the Department of Economic and Economic History at the Universitat Autònoma de Barcelona, thank you to all of you for your help in the administrative paperwork and the teaching. A special thank you to Dr. Pedro Fatjó for sharing with me your teaching material. *Moltes gràcies a totes i a tots!*

I would like to thank Dr. Jan Paul Heisig, your time and assistance with Stata was crucial when I thought it was an endless road.

I am grateful to Dr. Cecilia Albert, your knowledge about the Labour Force Survey and the Spanish educational system and labour market helped me to develop this thesis. *Gracias por acogerme en la Universidad de Alcalá de Henares*.

The Centre for Demographic Studies, and its former Director, Professor Anna Cabré, were essential to have access to the individual data from the Spanish Labour Force Survey. Without their complicity, some data analysis would not have been possible.

I also thank the scholarship from the Ministerio de Educación for the mobility of doctorate students (2008-2009). It was the first stone to build this thesis, since it was my first contact with the IoE.

Apart from this thesis, my experience at IoE rewarded me with three great friends. Diana Sousa, Rukmen Sehmi, and Will Essilfie. You helped me go through the worst moments of the journey. Thanks Diana for our daily chats and teas. Rukki, our long chats trying to disentangle the data were very productive and stimulating. Will, thank you ever so much for making me feel that I am not alone in this world. To all and each one of you, thank you ever so much for being there.

Last, but not least, I owe this thesis to the support and encouragement of my family. My parents, Ramon Mallorquí (*in memoriam*) and María Carmen Ruscalleda, you gave me the opportunity to enter the university and to introduce myself to a new world. My brother, Enric Mallorquí, my sister, Carme Mallorquí, and my brother in law, Manel Navarro, you have always supported me. And my husband, Thierry Helleboid, with your unconditional help, support and comprehension this thesis would not be possible.

A tots vosaltres, moltes gràcies per tot el vostre suport i ajuda!

Abstract

The supply of highly educated workers has increased in most developed countries since the 1970s. However, in many countries the demand for these workers has not increased as rapidly as the supply, and education and skill mismatches have appeared. In Spain, the increase of university graduates started in the 1980s, and nowadays, Spain has one of the highest proportions of over-educated workers.

Focusing on Spain, this thesis provides new empirical evidence on the education mismatch for university graduates in the Spanish labour market. First, I analyse the trend in the likelihood of university graduates being over-educated from 1995 to 2018. Second, I assess the effect that including cognitive skills heterogeneity has on the probability of being over-educated. Third, I foresee three possible scenarios for the future trend of the proportion of university graduates over-educated in the Spanish labour market.

The results of the research demonstrate that first, the Spanish economy, and especially the private sector in Spain, has not adapted its jobs to the increase of the higher educated workers, especially for women. Second, accounting for skills' heterogeneity has little effect on the probability for university graduates being over-educated, even if it is small. Third, given the forecasted trend in the probability for a university graduate to be over-educated in the Spanish labour market some changes in the educational and labour market policies are needed to propel an improvement of the education match for university graduates in the Spanish labour market.

Regarding policy implications, the results of this research suggest modifying the supply of and facilitating the demand for highly educated workers. On the supply side, those degrees which are more demanded in the labour market should be promoted, whereas it is recommended to change the education programme in those degrees which are less demanded in the labour market. In this sense, for instance, introducing some work experience might help to increase its suitability in the labour market. On the demand side, skill-biased technological change increasing the demand for skilled workers should be promoted.

Impact Statement

This research, based on large-scale empirical data analysis, provides new evidence on the extent of education mismatches for university graduates in Spain and specifically on skill variation in these mismatches. The analyses in this thesis evidence the over-education problem for university-graduates in the Spanish labour market. It has been a problem in the past, it is being a problem nowadays, and it is expected to be a problem in the future. Clear policy implications follow from these results.

These results suggest modifying the supply of and facilitating the demand for highly skilled workers in the Spanish labour market to reduce the education mismatches, but also to improve the productivity of the economy. Therefore, in terms of policy implications, this research suggests focusing on educational and labour market policies.

Policies promoting the creation of highly skilled jobs should be supported. In this sense, policies should facilitate the creation of high-skill demanding industries or even the creation of technologic clusters to propel the creation of new high-skill demanding companies. Furthermore, public and private R+D investment might propel the creation of non-manual and non-routine jobs, promoting, therefore, jobs which require autonomy, decision making, and high skills.

Regarding educational policies, the employability of university graduates might be favoured involving university students in work experience. It would facilitate their specific training that combined with the general education from university would increase their job opportunities.

Table of Contents

Introduction..27

The objectives of this research 29

Mismatches in the labour market and their measurement..............................32

Women's participation in the labour market..34

Thesis structure 36

Chapter 1. Literature review 45

1.1. The economics of education in a historical perspective.......................46

1.1.1. From Classics to Human Capital Theory 46

1.1.2. Critiques of Human Capital Theory...52

1.1.3. General theories. Considering both sides of the labour market: supply and demand 58

1.1.4. The economics of education: final remarks 61

1.2. Education mismatches in the labour market 63

1.2.1. Concept and measurements...63

1.2.2. Empirical studies...76

1.3. Concluding remarks 88

Chapter 2. Educational level and labour market indicators in Spain 93

2.1. Introduction 94

2.2. Historical context..96

2.3. The Spanish labour market in an European context.........................100

2.4. Changes on the Spanish labour market over time............................117

2.4.1. Educational level and changes over time 117

2.4.2. Occupational levels and changes over time 127

2.5. Education match in the Spanish labour market 134

2.6. Spanish labour market at a glance 139

Chapter 3. To what extent has the Spanish labour market been able to accommodate the expansion of highly skilled workers?..........................143

3.1. Introduction 144

3.2. The supply of and demand for university studies in Spain................145

3.3. The source: The Spanish Labour Force Survey 153

3.4. Data *155*
3.5. Methodology *158*
3.6. Results *170*
3.6.1. Education mismatch in the Spanish labour market (1995-2018) 170
3.6.2. Probability for a university graduate to be over-educated in the Spanish labour market (1995-2018) 175
3.6.3. Predicted probabilities for a university-graduate to be over-educated in the Spanish labour market (1995-2018). 186
3.7. Has the Spanish labour market been balanced? *194*

Chapter 4. To what extent does skill heterogeneity explain education mismatch for university graduates in the Spanish labour market? 197
4.1. Introduction *198*
4.2. The source: The Survey of Adult Skills (PIAAC) *201*
4.3. Methodology *210*
4.3.1. Variation in individual's cognitive skills across countries 210
4.3.2. Measuring the cognitive skills effect on the probability of being over-educated 213
4.4. Results *226*
4.4.1. Variation in individuals' cognitive skills across countries 226
4.4.2. Measuring the cognitive skills effect on the probability of being over-educated 228
4.4.3. Conditional predicted probabilities for a 25-64-year-old university graduate to be over-educated 244
4.5. Has skills heterogeneity conditioned the likelihood of a university graduate being over-educated in the Spanish labour market? *248*

Chapter 5. The future employability of university graduates in the Spanish labour market 251
5.1. Introduction *252*
5.2. Sources, data, and variables *255*
5.3. Scenarios and methodology *259*
5.4. Over-education in the Spanish labour market: Future trends *272*
5.4.1. Stage 1: Results 273
5.4.2. Stage 2: Results 277

5.4.3. Stage 3: Results 281
5.4.4. Forecast results 284
5.5. An unexpected external shock 288
5.6. Conclusions 292
Chapter 6. Conclusions 295
6.1. Context 296
6.2. Sources, data, methodology and limitations 297
6.3. The results of the research 300
6.4. Policy implications 302
6.5. Further research 303
References 305
Appendix 1 321
Appendix 2 329
Appendix 3 351
Appendix 4 359

List of Figures

Figure 2. 1. Spanish labour force distribution by economic sector. 1950-2000. 98

Figure 2. 2. Distribution of population (25-64 years old) by country and educational level. 1995-2018 ... 103

Figure 2. 3. Activity rate by gender, 1995-2018 ... 109

Figure 2. 4. 25-64-year-old individuals by educational level and gender. Spain, 1995 and 2018 ... 122

Figure 2. 5. 25-64-year-old employees, by professional group and gender. Spain, 1995-2018. ... 129

Figure 3. 1. Map of the Spanish regions (NUTS-1) ... 167

Figure 3. 2. Education mismatch among university graduates in the Spanish labour market, 1995-2018. Three-year degree university graduates excluded (in percentages) ... 173

Figure 3. 3. Predicted probabilities for a university graduate to be employed and over-educated by age-group, nationality, and gender. Three-year degree university graduates excluded. Spain, 1995-2010 ... 190

Figure 3. 4. Predicted probabilities for a university graduate to be employed and over-educated by age-group, nationality, and gender. Three-year degree university graduates excluded. Spain, 2011-2018 ... 191

Figure 3. 5. Predicted probabilities for a 25-64-year-old university graduate to be employed and over-educated by activity sector and gender. Three-year degree university graduates excluded. Spain, 1995-2010 ... 191

Figure 3. 6. Predicted probabilities for a 25-64-year-old university graduate to be employed and over-educated by activity sector and gender. Three-year degree university graduates excluded. Spain, 2011-2018 ... 192

Figure 3. 7. Predicted probabilities for a 25-64-year-old university graduate to be employed and over-educated by region and gender. Three-year degree university graduates excluded. Spain, 1995-2010. ... 193

Figure 3. 8. Predicted probabilities for a 25-64-year-old university graduate to be employed and over-educated by region and gender. Three-year degree university graduates excluded. Spain, 2011-2018. ... 193

Figure 4. 1. Mean literacy and numeracy PIAAC scores by country. ... 205

Figure 4. 2. Dispersion between the highest and the lowest 5 percent PIAAC literacy and numeracy scores values by country. 207
Figure 4. 3. Gender difference in PIAAC literacy and numeracy proficiency scores by country. 208
Figure 4. 4. Mean of PIAAC literacy and numeracy proficiency scores for 25-64-year-old individuals by country 212
Figure 5. 1. Gross Domestic Product and university graduates employed 261
Figure 5. 2. Research and Development investment and university graduates employed in highly skilled jobs 263
Figure 5. 3. Scenarios and hypothesis for the projection 264
Figure 5. 4. Outline of the projection results.. 271
Figure 5. 5. 25-64-year-old university graduate men in Spain (in thousands). 273
Figure 5. 6. 25-64-year-old university graduate women in Spain (in thousands) 274
Figure 5. 7. 25-64-year-old university graduate men employed in the Spanish labour market (in thousands) 277
Figure 5. 8. 25-64-year-old university graduate women employed in the Spanish labour market (in thousands) 278
Figure 5. 9. 25-64-year-old university graduate men employed in HSJ in the Spanish labour market (in thousands) 282
Figure 5. 10. 25-64-year-old university graduate women employed in HSJ in the Spanish labour market (in thousands) 283
Figure 5. 11. 25-64-year-old over-educated university graduate men in the Spanish labour market (in thousands) 285
Figure 5. 12. 25-64-year-old over-educated university graduate women in the Spanish labour market (in thousands) 285
Figure 5. 13. 25-64-year-old over-educated university graduate men in the Spanish labour market (in percentage) 286
Figure 5. 14. 25-64-year-old over-educated university graduate women in the Spanish labour market (in percentage) 286
Figure A2. 1. Education mismatch among university graduates in the Spanish labour market. Three-year degree university graduates. Spain, 1995-2018 (in percentages). 340

Figure A2. 2. Predicted probabilities for a university graduate to be over-educated by age-group, nationality and gender. Three-year degree university graduates included. Spain, 1995-2010 348
Figure A2. 3. Predicted probabilities for a university graduate to be over-educated by age-group, nationality, and gender. Three-year degree university graduates included. Spain, 2011-2018 348
Figure A2. 4. Predicted probabilities for a university graduate to be over-educated by activity sector and gender. Three-year degree university graduates included. Spain, 1995-2010 349
Figure A2. 5. Predicted probabilities for a university graduate to be over-educated by activity sector and gender. Three-year degree university graduates included. Spain, 2011-2018 349
Figure A2. 6. Predicted probabilities for a university graduate to be over-educated by region and gender. Three-year degree university graduates included. Spain, 1995-2010 350
Figure A2. 7. Predicted probabilities for a university graduate to be over-educated by region and gender. Three-year degree university graduates included. Spain, 2011-2018 350

List of Tables

Table 1. 1. Summary of the main characteristics of the three measurements for over-education 74
Table 2. 1. Main labour market indicators by country. Total population, 1995, 2005, and 2018 106
Table 2. 2. Main labour market indicators by country. Men, 1995, 2005, and 2018 107
Table 2. 3. Main labour market indicators by country. Women, 1995, 2005, and 2018 108
Table 2. 4. Classification of educational levels groups 120
Table 2. 5. 25-64-year-old men by age group and education level. Spain, 1995 and 2018 (in percentages) 124
Table 2. 6. 25-64-year-old women by age group and educational level. Spain, 1995 and 2018 (in percentages) 125
Table 2. 7. Description of the classification of professional groups 128
Table 2. 8. 25-64-year-old men by age group and occupation. Spain, 1995 and 2018 (in percentages) 132
Table 2. 9. 25-64-year-old women by age group and occupation. Spain, 1995 and 2018 (in percentages) 133
Table 2. 10. Population by educational level, gender and occupation. Spain, 1995 and 2018 (in percentages) 135
Table 3. 1. Universities in Spain by public-private ownership, 1960-2019 147
Table 3. 2. University graduates contributing to the national insurance in the first year after graduation by gender, field of study, and university ownership, 2013-2014 (in percentages). 148
Table 3. 3. University graduates paying as working on university graduate jobs in the national insurance in the first year after graduation by gender, field of study, and university ownership, 2013-2014 (in percentages). 149
Table 3. 4. University students by university ownership 150
Table 3. 5. Education mismatch in the Spanish labour market by gender. Excluding three-year degree university graduates (in percentages) 171
Table 3. 6. Gender distribution of university graduates by type of degree (in percentages) 172

Table 3. 7. Estimated coefficients for the probit model. Three-year degree university graduates excluded. Spain, 1995-2010 176
Table 3. 8. Estimated coefficients for the probit model. Three-year degree university graduates excluded. Spain, 2011-2018 178
Table 3. 9. Estimated coefficients for the selection equation of the two-step Heckman model. Three-year degree university graduates excluded. Spain, 1995-2010 180
Table 3. 10. Estimated coefficients for the selection equation of the two-step Heckman model. Three-year degree university graduates excluded. Spain, 2011-2018 181
Table 3. 11. Estimated coefficients for the outcome equation of the Heckman two-step model. Three-year degree university graduates excluded. Spain, 1995-2010 184
Table 3. 12. Estimated coefficients for the outcome equation of the Heckman two-step model. Three-year degree university graduates excluded. Spain, 2011-2018 185
Table 3. 13. Predicted probabilities for a 25-64-year-old university-graduate being over-educated by year. Three-year degree university graduates excluded. Spain, 1995-2018 188
Table 4. 1. Main characteristics of the Spanish sample 203
Table 4. 2. Main characteristics of the sample by country 204
Table 4. 3. Observations' distribution by country 212
Table 4. 4. Main statistics of the sample 216
Table 4. 5. Pairwise comparison of the means of the literacy and numeracy scores by country. 227
Table 4. 6. Education mismatch in the Spanish labour market by gender and educational level. Statistical mean approach (in percentages) 228
Table 4. 7. Education mismatch in the Spanish labour market by gender and educational level. Subjective approach (in percentages) 229
Table 4. 8. Estimated coefficients of the probit model. BS and statistical mean approach 231
Table 4. 9. Estimated coefficients of the probit model. LS and statistical mean approach 232

Table 4. 10. Estimated coefficients of the probit model. NS and statistical mean approach ... 233
Table 4. 11. Estimated coefficients of the probit model of the selection equation [4.2]. BS ... 237
Table 4. 12. Estimated coefficients of the probit model of the selection equation [4.2]. LS ... 237
Table 4. 13. Estimated coefficients of the probit model of the selection equation [4.2]. NS ... 238
Table 4. 14. Estimated coefficients of the Heckman two-step model. BS and statistical mean approach ... 240
Table 4. 15. Estimated coefficients of the Heckman two-step model. LS and statistical mean approach ... 241
Table 4. 16. Estimated coefficients of the Heckman two-step model. NS and statistical mean approach ... 242
Table 4. 17. Predicted probabilities for a university graduate being over-educated in the Spanish labour market by gender, age-group, and educational level. BS ... 245
Table 4. 18. Predicted probabilities for a university graduate being over-educated in the Spanish labour market by gender, age-group, and educational level. LS ... 245
Table 4. 19. Predicted probabilities for a university graduate being over-educated in the Spanish labour market by gender, age-group, educational level. NS ... 246
Table 5. 1. Population and university graduates by age and year. Spain, 2014-2018 ... 256
Table 5. 2. Population by gender and year. Spain, 2014-2030 ... 257
Table 5. 3. NUTS-1 and NUTS-2 correspondence classification ... 258
Table 5. 4. Variation in the number of university graduates by gender and region. Spain, 2017-2018 (in percentages) ... 266
Table 5. 5. Variation in the number of university graduates employed by gender and region. Spain, 2017-2018 (in percentages) ... 268
Table 5. 6. Variation in the GDP by region. Spain, 2017-2018 (in percentages) ... 268

Table 5. 7. Variation in the number of university graduates employed in HSJ (adequately matched) by gender and region. Spain, 2017-2018 (in percentages) 270
Table 5. 8. Variation on the R+D investment. Spain, 2017-2018 (in percentages) 271
Table 5. 9. University graduates by gender and year. Spain, 2014-2030. 275
Table 5. 10. University graduates out of the population by gender and year. Spain, 2014-2030 (in percentages) 276
Table 5. 11. University graduates employed by gender and year. Spain, 2014-2030 (Hypothesis 1 – W1) 279
Table 5. 12. University graduates employed by gender and year. Spain, 2014-2030 (Hypothesis 2 – W2) 280
Table 5. 13. University graduates employed out of university graduates by gender, year, and hypothesis. Spain, 2014-2030 (in percentages) 281
Table 5. 14. University graduates employed in a HSJ (adequately educationally matched) by gender, year, and hypothesis. Spain, 2014-2030 (in percentages). 284
Table 5. 15. Over-educated university graduates out of university graduates employed by gender, year, and scenario. Spain, 2014-2030 (in percentages) 287
Table A1. 1. Educational Classification used by SLFS from 1995 to 1999 323
Table A1. 2. Educational Classification used by SLFS from 2000 to 2010 326
Table A2. 1. Educational Classification used by SLFS from 1995 to 1999 and correspondence in years of education 331
Table A2. 2. Educational Classification used by SLFS from 2000 to 2010 and correspondence in years of education 334
Table A2. 3. National Classification of Economic Activities (CNAE – 2009) 336
Table A2. 4. Education mismatch in the Spanish labour market by gender. Three-year degree university graduates included (in percentages). 339
Table A2. 5. Estimated coefficients for the probit model. Three-year degree university graduates included. Spain, 1995-2010 341
Table A2. 6. Estimated coefficients for the probit model. Three-year degree university graduates included. Spain, 2011-2018 342

Table A2. 7. Estimated coefficients for the selection equation [3.2] of the Heckman two-step model. Three-year degree university graduates included. Spain, 1995-2010 343
Table A2. 8. Estimated coefficients for the selection equation [3.2] of the Heckman two-step model Three-year degree university graduates included. Spain, 2011-2018 344
Table A2. 9. Estimated coefficients for the outcome equation [3.1] of the Heckman two-step model. Three-year degree university graduates included. Spain, 1995-2010 345
Table A2. 10. Estimated coefficients for the outcome equation [3.1] of the Heckman two-step model Three-year degree university graduates included. Spain, 2011-2018 346
Table A2. 11. Predicted probabilities for a 25-64-year-old university graduate to be over-educated by year. Three-year degree university graduates included. Spain, 1995-2018 347
Table A3. 1. Estimated coefficients of the probit model. Baseline specification and subjective approach 353
Table A3. 2. Estimated coefficients of the probit model. Literacy specification and subjective approach 354
Table A3. 3. Estimated coefficients of the probit model. Numeracy specification and subjective approach 355
Table A3. 4. Estimated coefficients of the Heckman two-step model. Baseline specification and subjective approach 356
Table A3. 5. Estimated coefficients of the Heckman two-step model. Literacy specification and subjective approach 357
Table A3. 6. Estimated coefficients of the Heckman two-step model. Numeracy specification and subjective approach 358
Table A4. 1. Forecast's results for 25-64-year-old university graduate men living in the North-West area 361
Table A4. 2. Forecast's results for 25-64-year-old university graduate women living in the North-West area 362
Table A4. 3. Forecast's results for 25-64-year-old university graduate men living in the North-East area 363

Table A4. 4. Forecast's results for 25-64-year-old university graduate women living in the North-East area 364
Table A4. 5. Forecast's results for 25-64-year-old university graduate men living in the Madrid area 365
Table A4. 6.Forecast's results for 25-64-year-old university graduate women living in the Madrid area 366
Table A4. 7. Forecast's results for 25-64-year-old university graduate men living in the Central area 367
Table A4. 8. Forecast's results for 25-64-year-old university graduate women living in the Central area 368
Table A4. 9. Forecast's results for 25-64-year-old university graduate men living in the East area 369
Table A4. 10. Forecast's results for 25-64-year-old university graduate women living in the East area 370
Table A4. 11. Forecast's results for 25-64-year-old university graduate men living in the South area 371
Table A4. 12. Forecast's results for 25-64-year-old university graduate women living in the South area 372
Table A4. 13. Forecast's results for 25-64-year-old university graduate men living in the Canary Islands 373
Table A4. 14. Forecast's results for 25-64-year-old university graduate women living in the Canary Islands 374

"The decisive factors of production in improving the welfare of poor people are not space, energy, and cropland; the decisive factors are the improvement in population quality and advances in knowledge"

Theodore W. Schultz (1981, p. 4)

Introduction

In most developed countries, post-compulsory education, and especially, tertiary education, became increasingly widespread in the second half of the 20th century. After the Second World War, the state started to play a major role in the economic activity. With regard to education, this typically meant that primary and secondary education became compulsory until the age of 14 or 16, and was wholly (or at least, partly) funded by the state. Later, the increase in living standards enabled the investment in higher (post-compulsory) educational levels. In the United States, for instance, the growth of university education was in the 1960s, in Spain in the 1980s.

Since 1980 the number of students enrolled in Spanish universities has increased greatly. On the one hand, the number of public universities grew and spread to smaller regions, making it easier for students to access higher educational levels. On the other hand, children from the Spanish baby boom of 1960-75 were the entrants to these universities, which supposed a rise in the total number of new students, but also an increase in the proportion of the age group going to university, especially for women.

Apart from enhancing their knowledge and broadening their understanding, for most of these new students, there were three main social and economic forces behind their decision to enrol at university. The first was the high levels of unemployment in the labour market which reduced their likelihood of securing a job. In 1990, the youth unemployment rate was 36.5 percent for 16-19-year-olds, and 30.1 percent for the 20-24-year-olds. By gender it was 30.8 percent for 16-19-year-old men and 42.6 percent for women, and 23.8 percent for 20-24-year-old men and 37.7 percent for women.[1]

A second factor was the relative low cost of studying a degree. In 1995, the expenditure per student on public and private universities in Spain was US $4,944 converted using purchasing power parity (PPP) and US $12,018 on average in the OECD.[2]

And third, the expectations of moving up the social ladder thanks to a better job and better opportunities in the labour market after graduating from university were an important motivating factor. University graduates have lower unemployment rates, and they generally receive higher wages and enjoy better labour market conditions. Regarding unemployment rates, for instance, in the second quarter of 2018 the unemployment rate for men with primary education was 25.9 percent and 30.6 percent for women, whereas for those holding university degrees the unemployment rate was 7.6 percent for men and 10.1 percent for women.

Indeed, at the beginning of the 21st century in Spain, university studies are still promoting upward social mobility, and furthermore, they reduce the likelihood of being part of the working class (Requena, 2016, p. 24).

---

[1] http://www.ine.es.
[2] https://read.oecd-ilibrary.org/education/education-at-a-glance-1998_eag-1998-en#page118.

Despite the generally positive returns for university graduates, by the second half of the 1990s, the Spanish labour market started to have some difficulty accommodating them. Education mismatch appeared in the Spanish labour market, as it had in other developed countries only some years before (e.g. in the United States in the 1970s, in the United Kingdom and France in the 1980s).

The objectives of this research

This research focuses on the employability conferred by a set of individual attributes, some of which are acquired in the educational system. Education (formal and informal) is an important part of human development, both as an individual and as a member of society. For individuals, education improves personal abilities and it also helps them to become more independent and better able to develop their capabilities within the society they are living in. As for society, well-educated people will likely contribute more to society (e.g. community work, democratic participation) as well as in the productivity of the economic system. Therefore, part of this knowledge will also be used in the labour force, especially nowadays when there is a constant and rapid technological upgrading and organisational change in the economy.

The Skill Biased Technological Changes hypothesis assumes a positive relationship between technological change and the demand for more skilled workers. However empirical research has demonstrated that this positive relationship has not always existed. The 19th century Industrial Revolution had a higher demand for unskilled workers, whereas in the 20th century the demand for skilled workers increased (Acemoglu, 1998; Goldin and Katz, 2007). Indeed, since the introduction of personal computers and other technological gadgets, and especially in more recent years, changes in job's requirements have demanded more versatile workers. Technology is constantly improving and changing both the sets of tools needed for working and the way workers use them. Therefore, nowadays workers should be able, first, to do different tasks either in the same occupation or in a different one, due to the organisational changes, and second, to learn quickly the new requirements of the workplace including the technological upgrading of skills.

Empirical studies have demonstrated that holding a university degree is related to lower rates of unemployment and it increases the ability of the worker to adapt to new tasks or increased demands on his or her skills. However, university graduates are also the most affected by education mismatch in the labour market. The higher the educational level, the higher the proportion of over-education.

The literature on education and skill mismatch has mainly focused on measuring them and their consequences for individuals and society such as lower wages, lower productivity, less health or happiness. Much less attention has been paid to the main characteristics of the over-educated or over-skilled workers.

This research aims to shed some light on the studies of the Spanish labour market. I particularly focus on university graduates and their likelihood of being over-educated in the Spanish labour market. Since the increase in the supply of and demand for university studies, the proportion of university graduates in the Spanish labour force and their employability in the Spanish production structure have changed significantly. To better understand the phenomenon and to contribute to planning future educational or labour active policies, *this research analyses the past, present, and future of the education mismatch for university graduates.*

Focusing, therefore, on the education mismatch for university graduates in the Spanish labour market, I analyse first, the probability of being over-educated during the last two decades; and second, to what extent variations in cognitive skills can explain differences in the probability of being over-educated. Furthermore, the research also includes forecasts for the near future and the medium term of trends in the proportion of over-educated university graduates in the Spanish labour market. This latter analysis aims to offer policy makers different possible scenarios of future trends in the education (im)balance for university graduates in the Spanish labour market. The intention is to draw likely scenarios which could be considered as a baseline for designing active educational and labour market policies to improve the education balance in the labour market, and therefore to reduce waste of both private and public resources, from both the economic and the personal points of view.

Regarding estimates of the likelihood for university graduates to be over-educated, these are expected to increase over time. The number of university graduates is constantly increasing and the Spanish data do not seem to show a growing demand for highly skilled workers, at least, recently. However, the most interesting aspect of this analysis is to identify the main characteristics of over-educated people. Knowing who the potential over-educated are will help policy-makers to adjust their policies to reduce education mismatch and its consequences. For instance, gender differences in the probability of being over-educated and the characteristics of over-educated men and women can be identified. Knowing about the variation by gender is essential for policy makers to decide whether a gender perspective should be introduced in policy-making.

The analysis of the effect of the variation in skills on the probability of being over-educated will help to understand the phenomenon of over-education and to disentangle whether an educational active policy improving the skills acquired in the educational system will help to reduce the effect of education mismatch. It is worth remembering here that Spain is one of the European countries with lower proficiency results in the Programme for International Student Assessment (PISA) and the Programme for the International Assessment of Adult Competencies (PIAAC) surveys.

Finally, using recent trends in the number of university graduates, the number of university graduates employed, and the number of university graduates employed in highly skilled jobs, I forecast the proportion of over-educated university graduates in the Spanish labour market for the coming decade. Forecasting the near and mid-term future of social and economic variables is a good exercise to prepare different policy strategies either to prevent the worst-case scenarios, as in this case there would be a significant increase in the proportion of over-educated; or to try to achieve the other possible more favourable outcomes, for instance, reducing the proportion of over-educated university graduates or even eliminating them. In any case, it is important for policy makers to get some baseline models to foresee the tendency that their policies could have to promote or to correct.

Mismatches in the labour market and their measurement

Research and policy interest in education and skill imbalances in the labour market have increased as more recent generations have become more highly educated. In the 1970s in the USA, and later in some European countries, the increment in the number of university graduates was neither followed nor preceded by an equal upgrade in the skills required by the jobs. The result was the emergence of skill (education) imbalances in the labour market.

The literature identifies three main sorts of mismatches between supply of and demand for workers in the labour market: education, skill, and subject or domain mismatch. Education mismatches measure the education imbalance of supply of and demand for workers in the labour market. Thus, it measures the different educational levels between the formal education level held by the individual and the one required by the workplace s/he secures. Skill mismatches refers to the deficit or excess of skills of the worker related to the ones demanded by the job. The area mismatch (also called subject or domain mismatch) is another sort of mismatch, and it occurs when the imbalances are not in the level of education but the field of study. This is the case, for example, of a person who holds a degree in chemistry and is working as a clerk in a bank office. In this latter case, the more specific the skills learnt in the degree, the more important the domain mismatch. In this thesis, I will focus on the first mismatch, education mismatch.

Education mismatch could be due to a surplus or a deficit of education. As a result, an over-(under-)educated worker is the individual who holds a higher(lower) education level than the one needed by the workplace s/he is securing. The level of education can be measured by the highest diploma of formal education held by the individual, but it is usually measured by years of education. Therefore, an individual will be educationally matched if they have been at school the same number of years as needed for doing the tasks of their job. Otherwise, s/he will be over-(under-)educated if s/he has been at school more(fewer) years than what is needed for doing the task of the job they secure.

Accounting for these educational mismatches is not an easy task. First, there are different ways of measuring them, none of them uncontroversial. It is not always easy either to convert the level of education with years of schooling or to exactly know the years of schooling needed for a specific job. Second, a perfect data set including a good education mismatch indicator is not available yet, making the calculation and analysis of this imbalance even more complicated. As a result, the dataset usually determines which of the measurements is the most adequate.

Even when it is possible to measure the labour market mismatch, their level varies across countries and over time. For instance, regarding education mismatches, in the 1970s, Freeman (1976) already noticed increasing levels of over-education in the USA labour market. In Spain, it was not until the 1990s that the labour market started to show some level of over-education, although it was still lower than in the USA (Alba-Ramírez, 1993).

Nowadays, all developed countries show signs of over-education, although their degree of mismatch is still at different levels. For instance, Flisi *et al.* (2014, p. 33) used data from the Survey of Adult Skills (PIAAC) to calculate different over-education measurements and concluded that the level of over-education ranges from 10 percent to 50 percent in France, from 16 percent to 27 percent in Italy, from 14 percent to 47 percent in the United Kingdom, and from 12 percent to 43 percent in Spain.

The main problem with this surplus of more highly educated individuals is that it has serious negative consequences for the individuals concerned as well as for the economy in general. Empirical literature has demonstrated that over-education (surplus of higher educated individuals relating to the demand for them) has a negative impact on individuals' wages, productivity, job satisfaction, or even on their health (Allen and van der Velden, 2001; Badillo Amador, López Nicolás and Vila, 2008; Green, McIntosh and Vignoles, 2002). For society, the main consequences are the loss of productivity and the "waste" of social investment, especially in those countries where an important part of tertiary education is funded by the state, for example in Spain, where nowadays 65 percent out of the universities are publicly funded.

Women's participation in the labour market

Historically, women's participation in the labour market has been limited, but it has also been underestimated by the official data (Blau, Ferber and Winkler, 2006; Sarasúa, 2018). But in the last forty years there has been an important increase in the proportion of women in paid work in most developed countries, in part due to higher female participation in the education system and, especially, in tertiary education.

Women usually experience worse conditions in the labour market than men: higher unemployment rates, higher proportion of fixed-term contracts, etc. Besides, women show greater differences between generations than men, but lower within them. More highly educated women are less representative in the oldest women's generations (50-65 years old), whereas in the youngest generations the proportion of women graduating from tertiary education equals and even exceeds men's. In 2012, 143.2 women per 100 men graduated from tertiary education as an average for the European Union before the accession of Croatia (EU-27), whereas in 1998 there were 121.7.[3] Among the youngest generations, the early leavers from school and training are lower among women than men (between 3 percentage points in the average of the 15 European countries or France, and 7 percentage points in Spain),[4] reducing the difference in the educational levels of women within the same generation.

Women's educational upgrading is one of the reasons for the higher participation of younger women in the formal labour market, but they are still being concentrated in certain sectors, occupations, and with worse conditions than men, for example, low wages, and a higher proportion of part-time and fixed-term contracts. The recent economic crisis has reduced the labour market gender gap,

---

[3] http://ec.europa-eu/eurostat/data/database.
[4] Ibid

although this has not been due to an improvement of the women's situation, but rather a deterioration for men (Malo, 2015).

As nowadays the proportion of men and women with higher educational levels is similar (about 50 percent of each gender), especially in the youngest generations, some researchers have argued that one of the reasons for these different outcomes is the lower number of women who enrolled and graduated with the university degrees with higher demand in the labour market, namely engineering and technology. Data seems to partly corroborate this, since in 2016, only 4 percent of the women graduates in tertiary education were from engineering, manufacturing and construction, whereas 11 percent of the men graduated in those areas.[5]

But, there are also other factors, including cultural and social factors, which also influence women's participation in the labour market. The constraints for balancing work and family life as well as gender differences in wages or in job conditions seem to be some examples of these factors.

In contrast, recent studies shed some light on behalf of women and defend an upward trend in the demand for women for being employed at executive level (CEDEFOP, 2018; Matias Cortes, Jaimovich and Siu, 2016). The assumption behind these findings is that these jobs are changing and nowadays, employers are looking for workers with higher cognitive and socioemotional abilities. This trend may be favourable to women (CEDEFOP, 2018; Matias Cortes, Jaimovich and Siu, 2016). Nonetheless, it remains to be seen whether this perception becomes reality, and more women actually achieve these top-level jobs and if so, whether it is under the same conditions as men in terms of wages, etc., and without renouncing their family life.

However, in analyses to date, gender has received less attention than it deserves. Part of the literature on education mismatch considers only men. The literature considering all the population, has usually used gender as a covariate to see the differences between men and women, but they have rarely been analysed as two

[5] Ibid

possible subsamples with different behaviour. This might be due to the lower women's participation in the labour market. In order to contribute to filling this gap in the literature, and taking based on the sources used in this thesis, *gender differences are analysed in the chapters of this thesis*.

Thesis structure

This dissertation is divided as follows. After this introduction, the **First Chapter** provides a survey of the **Literature**. The chapter starts by presenting the main theories related to education and the labour market. The first part of the chapter looks at studies focused on the value of education and the ability of the labour market to accommodate employees according to their qualifications. A selection of empirical studies for Spain as well as other countries are discussed here. The second section of the literature review chapter covers the definition of education mismatch and a discussion of their measurements, followed by a survey of the international and national empirical literature on education mismatch. A final remarks section finishes this literature review chapter by emphasising the main gaps which this thesis aims to fill.

The **Second Chapter** describes the changes in the educational level of the Spanish workforce and the changes in the main indicators of the Spanish labour market in the last decades of the 20th century and beginning of the 21st. This is the period of the largest expansion in the post-compulsory education system, and especially university studies, increasing the supply of higher educated workers in the Spanish labour market.

The Spanish labour market has had some structural problems since the end of the Franco dictatorship and the restoration of democracy in Spain in the second half of the 1970s. The Spanish production structure has been mainly work-intensive rather than capital-intensive. It was mainly based on the construction sector and low-skilled services (especially tourism). As a consequence, the creation of jobs has been very cyclical.

> *The crisis [2007] has also shown the weaknesses of the economic specialization of Spain in construction and also on a labour market model based on the employment of temporary workers without skills and often concentrated in low-paid low-quality jobs"* Muñoz de Bustillo Llorente and Antón Pérez (2011, p. 436)

In fact, the Spanish labour market has not been able to create enough jobs even when an economic expansion occurred. Until today, Spain has shown poorly labour market indicators compared to other European countries. Even nowadays, Spain is still one of the European countries with the worst labour market indicators: its high unemployment rates of 17.3 percent, only surpassed by Greece at 21 percent, and 10 percentage points above the average of the EU-28, which was 7.6 percent in 2017;[6] high temporary rates, the second worst in Europe with 26.8 percent of the total number of employees versus 14.3 percent average of the EU-28 in 2017, and only surpassed by Montenegro with 27.8 percent are only some examples.[7]

Using data from the Spanish Labour Force Survey (SLFS), Chapter Two reports the distribution of the labour force by educational level and by occupation. This contextual chapter is essential to comprehend the effect that the improvement of the educational level of the labour force had on the indicators of the Spanish labour market. Furthermore, data is presented by gender, making the identification of the different trends by gender straightforward. It allows, for instance, the calculation of a gender dissemination index to measure the gender occupational segregation in the Spanish labour market, and its yearly trend.

Detailed data are provided for 1995, 2005, and 2018 to see the changes and to be able to compare the beginning and the end of the period analysed in Chapter Three. The data clearly illustrate the human capital improvement of the Spanish labour force, especially for women. Besides, the analysis of this preliminary data suggests a possible saturation of the Spanish labour market caused by the continuing entry of new university graduates into the labour market and the low

[6] http://ec.europa.eu/eurostat/data/database.
[7] Ibid.

ability of the Spanish productive system to accommodate these university graduates.

Building on this analysis, **Chapter Three** focuses on the extent to which the Spanish labour market has been able to absorb the increasing number of Spanish university graduates over time. According to the data, those individuals with higher human capital investment (graduates from tertiary education) are those with the best labour market indicators (higher employment rates, higher wages, and better labour conditions). In contrast, workers with lower levels of human capital investment (less than secondary education) tend to be disadvantaged in the labour market. For example, in 2017, when the general unemployment rate for the 15-64-year-old population in Spain was 17.3 percent, unemployment was 10 percent for those with tertiary education (ISCED levels 5 to 8), and 17 percent and 25.2 percent for those with lower educational levels (ISCED levels 3 and 4, and ISCED levels 0 to 2, respectively).[8]

Despite the fact that on average university graduates achieve better outcomes in the labour market than people with fewer educational qualifications, this is one of the most interesting groups to analyse. The reason is their extraordinary increase during the last forty years of the 20th century, both in absolute numbers and in terms of representation in the total labour force, and their higher likelihood of being over-educated.

Understanding the functioning of the labour market is essential to know whether there is a mismatch in the labour market and if so, to what extent. It is especially important when there is over-education. It has been broadly demonstrated that when over-education is a permanent phenomenon it has negative effects since it suggests a waste of personal and social resources (García Serrano and Malo, 1996; Tsang and Levin, 1985).

Most literature on over-education has focused on the consequences of this mismatch on wages, productivity or worker satisfaction. But*, has over-education*

[8] International Standard Classification of Education (ISCED): http://ec.europa.eu/eurostat/data/database.

increased in the Spanish labour market over time? This is one of the questions that this thesis intends to answer. The core of Chapter Three discusses the probability of mismatch between supply of and demand for highly skilled workers, defined as people who hold at least a university degree. The data are from the Spanish Labour Force Survey from 1995 to 2018.

Studying the changes of the probability of being over-educated during a period of time one can test the ability of the labour market to absorb the increasing supply of university graduates. This analysis will also allow us to define which personal, geographical and time characteristics contribute to reducing the likelihood of being over-educated.

Using statistical methods, I measured the mismatch in the Spanish labour market for each year of the sample (from 1995 to 2018). I used the statistical mean approach to classify university graduates as over-educated, under-educated or adequately educated or matched. I selected those individuals with at least a university degree who are aged between 25 and 64 to analyse the changes in their likelihood of being over-educated. The results clarify the changes in the probability of being over-educated depending on the particular characteristics of the individual such as gender, age, region of residence, and year.

The econometric baseline model applied here is the probit model. It accounts for the probability for a university graduate to be over-educated. However, it does not control for the possible non-random selection of employed people. In order to account for a possible sample selection bias (only those who are actually employed can be (mis)matched in the labour market and they are likely to hold the same set of characteristics), a second model using a variation of the Heckman two-step model is also estimated.

Nevertheless, university graduates are not a homogeneous group. Their particular productivity depends on their innate or potential abilities but also on the skills acquired at university and work: their educational level and their experience. And, as Alba-Ramírez (1993) pointed out:

> *"workers can have the adequate amount of human capital by combining formal education, on-the-job training, and experience"*
> Alba-Ramírez (1993, p. 271)

Individuals would use their labour experience as well as the degree studied, be it humanities, science or technical, to improve their innate or potential abilities in order to adjust their personal characteristics to the demand of the labour market. In Chapter Three I assumed some homogeneity between individuals. To analyse the effect that heterogeneity can have on the probability of being over-educated, in **Chapter Four**, and using different data, I introduced a new variable accounting for the variation in cognitive skills to estimate its effect on the likelihood of being over-educated.

The objective of Chapter Four is to measure the effect that accounting for observable cognitive skills has on the probability for a university graduate to be over-educated in the Spanish labour market. To do so, I use data from the Programme for the International Assessment of Adults Competencies (PIAAC) which included proficiency scores for literacy and numeracy.

Indeed, PIAAC data detected that in 2011-2012, and despite the high levels of post-compulsory education in Spain, Spaniards have low levels of skill in comparison with other European Countries. PIAAC data also depict a significant disparity within the country's skills, the individual who scored the highest score value performing twice as well as the individual with the lowest score value.

PIAAC is a very complete survey that was carried out in 24 OECD countries. It assesses adult proficiency in different information-processing skills. Its main goal is to detect the ability of adults to successfully participate in the current economies and societies, their ability to manage all the information they receive at home, at work and in their social life. PIAAC also includes data of the basic demographic characteristics and background of the respondents, educational and labour status information and social participation. Therefore, it includes some broad information about several areas.

Using these PIAAC data, Chapter Four uses pairwise comparisons by country to resolve to what extent the use of the literacy or numeracy PIAAC scores might be useful to account for cognitive skill heterogeneity among individuals. The goal behind this analysis is twofold: first, seeing whether the variation in skills in Spain follows the same pattern as other European countries. Second, disentangling whether it makes sense to use the cognitive proficiency skills from PIAAC as a measure of variation in skills to estimate the predicted probabilities for a university graduate to be over-educated in the Spanish labour market when their cognitive skills are included in the model.

Once the variation in these skills among university graduates in the Spanish labour market has been demonstrated, it makes sense to use them as an indicator of skills heterogeneity. Therefore, and to measure the effect that including cognitive skills heterogeneity in the model has on the probability for a university graduate to be over-educated, I use the statistical mean approach to measure the degree of education mismatch in the Spanish labour market in 2011-2012. Then, I analyse the eligibility of the Heckman two-step model to use it first without accounting for the cognitive skills, then considering the literacy PIAAC scores, and a third time introducing the numeracy PIAAC scores. The three specifications are also run with a probit model in order to use it as a baseline. Furthermore, all the specifications of both models, probit and the Heckman two-step model, have also been run using the subjective approach as a measurement of over-education.

As in Chapter Three, in Chapter Four, the predicted probabilities of success conditional to the selection equation, that is, the predicted probability for university graduates to be over-educated conditional to the probability of being employed in the labour market has been calculated for different variables and by gender. The results, with or without the cognitive proficiency scores, were compared to determine whether accounting for these observable skills has a negative effect on the predicted probabilities of being over-educated, that is to say, whether those higher-skilled workers with higher literacy and numeracy skills are less likely to be over-educated in the Spanish labour market. Overall, the comparison of the different results of the specifications and models are used to verify their robustness.

Chapter Five looks to the future. It aims to offer a baseline model for policy makers to foresee possible future educational or active labour market policies. The education imbalance in the Spanish labour market is discussed throughout the thesis. It was a problem in the past, it is a problem now, hence the last element that remains to discover is whether it will be a problem lasting for a long time.

Using data from the SLFS, from 2014 to 2018, and the demographic projection of population of the National Statistics Institute (INE), I estimated the proportion of over-educated university graduates that would be in the Spanish labour market in the next ten years. A regression was run to estimate the latest trend of the variables to use as a baseline to construct different foreseen scenarios.

The forecast is divided into three different stages: first, I calculate the number of university graduates in the next ten years, until 2030. In the second stage, I use the outcome of the first stage to calculate the number of university graduates employed. And equally as the previous stage, in the third stage, I use the number of university graduates employed obtained in the second stage to calculate the number of university graduates employed in highly skilled jobs. From these three outcomes, I finally extracted, for each scenario, predictions for the trend that the proportion of over-educated university graduates could plausibly follow in the next ten years. In all stages, data have been projected by single ages, gender, and regions (NUTS-1)

Three scenarios are forecast. In all scenarios, the number of university graduates is projected following the same tendency as in the last years (2014-2018). The difference is therefore marked by population movement as well as by the number of university graduates employed overall, and those employed in highly skilled jobs. Scenario 1 accounts for the continuity of the trend of the last years in both cases, the number of university graduates employed and the number of university graduates employed in highly skilled jobs. Scenario 2 accounts for an increase of half of the year-on-year variation of the GDP by region in the number of university graduates, and Scenario 3 considers besides this increase in the number of university graduates employed, a growth in the number of university graduates employed in highly skilled jobs. In this case, the investment in R+D has been used as a reference point.

To take account of the effects of the COVID19 on the Spanish labour market, Chapter Five ends with a section discussing its possible effects on the projected scenarios. It is well known that this pandemic is affecting the whole world and that its economic consequences could be even worse than the last economic crisis of 2007 or even the Great Depression of 1929.[9] Indeed, the pandemic has already started to affect national economies, its effect is not yet over, and nobody knows how long it will last. Therefore, it seems sensible to at least try to take account of how this unexpected external shock could affect the ability of the Spanish labour market to accommodate the labour force in the next years.

Finally, the **Conclusion** chapter sets out the main contributions as well as the limitations of this research. This section evaluates policy recommendations to improve the worker's education balance in the labour market, as well as, some reflections in relation to the gender differences in the labour market.

[9] In this sense, Barry Eichengreen, Professor of Economics and Political Science at the University of California-Berkeley, tweeted on 3rd April 2020 that "We have done in a fortnight what took them a year", comparing data from the growth of the unemployment rate in the US during the Great Depression of 1929 and these later weeks. (https://twitter.com/b_eichengreen?lang=ca)

> *"Economic theory has in recent years been extend to explain the accumulation of human capital, and the price and income effects of this form of capital. The theory has led to important new approaches in bringing economics to bear on human behaviour"*
>
> *Theodore W. Schultz (1981, pp. 79-80)*

Chapter 1. Literature review

Studies in the value of education and the ability of the labour market to accommodate workers according to their educational level are the focus of this chapter. The main theories of the Economics of Education, including the origins of the field and the main theoretical frameworks are presented first. The section on labour market mismatches discusses education mismatches, the concept and its different measurements. It is followed by the empirical literature on this sort of mismatches. A final section reviews the gaps in the literature on education mismatches in the labour market which will be considered in this research.

1.1. The economics of education in a historical perspective

1.1.1. From Classics to Human Capital Theory

Historically economists have identified education as a significant factor accounting for the differences among individuals' capabilities and their incomes. Adam Smith already introduced the idea that part of the difference in human abilities

> *"seems to arise not so much from nature, as from habit, custom, and education"* Smith (1961 [1776], pp. 19-20)

According to Smith,

> *"The difference of natural talents in different men is, in reality, much less than we are aware of; (…) the difference of talents comes then to be taken notice of, and widens by degrees."* Smith (1961 [1776], pp. 19-20)

Indeed, two main contributions by Smith will be used during the development of this thesis. First, the importance of improving individuals' innate abilities. In this sense, education and training are two of the main ways to develop individuals' capabilities. Second, individuals have different abilities. Some individuals are better at numeracy and others are better at literacy, the arts or other fields. Therefore, in this sense, some of Smith's ideas will be seen in the following pages.

Moreover, following Smith's idea, other classical economists such as John Stuart Mill (Mill, 1968 [1848]), Thomas Malthus (Malthus, 1968 [1836]) and neo-classical ones notably Alfred Marshall (Marshall, 1920 [1890]) considered investment in education to be valuable for the future occupational prospects of the individuals, as well as their future earnings. This was, hence, the conception of the idea that education was an investment and not an expenditure. As an investment, education will be rewarded later by the labour market. Indeed, as is well known, the economics of education is based on the idea that improving individuals' knowledge and individuals' abilities is an investment for getting better conditions such as a permanent contract, and higher returns, pecuniary or not, in the labour market.

Likewise, at the beginning of the 20th century, Irving Fisher stated that

> *"When a young man studies law, medicine, journalism, music, or prepares for any other profession, he is investing in his own person, with the hope that the sums thus invested may ultimately be returned to him (with interest). The same is true of physical training."* Fisher (1965 [1906], p. 170 §4).

These authors highlighted the relationship between more education or training and higher future earnings. However, they did not go more deeply into the reasons for the existence of this positive relationship between more training and more returns.

In this sense, Pigou, in *A Study in Public Finances* in 1929*,* clearly developed the idea that financing education or training should not be considered as an expenditure, but as an investment. He also defended the view that those individuals who have higher educational levels or more training will be rewarded in their future earnings. However, Pigou does focus on the relationship between acquiring more education or training and being more productive as the cause of this positive relationship:

> *"There is such a thing as investment in human capital as well as investment in material capital. So soon as this is recognised, the distinction between economy in consumption and economy in investment becomes blurred. For, up to a point, consumption is investment in personal productive capacity"* Pigou (1929, p. 29).

Nevertheless, it was not until the middle of the 20th century that education was transformed into one of the key concepts in the economy. The interest of economists in how education influenced the economy has progressively increased since then. Part of this interest was due to the inability of the theoretical models of the time to explain the surprising economic recovery of the European countries which had fought in the Second World War (Nuñez, 1993; Psacharopoulos, 1996). This is the main reason why a large proportion of the first-stage studies about the influence of education on the economy were macroeconomic analyses: studies about the influence of education on economic growth or development (Bowman, 1970; Englader and Gurney, 1994; Psacharopoulos, 1984).[10]

Robert M. Solow (Solow, 1957) was one of the authors who, focusing on the macroeconomic perspective, established a relationship between skills (human capital) and economic growth. He argued that there is an unknown residual factor (related to human capital and skills) with a great influence (positive correlation) on economic growth. Whether the improvement in the educational level of individuals comes out in an increase in their productivity (i.e. Pigou) or just in higher future earnings (i.e. Fisher, Marshall or Malthus), in both cases, it will result in economic growth even if the positive consequences will be higher in the former.

In the first case, if there is an increase in individuals' productivity, it will be translated in a growth in the production capacity of the production system, and therefore, an increase in the supply of goods and a reduction in their prices, and as a consequence, an increase in the demand for workers due to the increase in the wages related to the productivity increase. The final consequence would be economic growth. In the second case, the increase in the future earnings, even if there is not an increase in the productivity would lead to an increase of either the expenditure or the savings of the workers due to their higher wages related to their higher educational level. In short, in both cases, the economy will finally grow.

It was just some years later, in 1961, that Theodore W. Schultz (Schultz, 1961) contributed to settling the concept of human capital and its positive correlation with education in core research. He devoted his opening discourse as President at the 73rd Annual Meeting of the American Economic Association to the investment in human capital.[11] He put together the studies of previous economists to develop one unique argument which includes the positive relationship between improving individuals' skills and the growth of both personal income and national economic wealth. He first highlighted the importance of

---

[10] An important demand for this kind of studies was from international institutions such as the UN, UNESCO, European Commission or OECD. See, for instance, IRDAC (1991).
[11] Discourse reproduced in *The American Economic Review*, 1 (Vol. LI, March 1961).

investment in human capital to explain the part of the increase in national output which cannot be explained by other factors (e.g. physical capital); and second, the increase in real income for educated (trained) workers.

> *"Laborers have become capitalist not from a diffusion of the ownership of corporation stocks, (…), but from the acquisition of knowledge and skill that have economic value. This knowledge and skill are in great part the product of investment and, (…), predominantly account for the productive superiority of the technically advanced countries."* Schultz (1961, p. 3)

Despite all these precedents, **Human Capital Theory** (Becker, 1962; Becker, 1964) is considered the first economic theory based on the study of the investment of individuals in their human capital, and a central contribution to the microeconomic analysis of education. Based on the ideas of classical economic theory, Becker's theory emphasises the positive relationship between higher skills, as a consequence of higher educational levels, more training or more experience in the labour market, and the improvement of individuals' productivity. Moreover, this higher productivity will also be rewarded in the labour market with higher wages and better job conditions.

According to classical economic theory, in all markets, there is perfect competence and individuals behave rationally and try to maximise their utility, which reflects their preferences. When these conditions are satisfied, the price of a good is determined by its marginal productivity. Following the classical economic theory and focusing on education and its rewards on the labour market, Human Capital Theory states that the worker´s abilities or skills determine their productivity; and their productivity is what fixes their wage. Consequently, each individual will desire to increase their productivity as much as possible in order to maximise their returns (wage). They do so taking into account the discount tax of their investment and the opportunity cost of improving their abilities (increasing their educational levels or acquiring more experience in the labour market), since they have to decide in the present while taking the future earnings they expect to receive into account.[12]

In other words, according to Human Capital Theory the real productivity of an individual is fixed by their abilities and by their investment in skills. Individuals have potential abilities determined by their innate characteristics, which are more or less developed, depending on the individuals' investment in training or education. Hence, in practice, individuals can better their abilities by improving their education (formal or informal) or acquiring more experience (e.g. on-the-job training). The main problem, however, is that individuals' characteristics are not the only factor which will determine the wage. Other factors, such as the job characteristics, the supply of and demand for individuals with the same educational level, or even, cultural factors can also affect wages.

Becker's model also attempts to take into account the influence of some of these factors, as for instance, the impact that motivation and the job intensity can have on the employees' productivity.

> *"The productivity of employees depends not only on their ability and the amount invested in them both on and off the job but also on their motivation, or the intensity of their work"* Becker (1964, p. 36).

Becker's theory was turned into a mathematical formulation by Mincer (1962). Mincer's formulation allows researchers to calculate returns on education in terms of increases in wages due to higher educational levels or more experience, among other variables. Nevertheless, this formulation, as will be seen in the analysis section of this thesis, implies some homogeneity of skills among individuals holding the same degree or the same level of education, ignoring the possibility of variation in skill, either innate or acquired. This is the case, for instance, in the third chapter of this thesis. In contrast, Chapter Four attempts to provide new evidence on the effect that skills heterogeneity has on an individual's labour market returns, in particular, in the probability of being over-educated.

---

[12] According to classical economic theory there is perfect information in the market, and individuals know the discount tax of their investment as well as the opportunity cost of improving their skills with formal or informal education, training or experience.

Even though Mincer's framework does not account for individuals' skill heterogeneity, and I do not use it in this thesis, it has become one of the main models used to quantify the effect of education and training on wages (Hartog, 2000). In fact, multiple variations of this initial mathematical formulation have also been used by researchers to calculate the influence of different variables on wages. There is, for instance, the ORU model by Duncan and Hoffman (1981). This is one of the main models used to measure the effect of education mismatch on wages (Hartog, 2000).

Mincer's equations use the years of education to measure a qualification's effect on wages, whereas the ORU model introduces three dummy variables to account for possible education mismatch (Over-education, Required education, or Under-education), and measures the effect that these have on wages. As a consequence, Mincer's model is used mainly to estimate the returns on education, whereas the ORU model is used to estimate the pay penalty for educated mismatched workers.

Summarising, the economic value of education (or training) has been considered as part of economics since its earliest beginnings, although it became more prominent after the 1960s. The attempt to explain the fast recovery of the European economies after the Second World War followed by the Golden Age of capitalism placed education in the spotlight. The objective then was to explain to what extent the increase in the level of education of the population boosted economic growth. According to Human Capital Theory, there was a direct positive relationship, since education increased the productivity of the individual.

It is expected, therefore, that higher education trains the individual for their future job. S/he will acquire the general and the specific skills needed for a specific job during their formal education: secondary or even tertiary education. However, even if higher educational level does not increase the labour productivity of the individual, it does increase the individuals' ability to interact in the continuing changing real world. For instance, the ability to understand legal and economic documents is essential to be able to cope with today's world. Part of these general abilities is acquired with formal education, and especially, with higher formal education. Therefore, in this sense and even if at the beginning this was not the objective of this theory, one can assume that Human Capital Theory was right, at least if we consider the productivity as the ability to manage possible problems, which it does increase during peoples' lives.

However, Human Capital Theory accounts for only one side of the labour market, the supply side. Since then, more completed models and theories have attempted to better understand the (im)balance of the supply of and demand for workers in the labour market. In the next paragraphs, I analyse first the critiques of Human Capital Theory, and then other theories which include in their analyses both sides of the labour market.

1.1.2. Critiques of Human Capital Theory

Human Capital Theory is not the unique framework to understand the economic value of education. In fact, since its first publication in 1962, Human Capital Theory has been the subject of much debate. Most of the authors have directly criticized the assumptions of Human Capital Theory for being simplistic or unreal. These criticisms are not unjustified, and in fact, we have already discussed some of them.

Human Capital Theory is based on Classic Economy Theory, and therefore it assumes a perfect market, with perfect information, and with rational individuals. And it is well known that the real world is full of imperfections. However, as a simplification of reality, Human Capital Theory is a good exercise to try to better understand this imperfect world and these imperfect markets. Besides, it has been used as a basis to construct more complex theories or models.

In this section I will discuss the main theories that in one way or another have tested Human Capital Theory. The **Theory of Capabilities** (Sen, 1987; Sen, 1997), the **Feminist Theories** (Hartmann, 1976), the **Theory of Occupational Segregation** (Doeringer and Piore, 1971), the **Signalling Theory** (Spence, 1973) or the **Screening Theory** (Arrow, 1973; Stiglitz, 1975) are some examples.

The most important, from Theory of Capabilities to Screening Theory, are briefly summarised in the following paragraphs.

The **Theory of Capabilities** (Sen, 1987) is probably the least economic theory of all. It focuses on the most philosophical aspect of Human Capital Theory; that is on the definition of education. Sen's critique of Human Capital Theory argues that Becker's theory is narrow, in the sense that it does not consider all the benefits of education. According to Sen:

> *"If education makes a person more efficient in commodity production, then this is clearly an enhancement of human capital. This can add to the value of production in the economy and also to the income of the person who has been educated. But even with the same level of income, a person may benefit from education, in reading, communicating, arguing, in being able to choose in a more informal way, in being taken more seriously by others, and so on. The benefits of education, thus, exceed its role as human capital in commodity production. The broader human-capability perspective would record –and value– these additional roles. The two perspectives are, thus, closely related but distinct (…) The narrower view of human capital approach fits into the more inclusive perspective of human capability which can cover both direct and indirect consequences of human abilities."* Sen (1997, p. 1959)

Sen's reflections are important since they also positively encouraged economists, as well as other researchers such as sociologists or demographists, to look beyond the economic value of education. The positive effect that higher education levels have has already been discussed. Higher education levels increase, for instance, the individual's ability to deal with some paperwork or other social or legal constraints. Sen pointed this out and consequently, researchers are now looking at the impact of education on wellbeing, health or happiness, among others. For instance, Albert and Davia (2007) applied econometric techniques to analyse the effects of education on health at the end of the 20th century and the beginning of the 21st century in European countries, concluding that individuals with more than compulsory education are healthier, although the impact is different among countries.

The critique of **Feminist Theories** or **Theories of Gender** (Hartmann, 1976) of Human Capital Theory highlights the gender differences in the labour market.

According to these theories, gender gaps in the labour market (such as the proportion of women in the labour force or employed, the types of job held by women or the wages they receive) are a consequence of exogenous variables like history, economy or culture, variables that are usually taken as given by economists.

> *"The discipline of economics tends to view individual decision making as determined by economic incentives and individual preferences (or tastes). It generally does not analyse the formation of preferences, and choices are generally viewed as being at least to some extent voluntary (…) within the context of sociology or social psychology, individual choices are more likely to be seen as stemming from social conditioning or constraints rather than as voluntary"* Blau, Ferber and Winkler (2006, p. 159)

Feminist Theories, as part of their critique of classical economic theory, reject the central idea of Human Capital Theory that the worker's wage is equal to the individual's marginal productivity. If so, and according to the European Union Statistics on Income and Living Conditions (EU-SILC) in all activities, sectors, and jobs, men would be more productive than women, since it estimates an average unexplained gender pay gap of 11.1 percent in the OECD countries (Boll *et al.*, 2016).

As a whole, as has already been stated, the arguments of the previous theories against Human Capital Theory basically come down to its simplicity. The complex social, cultural, and historical contexts are not considered by Human Capital Theory and, therefore, it is not useful to interpret or to understand the real labour market.

Theory of Occupational Segregation or **Dual Labour Market Theory** (Doeringer and Piore, 1971) was developed at the end of the 1960s in an attempt to explain the critical economic situation of the moment, with high unemployment and poverty rates. It explains the labour market as divided into two markets which function independently. It is useful to explain different dichotomies: high versus low skill workers, men versus women, or even, national versus migrant workers.

According to the Theory of Occupational Segregation or Dual Labour Market Theory, there are two distinct markets for labour in the economy. The first one, the developed (or men or national workers) industry, uses high technology processes and requires highly skilled employees.[13] Its workers have better labour conditions including more stable jobs and higher wages. The second industry, the less technologically developed one (or women or migrant workers) requires lower-skilled workers and the workplaces have poorer labour conditions such as higher unemployment and lower wages. Workers cannot change the industry because each one uses its own productive system with different technological levels and, therefore the workers' demands are different: they need different skills levels or different characteristics such as gender or origin/nationality (Reich, 1984; Reich, Gordon and Edwards, 1973).

The Spanish labour market is, according to García-Serrano and Malo (2013) and Polavieja (2003), an example of a dual labour market caused by the coexistence of temporary and permanent contracts, especially after the legal reforms of the 1980s:

> *"In Spain, the extension of labour market segmentation is associated with a rather 'small' legal change easing temporary contracts in 1984. The gap in firing cost with respect to open-ended contracts (…) and the historical context fuelled temporary contracts as the main form of flexibility for firms"* García-Serrano and Malo (2013, p. 28)

This theory is not tested in this thesis, since data used are cross-sectional and it is not possible to follow individuals in different periods. Even so, it is important to bear in mind the previous results. They will help to understand the persistence of the educational imbalance in the Spanish labour market.

Signalling Theory (Spence, 1973) and **Screening Theory** (Arrow, 1973; Stiglitz, 1975) disagree with Human Capital Theory in its assumption of perfect market. The key idea of the Signalling and the Screening Theories is that there is imperfect information in the labour market. The Signalling Theory focuses on labour market supply, whereas the Screening Theory focuses on labour market demand. Both theories accept the relationship between education and

[13] In the particular case of the Feminist Theories this would be the men's labour market.

productivity (usually more educated people are more productive) but both reject the causal relationship that Human Capital Theory defends (people are more productive because they are more educated).

According to Signalling and Screening Theory, the main function of the educational system is to select the most skilled workers, because only more productive workers are likely to invest in more education. The reason is the relatively cheaper investment, pecuniary and not pecuniary, for these more able workers than for the least skilled ones, since only more able workers will obtain higher returns in the labour market. If so, educational level can be used by workers to signal (Signalling Theory) or by the employers to screen (Screening Theory) the workers who seem to be the most able –productive– workers, rather than increase their real productivity.

These critiques of Human Capital Theory agree upon the importance of education for the individual, although they differ in their explanations. Each of them is partly describing the actual effects that increasing the educational level has for an individual. Theory of Capabilities highlights how important education is for all aspects of daily life; Theories of Gender focus on the cultural and social differences between men and women; Signalling and Screening Theories centre on the imperfection of the labour market, and especially on those related to information. In contrast, Theory of Occupational Segregation or Dual Market Theory especially focuses on the division of the labour market into two sectors (skilled/unskilled; men/women; national/migrants, etc.). As a consequence, for the latter, education will be important when the segmentation of the labour market is related to the skill or aspects not related to personal characteristics: gender, nationality, race, etc.

In most cases the critique is not directly of Human Capital Theory, but of the Classical Economy Theory and its vision of the perfect market. The empirical studies demonstrate that, whatever the cause is, for example, higher productivity, signalling, or screening, a positive relationship exists between educational level and better labour conditions since individuals with higher educational levels have more opportunities in the labour market including higher occupational rates, better occupations, and higher wages (Barceinas *et al.*, 2000; Calvo, 1988; Greenwood, Jenkins and Vignoles, 2007).

The problem appears when signalling or screening mechanisms provoke a ‘process of qualification escalation’ or, in other words, ‘qualification inflation’

> *“In the process of qualification, by contrast, the pupil is concerned not with mastery, but with being certified as having mastered. The knowledge that he gains, he gains not for its own sake and not for constant later use in a real life situation – but for the once-and-for-all purpose of reproducing it in an examination. And the learning and reproducing is all just a means to an end – the end of getting a certificate which is a passport to a coveted job, a status, an income. If education is learning to do a job, qualification is a matter of learning in order to get a job.”* Dore (1997, p. 8)

However, the previous theories and models, critiques as well as Human Capital Theory, only consider one of the sides of the market: in most cases the supply of educated workers. Human Capital Theory and Signalling Theory focus on the supply side, whereas Screening Theory analyses the labour market from a demand perspective.

In supply side theories future workers decide what the adequate level of investment in education or experience is in order to obtain higher wages and better jobs. Human Capital Theory believes in the improvement effect that these investments have on the individual’s productivity, whereas Signalling Theory questions the existence of productive increases. Spence (Signalling Theory) claims that most abled individuals would use their higher educational or experience levels to show the employers that they are better candidates to do the job. Screening Theory neither believes in the productivity improvement of the education or experience, but contrary to the Signalling Theory, Arrow and Stiglitz give the power to the employer, who decides on the education or experience level that the future employees need to hold to be hired in a particular job.

However, to what extent an individual with a particular set of characteristics will benefit from the labour market also depends on the demand in the market for these characteristics. General theories considering both sides of the labour market and accounting for this match between supply of and demand for workers are considered in the next section.

1.1.3. General theories. Considering both sides of the labour market: supply and demand

The feebleness of partial theories to understand and explain labour market patterns has encouraged other researchers to undertake a more global vision, analysing both sides of the labour market together. The two most relevant examples are the **Job Competition Model** (Thurow, 1970), and the **Assignment Theory** (Tinbergen, 1956; Sattinger, 1980; Hartog, 1981).

Job Competition Model (Thurow, 1970) uses a game theory framework, to describe the accommodation of a workplace with the best candidate for that job. In this sense, the model considers the possible heterogeneity of individuals' skills, but also among the characteristics of the workplaces. According to Thurow's model, workplaces as well as workers are heterogeneous.

Workplaces differ in the quantity and quality of tasks that they require. Equally, workers differ in their abilities (a combination of innate and acquired abilities). The key idea of the model is that each workplace has a *labour queue* where workers are classified by their skilled appropriateness for the job. For instance, if there are two different types of workplaces, then 'type A' requires more skilled workers and 'type B' requires less-skilled workers. The best worker for the 'type A' workplaces is the first one in the *labour queue* of the 'type A' workplaces. Workers will be hired in all available workplaces following the order of this *labour queue*. The problem arises when there is oversupply of workers for this type of 'A' workplaces. Then, part of these excesses of supply (skilled workers) would be employed in less skilled workplaces where they would be over-educated or over-qualified, and less-skilled workers, who were in the first positions of the *labour queue* for the 'type B' workplaces, would be displaced.

In this sense, Thurow considers that the problem lies not as much in the lower returns for over-educated workers as in the crowding-out of less educated (less skilled) workers, who reduce their probabilities of being employed, and therefore, the education inflation defined by Dore (1997) emerges:

> *"As the number of certificate leavers grows far larger than the number of clerkships that are available, (…). The bus company gives them preference. Soon all the available conductor slots are filled by senior certificate holders: a senior certificate has become a necessary qualification for the job"* Dore (1997, p. 7)

But some clarification should be made here. As Borghans and de Grip (2000) argue, a distintion between crowding-out or bumping down and upgrading should be made. The difference between these two concepts lies in the extent to which the qualification required for the job is really needed to perform the tasks required by the job. When a skill biased technological change has occurred and subsequently a particular job cannot be performed without a higher level of qualifications, it is said that there has been an upgrade of the job. In contrast, when there has only been an increase in the qualifications demanded by the employer, then it is said that there has been a crowding-out or a bumping down. The problems in the labour market will come out when the latter phenomenon appears.

Assignment Theory (Tinbergen, 1956; Sattinger, 1980; Hartog, 1981) also focuses on the match between the characteristics of supply of and demand for labour. It claims that on the one hand, part of the worker's productivity depends on the characteristics of the workplace, and on the other hand, a wage is the instrument used to allocate workers to workplaces. Therefore, according to Assignment Theory, the best match would be the one that maximises the productivity of the worker, and in that case, they would receive the highest reward. If, on the contrary, there is a mismatch between the skills of the worker and the skills required in the workplace, then there is an under-utilisation of the resources (since the worker would not maximise his or her productivity) and a lower return for the worker (since wages depends on individual's productivity). There are two main sorts of mismatch: over-education (over-qualification) when the individual's educational level (qualification) is higher than the one required by the workplace, and under-education (under-qualification) when the individual's educational level (qualification) is lower than the one required by the workplace.

In short, both general theories Job Competition Model and Assignment Theory conclude that there is not a perfect match between supply of and demand for workers due to the difference in the qualifications (skills) level they hold and the ones demanded in the labour market. In the Job Competition Model the oversupply of qualified workers or the under-demand for these qualified workers is the main point of instability. According to Assignment Theory, the qualification mismatch in the labour market is due basically to imperfect information. The worker does not perfectly know the level of skills required in the workplace and the employer does not perfectly know the worker's skill level. These theories have become more used since the increase in the number of university graduates in the developed countries in the last decades of the 20th century. In fact, the calculations of the main chapters of this thesis are based on quantifying the education mismatch in the Spanish labour market.

The **Theory of Career Mobility** or **Job-Matching Theory** (Jovanovic, 1979a; Jovanovic, 1979b; Sicherman, 1987; Sicherman and Galor, 1990) gives a different vision of qualification mismatch. It focuses on explaining the particular case when the worker's qualifications are higher than the ones required in the workplace, in other words, where there is over-education. According to the Theory of Career Mobility, over-education is a temporary situation which is usually related to young workers with high education levels having little or no on-the-job training. These young educated workers will improve their total human capital with the skills acquired in this (usually first) temporary job. Therefore, young educated workers agree to work in these less remunerated and less skilled jobs, that is, young workers agree to be temporarily over-educated, for the first years of their career for two main reasons: first, they expect to improve their experience; and second, they expect to be more likely in the future to increase the occupational ladder and get a better job with better conditions commensurate with their education level, either in the same company or in another one (Jovanovic, 1979a; Jovanovic, 1979b; Sicherman, 1987; Sicherman and Galor, 1990). Therefore, for the Theory of Career Mobility, over-education is one of the many steps for a worker to get an adequate job, rather than a problem.

Therefore, Job Competition Model, Assignment Theory, and Theory of Career Mobility or Job-Matching Theory, are all theories which attempt to explain the education (im)balance in the labour market. All of them agree on the fact that the current labour markets are having problems accommodating all the workers in the appropriate jobs. Job-Matching Theory or Theory of Career Mobility states that this is a temporary problem due to the imperfection information of the labour market and due to the adverse selection of the workers regarding their real productivity level. Similarly, the Job Competition Model also suggests the adverse selection of the workers. In contrast, Assignment Theory does not focus on the causes of this imbalance, but on the consequences of this education mismatch.

1.1.4. The economics of education: final remarks

All in all, this section shows that the focus of the research on the economics of education has been changing in parallel to the labour market reality. The first economists (e.g. Smith, Malthus, or Marshall) were concerned about the economic value of education, since they were the first ones who considered education as an investment and not as an expense. In this sense, the positive relationship between educational levels and productivity assumed by these first economists has largely been discussed here.

Once economic researchers were able to explain how the educational level influences wages (e.g. Becker and Mincer), the proportion of population with higher educational levels was increasing and the first mismatches in the labour market appeared.[14] As a consequence, part of the research shifted to these new

[14] Since the 70s, 80s and 90s, depending on the country, there was an extraordinary increase in the number of university graduates. The increase of their share in the active population made it necessary to develop a deep knowledge of their employability in the labour market.

labour market imbalances (e.g. Thurow, Sattinger, Hartog or Sicherman and Galor).

The empirical literature demonstrates there are mismatches in the labour market: over- and under-education. However, they do not agree on whether this mismatch is an “assignment” or a “competition” problem. The empirical studies do not agree on whether education mismatch is a temporary or a permanent phenomenon for the individual. A complete longitudinal dataset would be needed to verify the Theory of Career Mobility. Knowing the duration of the education mismatch is essential to develop good policies to reduce their consequences for both individuals and society.

As the labour market imbalances appeared, researchers realised that individuals were not homogeneous and these influences were different even between individuals of the same educational level (e.g. Duncan and Hoffman). Therefore, the difficulty of measuring individuals’ skill heterogeneity using the level of education became more evident.

However, the lack of data measuring individuals’ skills drives on the use of education level as a proxy of individuals’ abilities to analyse the level of labour market mismatches or their effects on wages. However, if these individuals’ heterogeneity is nowadays more than assumed, finding the perfect measurement for this heterogeneity is not always possible. Indeed, this is one of the most difficult problems for today’s researchers, to find a good indicator for individuals’ skill heterogeneity.

Recent data have allowed researchers to account for some of individuals’ skills, introducing skill heterogeneity in the analysis or, even, moving the core of the studies related to labour market mismatches from education mismatch to skill mismatch. However, a concrete definition of which skills should be considered or how to quantify them in or out of the labour market is still missing:

> *“‘Skill’ is a notoriously difficult concept to define, let alone measure (Department of Labor 1999). These problems are not restricted to human capital theorists, but they confront particular problems because of their assertion that skills can be accurately measured in quantitative terms. Skill is variously defined as the expertise, ability, or competence to undertake specific activities often acquired through formal instruction*

> *or work experience. An obvious problem is how to measure the skill levels of different jobs."* Brown (2001, p. 23)

In this sense, Chapter Four of this thesis aims to contribute to the analysis of the variation in skills among individuals and their effect on the probability of being over-educated. In general terms, empirical literature has analysed the effect that accounting for the individuals' skill heterogeneity has on wages. However, in this thesis I use data from the Survey of Adults Skills to see whether individuals' proficiency level of literacy and numeracy skills affects their likelihood of being over-educated.

1.2. Education mismatches in the labour market

1.2.1. Concept and measurements

Over-education (or over-qualification) and under-education (or under-qualification) are the two main education mismatches that can be found in a modern labour market.[15] There is a third education mismatch which is the area or field of study mismatch which, as its names states, appears when the individual is working in a different area from the one of his or her education. For instance, it appears when a veterinarian is working in sales or an artist as a clerk. However, less attention has been paid to this latter mismatch, in part due to the lack of data regarding fields of study.

Most of the literature on the causes and consequences of over-education has been written since Freeman introduced the concept in 1976 (Freeman, 1976). Since the publication of *The Overeducated American* (Freeman, 1976), different authors have studied these qualification mismatches for different countries and periods.[16] Most of these articles have focused on the measurement of these

[15] Modern labour market refers here to a market where individuals voluntary sell their work to employers. In return, they receive a monetary wage which they can use to buy the goods or services they need or desire.
[16] A summary of the main literature on education mismatch can be found, for instance, in Sloane (2002) who offers very detailed data about the studies, or in Badillo-Amador, García-Sánchez and Vila (2005), within others.

education mismatches and their effects on earnings (Groot and Maasen Van Den Brink, 1997; Hartog, 2000; Sicherman, 1987). Articles focusing on their non-pecuniary consequences, for example, job satisfaction or health, have also been published (Allen and van der Velden, 2001; Battu, Belfield and Sloane, 1999; Tsang, Rumberger and Levin, 1991; Verhaest and Omey, 2006). Other authors have particularly centred on the effects of over-education on university graduates, the population group most likely to be over-educated (see, for instance, Barone and Ortiz (2011) for Europe; or Battu, Belfield and Sloane (1999), Chevalier (2003), Chevalier and Lindley (2007), Dolton and Vignoles (2000), and Green, McIntosh and Vignoles (1999) for the United Kingdom). Nonetheless, there is still a debate about a precise definition and the best way of measuring it (Green, McIntosh and Vignoles, 1999).

According to Rumberger (1981), there are three different circumstances which can cause over-education:

(a) whether the economic position of people with higher educational levels has deteriorated in relation to the position that individuals with the same educational level have historically held;
(b) whether the requirements of the job dissatisfy the expectations of the higher educated workers who secure them, in the sense that the job requirements are too easy;
(c) whether the educational skills of the workers are higher than the ones required in the workplace.

The third one is the definition that is most commonly used, although the three of them are highly correlated because, according to empirical analysis, when the educational skills of the workers are higher than the ones required in the workplace (definition c) the worker usually receives a lower wage (definition a) and it is likely that s/he finds the work dissatisfying (definition b).

Therefore, over-education commonly defines a situation in the labour market where "*the individual is simply supposed to have been at school longer than necessary for this job*" (Hartog and Oosterbeek, 1988, p. 187). Indeed, this is the definition that I will use in this thesis to estimate and analyse the likelihood for a university graduate to be over-educated in the Spanish labour market.

The opposite situation, under-education, is supposed to occur when the individual does not have enough formal education to hold the skills needed to do the job. Hence the focus of the definition of over- (or under-) education is on the number of years of formal education or the level of formal education that a worker holds.

Knowing the grade of labour market over-education is essential, since over-education implies a waste of personal and social resources. According to the main empirical studies, over-education has three main negative consequences:[17]

(a) a squandering of personal and social resources (the latter since most of the non-compulsory education in European countries is subsided by public finances. See the Spanish case for instance where in 2000, 39 out of 60 universities were financed by public resources) (García Serrano and Malo, 1996; Tsang and Levin, 1985),

(b) a displacement of less skilled workers, this is the bumping down or crowding-out effect,

(c) an underuse of most of the workers' abilities (reducing the personal dedication, interest and productivity at work) (Hartog, 2000; Hartog and Oosterbeek, 1988; Rumberger, 1987a).

As has already been seen in the previous section, most researchers tend to agree to use formal educational level as a good proxy of the individual's abilities, both innate and acquired. Education mismatch does not account for the individuals' heterogeneity. It supposes that all individuals with the same educational level are equally skilled. Theoretical as well as empirical studies have refuted this argument (Alba-Ramírez and Blázquez, 2003; Allen and van der Velden, 2001; Battu, Belfield and Sloane, 1999; Berg, 1973; Chevalier and Lindley, 2007; Green and McIntosh, 2007; Green, McIntosh and Vignoles, 1999). However, until

[17] See, for instance, Battu, Belfield and Sloane (1999); Green, McIntosh and Vignoles (1999).

recently, the lack of adequate data to measure the actual skills of individuals have forced researchers to use the educational level as a proxy of individuals' skills level and to use education mismatch instead of skills mismatch (Quintini, 2011). In fact, I will use the number of years of education to measure the degree of over-education in the Spanish labour market in Chapter Three and Four.

As has already been stated, the fact that individuals are not homogeneous and that they hold different abilities (either innate or acquired) has been present in the literature since Adam Smith, who emphasised the importance of the acquired abilities more than the innate ones:

> *"The difference between the most dissimilar characters, between a philosopher and a common street porter, for examples, seems to arise not so much from nature, as from habit, custom, and education. When they came into the world, and for the first six or eight years of their existence, they were, perhaps, very much alike, and neither their parent nor playfellows could perceive any remarkable difference. About that age, or soon after, they come to be employed in very different occupations. The difference of talents comes then to be taken notice of…"* Smith (1961 [1776], pp. 19-20)

Nevertheless, and according to Smith, educational level is a good indicator of potential individual skills. In fact, the Assignment Theory also predicts a strong correlation between education and skill mismatches (Allen and van der Velden, 2001). Even so, educational level does not account for individuals' skills heterogeneity. It does not take into consideration the differences between the schools that the individuals have attended, the areas in which they have studied, or the proficiency levels of individuals, among others. Whether the individuals' skill heterogeneity is considered, the relationship between education and skill mismatches may not be assured.

Those studies that have analysed both mismatches, education and skill, using the same set of data, have confirmed the low correlation between the two measurements. For instance, Halaby (1994) already argued that

> *"overeducation functions do not reliably identify workers whose skills are underutilized", and therefore "the validity of overeducation*

> *as an indicator of skill mismatch is weakened*". Halaby (1994, p. 56 and 58)

More recently, Flisi *et al.* (2014) used PIAAC data for different European countries and found that

> *"the share of people who are mismatched both in education and skill ('severely' and 'mixed mismatched') over the total population is pretty low; on average when summed up the two groups count for 15% of the population".* Flisi *et al.* (2014, p. 46)

Similar conclusions were found by Levels, Van der Velden and Allen (2013, p. 21) for whom "*part of the educational mismatches is just apparent and do not necessarily imply that worker's skills are heavily underutilised or overutilised*". Using data from 2001 British population, Green and McIntosh (2007, p. 437) estimated that "*less than half of the over-qualified are over-skilled*". And for the Spanish case, Nieto (2014, pp. 54-55) estimated that "*23 percent of overeducated workers are also overskilled*", concluding that there is "*a weak correlation between overeducation and overskilling*".

The main reason for the weak relationship between education and skill mismatch is, according to Flisi *et al.* (2014) that

> *"the two dimensions of skill and education mismatch tell a different story and provide different pieces of information"* Flisi *et al.* (2014, p. 46)

However and despite the different nature of the two measurements, Allen, Levels and Van der Velden (2013) justify the weak relationship between the two measurements because

> *"the link is indeed somewhat stronger in the case of skill use than in the case of skill level, which is why we observe a residual relationship between educational mismatches and skill mismatches".* Allen, Levels and Van der Velden (2013, p. 27)

Carrying out new surveys including some sort of measurement of skills (in most cases cognitive skills; i.e.: literacy, numeracy) has changed the research on the labour market mismatches, since it allows the researchers to account for individuals' skill heterogeneity, at least for a part of it. A series of international projects have provided new data including some sort of measurement for the skills. There is, for instance, the International Adult Literacy Survey (IALS)[18] carried out between 1994 and 1998 in 22 countries and regions or the Adult Literacy and Lifeskills Survey (ALL)[19] conducted first in 2003 with the participation of seven countries/territories, and later between 2006 and 2008 with the participation of four countries. One of the most recent surveys including individuals' proficiency on literacy, numeracy, and problem solving for different OECD countries is the Survey of Adults Skills (PIAAC). Besides, this survey includes important information (background, demographic characteristics, educational level, labour-force status, within others) which allows measuring the education as well as the skill mismatches. A detailed description of the methodology and results of this survey can be found in OECD (2013a).

A main advantage of these international surveys is the possibility of comparing different countries, since the same questionnaire and codification was employed and the samples harmonised. In fact, some authors have done this (see Flisi *et al.*, 2014; Levels, Van der Velden and Allen, 2013).

Notwithstanding this, in most of these cases only cognitive skills are being measured, and a measurement for non-cognitive skills is still missing. These latter are also important for understanding the skills heterogeneity among individuals and their effect on labour market mismatches or even on wages (Heckman, Stixrud and Urzua, 2006). Nevertheless, according to recent studies, as for instance Green (2012) using data for Britain, the late changes in the use of skills in the workplace indicates an upgrade of the use of cognitive skills, with literacy higher than numeracy, rather than the physical ones. These results suggest that the measurement of cognitive skills, as literacy or numeracy, may became a good estimation for the skills level of individuals, and therefore to account for their heterogeneity even when they hold the same educational level. In line with these results, the literacy and numeracy proficiency levels of PIAAC are used in Chapter Four to measure the effect that cognitive skills heterogeneity

[18] https://nces.ed.gov/surveys/ials/.
[19] https://nces.ed.gov/surveys/all/.

has in the probability for individuals to be over-educated in the Spanish labour market.

There is clear empirical evidence that not all students with the same qualifications hold the same skills. Besides the innate characteristics which define each individual and determine their skills, there are also differences in the education acquired, such as different subjects, different studies or different institutions which can markedly influence the final skills or abilities acquired. For instance, the articles by Green, McIntosh and Vignoles (1999) for the U.K. or Barceinas *et al.* (2000), and Oliver, Ramos and Raymond (1999) for Spain demonstrate that, in some cases, the type of degree is more important in determining whether or not an individual is over-educated than the fact of merely being a university graduate. They concluded that holding a degree in technology or science rather than in humanities increased the probability of being employed and reduced the likelihood of being over-educated. This displaces attention from the measurement of educational levels to the content of education rather than the years of schooling or the grade held.

However, measuring over-education as *"the individual is simply supposed to have been at school longer than necessary for this job"* (Hartog and Oosterbeek, 1988, p. 187) also implies measuring the educational level needed (or required) in the workplace. The concepts of required and needed education have different meaning, especially when the Signalling Theory or the Screening Theory is taken into account. The required education is the education level that the employer asks for, and it can be equal, higher or lower than the real education level needed to do the task in the workplace. For instance, it is quite common for the employer to require a higher educational level for a workplace when there is high unemployment, and a high supply of workers. In this case, some credentialism will appear.

However, here the two concepts, required and needed, will be used interchangeably, because empirical studies use one or the other depending on which one is available in the dataset. Furthermore, the required and the needed educational level for a job will be different only if there is credentialism in the labour market, that is an oversupply of workers and, therefore, employers demand for the ones holding the highest levels of formal education even if they do not need them for the job.

Researchers use different ways to measure the educational level needed (or required) in the workplace, and therefore, different ways to measure over-education. These measurements are usually determined by the data set because the definition and the way to measure it would differ, whether the source of the data is a survey where employees are asked about their own perception of their job, or whether it is official data where only a general description of the job or an occupational classification is available. In the first case, the grade of education mismatch is reported by the worker, being therefore subjective, depending on the perception of the worker. In the second case, the degree and the kind of education required (needed) in the job are determined by external sources, usually less specific and more rigid. And, therefore, it is more difficult to account for the changes in the tasks of the workplaces over time.

Summarising, there are three main ways to measure the kind or level of education needed by the workplace:

(a) self-assessment techniques: based on subjective methods such as the worker´s opinion on the educational level required by the employer;
(b) statistical measurements: which determine the educational level needed on the job as a statistical measure –usually mean or mode– of the educational level of the workers of the same occupational category.
(c) external measurements: such as the average years of education of the workers or a Dictionary of Occupational Titles;

However, none of these three measures gives a perfect value (Green, McIntosh and Vignoles, 1999). Therefore, studies usually use one technique or another depending on the set of data available.

Subjective methods (or worker's self-assessment techniques) require a data set with the personal evaluation or opinion of individuals about their own knowledge and the one required in their workplace. Despite trying to measure whether the educational skills of the workers are higher than the ones required by the workplace (definition c), it seems to me that the utilisation of the subjective methods is more related to the second definition of over-education given by Rumberger, when the worker is dissatisfied due to the low achievements of the workplace (Rumberger, 1981). But, whatever the definition of over-education they use is, these methods have been used in several national as well as international empirical studies to measure the degree of over-education. I will also use them in the fourth chapter of this thesis.

One of the first examples of the measurement of over-education using subjective methods is, for instance, the Career after Higher Education: a European Research Study (CHEERS) whose results for Spain were published at the end of the 20th century (García-Montalvo, 2001). The CHEERS project started in 1997 with the collaboration of institutions from eight European countries (France, Finland, Germany, the Netherlands, Italy, Norway, Spain, and the United Kingdom). Subsequently, Sweden, the Czech Republic and Japan joined the project. This international collaboration in one unique project was motivated by the lack of information regarding the employability of university graduates.

After designing the survey sheets and realising the interviews, the CHEERS project gathered information about eleven European countries plus Japan.[20] The researchers aimed to unify the main questions to one unique general questionnaire. They also included some specific country questions. In the Spanish case, for instance, due to the higher unemployment rates, even among university graduates, they included some questions to be answered only by unemployed graduates (García-Montalvo, 2001, p. 26).

CHEERS was important basically for two main reasons. First, the same questions were asked in different countries.[21] This allowed cross-country comparisons. Second, the survey was particularly relevant for research on over-education. Thus, for each individual interviewed there is information about their personal characteristics (including the economic and social familiar background), the

[20] See Teichler (2007) for a detailed description of the survey and some of its results.
[21] They were very careful to translate the questions in order to have the same meaning in all countries.

university curricula (the degree studies, rather than the university where they were enrolled as in the USA), one's transition from university to the labour market, and the economic structure of the countries, particularly its labour market conditions and the impact of technological development on the workplace (García-Montalvo, 2001).

More recently, the OECD has created the Programme for the International Assessment of Adult Competences (PIAAC). This is a survey of adult's skills which focuses on the skills of the worker and of the workplace. For each country, PIAAC covers all the population from 16 to 65 years old, and includes objective information about the formal educational level and subjective information, since interviewees are asked about the level of skills needed and used in their everyday workplace. It also measures the literacy and numeracy proficiency skills of the individuals surveyed. A more detailed description of this survey can be found in Chapter Four where I analyse the data.

The second over-education measurement, **statistical methods** is based, as its name indicates, on using statistical measurements (mainly mean or mode of the occupational group, although it can also use other statistical measurements) to determine the level of education required to do the job. These methods are based on the supposition that supply of and demand for work match in the actual labour market. Therefore, they might not be very feasible if there is any sort of mismatch in the actual current labour market. The advantage of this methodology is that it does not need the worker's self-assessment or any other extra information rather than the dataset used as a sample.

One of the first articles in considering statistical measurements was Verdugo and Verdugo (1989). Using a sample conducted in 1980 in the US of white men aged between 25 to 64 who worked full-time as salaried employees, they estimated the effect of an excess or default of years of education on wages. Applying statistical methods, they classified the mean of years of education for each of the three-digit census occupations using all the population (not only white men). The mean and standard deviation found for each three-digit occupation were then used to define the population of the sample (only white men) as follows:

> *"Workers with educational attainment greater than one standard deviation above the mean for their specific occupational are defined as overeducated. Conversely, undereducated workers are those whose education is less than one standard deviation below the mean. Adequately educated workers comprise the middle ranges: within +/- one standard deviation of the mean."* Verdugo and Verdugo (1989, p. 633)

The subjectivity of the measurement used by Verdugo and Verdugo (1989) has been questioned with the argument that there is not a formal reason why the limit of the differentiation between adequately educated and over-(under-) educated should be plus(minus) one standard deviation (Green, McIntosh and Vignoles, 1999). Even so, a version of these statistical methods is applied to different Spanish data in Chapters Three and Four of this thesis.

Finally, **objective or external measurement methods** use impartial sources to quantify the degree of mismatch. The Dictionary of Occupational Titles (DOT) published in 1965 by the US Department of Labor, for instance, is one of the sources used in the United States. Rumberger (1987a) used this source to estimate the degree of over-education in the 1980s USA labour market. As Rumberger defines in his article

> *"The DOT, developed by the U.S. Employment Service, contains detailed descriptions of all occupations in the U.S. economy and information on a number of occupational characteristics. These characteristics include the amount of general and specific training needed for 'average performance' in each job situation."* Rumberger (1987a, p. 30)

However, an equally objective source is not available for all countries and when there is one, it is updated only once in a while. Therefore, despite being an objective measurement it is not free of imperfections. For instance, it is difficult to update the tasks of the jobs or to account for the changes in technology, and the cost of updating the information is high (Hartog, 2000). Besides, in some countries, such as Spain for example, there is no such dictionary to define the skills required in one occupation. This is one of the main reasons why there are few studies which use objective measurements.

As is usual in social sciences, none of the measurements is perfect, and the choice between using one or another is usually defined by the data set. Objective methods seem to be more impartial in the definition of the tasks required in the workplace, but they are more expensive and rigid. Subjective methods seem to be more accurate to measure the real skills of the workplace and the individual’s abilities, but the results depend on the workers' judgements. Finally, statistical measurements are based on the current adjustment between workers and workplaces, taking it as matched. As has already been stated, in this thesis I used two of the three measurements: statistical and subjective.

Table 1. 1. Summary of the main characteristics of the three measurements for over-education

	Objective	**Subjective**	**Empirical o Statistical**
Measurement	Clear definitions for occupations and detailed measurement methodology	Worker’s self-assessment	Based on criteria of frequency (mean, mode, etc.)
Assumptions	Education and abilities are not substitutes (all workers need the same education level to do the job)	Employees report actual requirements much more accurately than employers	Assumes that most of the workers are matched
Occupation	General description of the occupation. Different jobs can be grouped in one unique occupation	Specific description of the tasks required by the job rather than a general one for the occupation	Based on the current worker-workplace adjustments of the labour market
Education	Equals all types of education (years of schooling or dummies by educational levels)	Credentialism can be underestimate (workers state current hiring standards requirements)	Different educational levels can suit one occupation
Responsiveness	Difficult to introduce technological and organisational changes (classifications are used for long periods)	Changes in the job requirements are reported when they occur.	Sensitive to labour market conditions and technological developments

Source: Borghans and de Grip (2000, pp. 15-16)

Following Borghans and de Grip (2000), Table 1.1 summarises the most important differences between the three measurement methods. The table describes the source used to quantify over-education and the assumptions about education and abilities that the measurement uses. It also includes information about the way they deal with occupation and education definitions, as well as, their responsiveness.

Despite the fact that the decision of using one or another method depends on the data available, part of the literature on over-education has focused on finding which of the measurements is more accurate to estimate the degree of mismatching.

It is almost impossible to estimate the degree of over-education for the same data using different measurements. Their definition makes it impracticable. Only the new survey of the OECD, the PIAAC allows it (Martínez García, 2013). In other cases, it has been possible to compare data from the same regions and the same periods, but using different sources. That allows a comparison between them, even if the original source and the interviewees are not the same. In the cases where the comparison has been possible, the results indicate that subjective measurements tend to overestimate the mismatches, because of the *"workers' inclination to inflate the standing of their job"*, whereas objective measurements underestimate them (Hartog, 2000, p. 133).[22] But there is not a predictable result for statistical measurements. Statistical methods estimate the mode or mean of the allocation distribution in a particular time, the time of the dataset. Therefore, the outcome would be over or underestimated, depending on the current labour market situation. Indeed, the comparison between the result of statistical mean approach and subjective model is possible in the fourth chapter of this thesis.

Whatever the measurement method used, we should be cautious when a period is analysed. As will be seen in the next chapters, the upgrading of the population's educational level as well as the innovation in the production or organisation of the job can change the current equilibrium between demand for and supply of quality of the labour force. For instance, some technological innovation can either

[22] Even though, in their article using data from the school-leavers on the Dutch labour market in 1993, Van der Velden and Van Smoorenburg (1997) found the opposite results: the objective analysis overestimated the level of over-education, whereas the subjective method was more reliable.

increase the demand for highly educated workers, for instance, if the new tasks required a highly specific knowledge such as solving the technical problems of the machine, or on the contrary, increase the demand for less-educated workers if the technological innovation developed upon more tedious work that only needs someone to supervise the machine works (Acemoglu, 1998; Acemoglu, 2002a; Acemoglu, 2002b; Goldin and Katz, 2008). Therefore, depending on the direction of these changes the labour force situation may be more or less balanced with the workplaces they hold (Green, 2012; Mañe Vernet, 2001).

1.2.2. Empirical studies

The literature does agree on the importance of considering and measuring education mismatches, and especially over-education, in order to understand the labour market better. As pointed out by numerous empirical studies, knowing the degree of over-education in the labour market is essential, since over-education supposes a waste of personal and social resources (García Serrano and Malo, 1996; Tsang and Levin, 1985). For instance, it has been broadly demonstrated that when over-education is a permanent phenomenon it has negative effects on wages as well as on the productivity of the workers (Hartog, 2000; Hartog and Oosterbeek, 1988; Rumberger, 1987a).

In order to contrast which theory is the most appropriate one to analyse the consequences of mismatches, and in particular over-education, on the labour market, most empirical studies focus on:

(a) the effect that over-education has on wages,

(b) whether the wage differential between individuals with the same set of characteristics are due to differences in productivity or a signalling or screening effect,

(c) the length of this education mismatch (whether it is a permanent or a temporary phenomenon),

(d) how education mismatch would affect the job mobility intra- or inter-firm.

Regarding the effects of over-education on wages (a), the ORU model (Duncan and Hoffman, 1981) is the methodology most used to measure these effects. As explained earlier on, it is based on Mincer's wage function (Mincer, 1962; Mincer, 1974), but it splits years of education into three dummy variables: Over-educated, Required education, and Under-educated. Mincer's wage function uses the logarithm of wages as the dependent variable, and gender, years of education, and experience among others as control variables. Therefore, the results measure the correlation between years of education and wages, but it does not say anything about those who are mismatched.[23] The introduction of these three new variables (ORU) allows the researchers to measure the returns on educational mismatched workers (Over-educated as well as Under-educated) compared to the workers who were educationally adequately matched (Required education).

International studies about education mismatch and the problems that this phenomenon provokes to the good performance of the economy appeared with the rise of university graduates in the United States. Freeman's book drew attention to the effects on wages due to the increase in the number of educated workers entering the labour market (Freeman, 1976). After that, the interest of economists in measuring these effects mushroomed around the world, considering first those countries where the expansion of higher education students was more evident.

[23] However, it does allow the researchers to measure the negative impact of over-education on wages when data of different times are available.

Rumberger (1987a) wrote one of the first studies evaluating the effects of over-education on productivity and earnings. He used the Dictionary of Occupational Titles (objective approach) to identify those workers mismatched in the labour market in the United States for the years 1969, 1973, and 1977. According to his calculations, the proportion of over-educated workers was 35 percent in 1969, 27 percent in 1973, and 32 percent in 1977; and the pay penalty for those mismatched workers was 30 percent in 1969, 40.4 percent in 1973, and 34.8 percent in 1977 for men. For women, the pay penalty was higher all the years, except in 1969. It ranges from 26.1 percent in 1969, to 43 percent in 1973, and to 36 percent in 1977. The pay penalty also differed among occupation groups for both, men and women, being higher for those in professional or managerial occupations.

The results of other studies also using data from the United States in this period are consistent with Rumberger (1987a). Using a statistical approach (mean plus one standard deviation), Clogg and Shockey (1984) analysed the US data from 1969 to 1980 and concluded that education mismatch increased during this period, although they estimated a degree of over-education around 15 percent among men and around 9 percent among women. They particularly analysed the cohort effect and stated a higher average increase for the youngest generations. These results confirm the general educational upgrading of younger generations and the remainder of the occupational distribution.

In Europe, seminal work by Hartog and Oosterbeek (1988), using data for the Netherlands, estimated a rise in the proportion of over-educated workers from 7 percent in 1960 to 25.7 percent in 1977. As in other countries when the number of higher education students increased,

> *"The development towards increased levels of schooling has even been strengthened by the composition of labor force withdrawals at higher ages. Withdrawals on account of early retirement and*

> *disability have fallen disproportionately on the lower educated.”* Hartog and Oosterbeek (1988, p. 186).

According to their data, in the Netherlands in 1977, the pay penalty for over-educated workers in relation to those who were adequately educated was 20 percent for each year of mismatch: 29 percent for women and 14.5 percent for men. Analysing the UK graduate labour market from 1980 to 1986, Dolton and Vignoles (2000) found a proportion of over-educated graduates of around 30 percent.

According to the results of the empirical studies, on average, over-educated workers still get positive returns for their higher educational levels, although they earn less than the workers with their same educational level who are adequately educated in the workplace. Under-educated workers get negative returns, since they earn less than those higher educated workers who are adequately educated. However, under-educated workers still earn more than the workers with their same educational level who are adequately educated. Therefore, they claim that the fact that the returns for over-educated people are still being positive is the main reason why individuals continue investing in education (Chevalier and Lindley, 2007; Dolton and Vignoles, 2000; Green, McIntosh and Vignoles, 1999).

These results indicate the importance that the characteristics of the workplace have in determining wages. The influence will be greater in markets where the price (wage) has a threshold. This is, for instance, the case of the Spanish labour market where most occupations have a minimum wage imposed by a sector/region labour agreement. Where there is a minimum threshold, workers will receive at least the minimum wage even if they are under-educated, meaning their level of education is lower than the one required for the job.

Empirical studies based on Human Capital Theory agree with Freeman’s insight: the main problem of over-education would be its negative impact on the workers’ earnings and opportunities in the labour market (Freeman, 1976, p. 4), although this impact has not been as negative as he expected. According to most of the empirical studies it seems that over-educated workers were still having some education returns, albeit lower than the ones that are matched.

In order to test Human Capital Theory and Signalling or Screening Theory (b), there is a group of authors who have focused on trying to disentangle if the returns to higher educational levels are really a reward for higher productivity as the Human Capital Theory predicts or if, on the contrary, they are a consequence of a signalling or screening effect. This is, for example, the case of the studies done by Green, McIntosh and Vignoles (1999); Chevalier (2003); and Chevalier and Lindley (2007) for the United Kingdom; or Alba-Ramírez and Blázquez (2003) for Spain. The main results of these researches confirm the relationship between higher educational levels and the improvement of the productive capacity of individuals, but they also point out the signalling or screening effect of this higher education levels (Iglesias, 2005), although as Green, McIntosh and Vignoles (1999, p. 25) stated "*there is [in the UK] little evidence of widespread of qualification inflation*".

Obviously, the effects of over-education in the labour market might be very different whether it is a temporary phenomenon or a permanent one (c). An implication of Human Capital Theory and Theory of Career Mobility is that over-education is not really a serious problem, since it is a temporary phenomenon which will disappear in the long term.

According to Human Capital Theory, which, as has already been seen, derives from the Classic Economy, over-education is a temporary phenomenon caused by the excess of supply of highly educated people (the market has a disequilibrium for a while). The oversupply of highly skilled people will be solved by the perfect competence of the labour market, although it could take some time. As it occurs in a goods market, the excess of supply of highly educated people will reduce their price or remuneration (wage). The downward trend in wages due to the abundance of educated people will reduce the economic returns to education (wage). Then, it is expected that the loss of returns to education will reduce the interest of individuals in investing in education. Individuals are supposed to be acting in a rational way trying to maximise the return to their investments, in this case the monetary and non-monetary investments in education.

Therefore, according to Human Capital Theory, in the long term, the perfect competition of the market will reduce the wages until the supply of and demand for highly skilled workplaces are balanced, that is, until the supply of and the demand for highly educated people match: over- as well as under-education disappear. The problem is that perfect competition does not exist, and there are always either some frictions or some imperfect information in the actual labour market. The other big issue with this approach is the definition of what can be considered as a long time in order to expect the labour market to recover its equilibrium, since current data are demonstrating that the mismatch continues year after year (see Chapter Three on the Spanish case).

The Career Mobility Theory also implies that over-education is a short term phenomenon. As seen above, according to this theory, over-education usually affects young people, the ones who despite having high educational levels enter the labour market for the first time and do not have experience. According to Sicherman and Galor’s theory, the youth have high formal education but they lack the specific training that is required to do the job (Sicherman, 1987; Sicherman and Galor, 1990). According to the Theory of Career Mobility this is the main reason why young people accept a job for which they are over-educated. They accept this lower skilled job for a short period, while they are acquiring the specific training that the work experience can give them. However, they also expect to move up the occupational ladder as soon as they become experienced workers, since they have attained a high educational level and they are able to work in more highly skilled jobs with higher wages and better conditions. Consequently, over-education will usually be a problem for younger generations, and over-educated workers will be more likely to change jobs, both within the company or between companies (Dolton and Vignoles, 2000; García Serrano and Malo, 1996; Groot and Maasen Van Den Brink, 1997; Sicherman, 1987).

In sharp contrast, Signal and Screening Theory identifies over-education as a permanent problem, basically caused by the growing importance of other high-level qualifications as a method of identifying more productive workers. According to Dore (1997), this credentialism will provoke a qualification escalation, because

> *"the mechanism of 'qualification escalation' ensures that once one is in the modern-sector-qualification range, the higher the educational qualification one gets the better one's chances of getting some job."* Dore (1997, p. 4)

Therefore, even if there is credentialism, in individual terms, increasing one's own educational level is still the best option, especially when the main direct cost is subsided by the state, despite the fact that it is not the best solution in social terms.

In most of the European cases, and the Spanish case in particular, over-education implies a waste of public investment, since, most of the current tertiary education institutions in Spain are from the state.[24] Therefore, on the one hand, public institutions, and individuals and their families would have been investing in education which will not be used in the labour market, and, on the other hand, workers' abilities and skills are underutilised by the firms. This makes firms lose productivity and efficiency. Besides, over-educated workers tend to be less productive and less satisfied with their jobs (Allen and van der Velden, 2001; Badillo Amador, López Nicolás and Vila, 2008; Green, McIntosh and Vignoles, 2002).

---

[24] In 2000, 39 out 60 of the Spanish universities were state universities.

One of the first studies analysing the effect of over-education on labour mobility was Sicherman (1987). Using data from the USA from 1976 to 1981, he analysed the over-education phenomenon to conclude that

> *"over educated workers have higher rates of firm and occupational mobility than other workers with similar characteristics (…)They are more likely to move to a higher level occupation."* Sicherman (1987, p. 18).

In the same sense, the results of Dolton and Vignoles (2000) reject the main hypothesis of the Human Capital Theory and support the Assignment Theory for over-educated graduates in the UK labour market. They confirm that

> *"both the human capital of the individual and the requirements of the job determine earnings"* Dolton and Vignoles (2000, p. 193).

They found that some over-educated graduates in their first job became adequately educated six years later. In particular, they found a reduction of 8 percentage points (from 38 percent to 30 percent) in the proportion of over-educated graduates between their first job and their labour market situation six years after (1980 to 1986).

Also using data from the UK, but in 1991, Groot and Maasen Van Den Brink (1997) confirmed the hypothesis of an adjustment period in the labour market for the new entrants. Therefore,

> *"even if over-education appears to be a permanent feature in Western economies, for individual workers it seems to be only a temporary phenomenon"* Groot and Maasen Van Den Brink (1997, p. 181).

Their results were similar to those of Dolton and Vignoles (2000). It is also worth noting that, according to Groot and Maasen Van Den Brink (1997) there is a gender difference in the reduction of over-education when the age and the experience in the labour market increase.

> *"A one-year increase in age and experience decreases years of over-education for men by 9.0% and for women by 1.5%."* Groot and Maasen Van Den Brink (1997, p. 179).

In Spain, most empirical analyses of over-education have centred on two main goals: first, calculating the degree or changes in education mismatch, and secondly, measuring the effect of this education mismatches on wages. The following paragraphs will focus on empirical studies on the Spanish labour market and on these two topics.

Most Spanish empirical literature focused on measuring the degree of education mismatch in the country has used subjective methods. One of the first studies on education mismatch using Spanish data is Alba-Ramírez (1993), which used a representative and exceptional survey of the Spanish labour force (Living and Working Conditions Survey, 1985) to quantify the over-education of the Spanish labour market. According to the answers of the workers, around 17 percent of the salaried people working in the Spanish labour market were over-educated in 1985.[25] In particular, individuals who were young, less experienced, or with less on-the-job training, as well as those who hold a job with low tenure were more likely to be over-educated. These results should be viewed with caution, since the data is from the Spanish labour market in 1985, when the expanding numbers of university graduates had not entered the labour market yet. More recent results do not differ substantially from this first analysis (see, the third chapter of this thesis).

Along with the increase of the university graduates at the end of the 1990s and the beginning of the 2000s, the Spanish literature focused on labour market

[25] Living and Working Conditions Survey is a special nation-wide representative household survey done in the last quarter of 1985. It contains workers' self-assessments about their workplace and their own skills or educational levels, as well as, the degree of match between them. The sample included more than 60,000 individuals older than 14 (Alba-Ramírez, 1993). See also http://www.ine.es.

mismatches becoming more substantial.[26] Subjective measures have concluded a share of over-educated people between 25-30 percent at the beginning of the 90s (Beneito *et al.*, 1996; García Serrano and Malo, 1996) and 53.8 percent in 1995 (Alba-Ramírez and Blázquez, 2003). When objective measurements were used, the share of over-educated people was estimated to be below 10 percent of salaried people: 3.7 percent in 1985, 6.6 in 1991 and 7.7 in 1993 (García-Montalvo, 1995). [27]

In order to interpret the previous results, it is worth being aware that subjective methods tend to overestimate the degree of over-educated people, whereas the objective measurements tend to underestimate it. It appears quite reasonable to conclude that:

(a) there was in the Spanish labour market a proportion of workers who were over-educated during the 90s, even if it is not easy to estimate accurately these proportions;

(b) the proportion of over-educated workers in the Spanish labour market increased during the decade, alongside the increase in the supply of highly educated workers. These conclusions have been also confirmed for a more recent period.

---

[26] In this period, the number of university students almost doubled. At the end of the academic year 1999-2000 there were 211,312 university graduates, 83.5 percent more university graduates than in 1988-1989 (115,135). The highest increase (98.4 percent) was in the graduates with a short-degree (three years of university): from 42,106 short-degree graduates in 1988-1989 to 87,173 in 1999-2000. Degree graduates (four to five years of university) increased by 71.5 percent, and 62 percent had doctorates (tertiary education).

[27] For a summary of the literature measuring the mismatch in the Spanish labour market since 1993 see Aguilar Ramos and García Crespo (2008); Fernández (2004) and Sanromá and Ramos (2004). Their classification is based on (a) the methodology to measure the job mismatch used (objective, subjective or statistical), (b) the survey, and (c) the year that the survey was carried out. The three studies consider the proportion of over-educated, under-educated and adequately educated. They all agree that the proportion of mismatches in the Spanish labour market, highly depend on the way the mismatch was measured. Subjective and statistical measurements usually give higher proportions of mismatch than objective measures.

Using a different dataset Parellada, Duch and Alvarez (2009) have evaluated the match for university graduates in the Spanish labour market in 2006. Their dataset uses the job offers by the firms. The source is “*the most important virtual employment Website in Spain*” (Parellada, Duch and Alvarez, 2009: 37, endnote 4). The total dataset contains more than two million job vacancies and more than three million applicants (15.5 percent of the Spanish labour force in 2006). The authors focus on the supply of and the demand for university graduates, 18.4 percent of the two million vacancies. With an average of 53.7 percent of university graduates applying for one university job offer, they conclude that in 2006 there was over-education in the Spanish labour market. According to them, this over-education was caused by a surplus of supply of university graduates (Parellada, Duch and Alvarez, 2009). These results confirm the evidence found by other studies from previous periods.

Using data for workers aged 25-64, in the same period (2006-2007), Nieto (2014) applied statistical methods. She found that, when the mode (mean) was used, 25 percent (20 percent) of the population was over-educated in the Spanish labour market. Using data from the PIAAC survey, Nieto (2004) calculated for a more recent period, 2011-2012, that the proportion of over-educated individuals in the Spanish labour market rose to 35 percent when statistical mode approach was used and fell to 15 percent when statistical mean approach was applied. The proportion of over-educated people using the same data with subjective methods is also around 35 percent.

It seems evident, that the proportion of over-educated people in the Spanish labour market has been rising along with the incorporation to the labour market of the more educated youth from the baby boom, as well as the higher participation of women into the labour market (San Segundo, 1997).

Murillo, Rahona and Salinas (2010) estimated education returns for Spanish data from 1995, 2002, and 2006. They apply first the Mincer’s equation to analyse the changes in the pecuniary returns to education, to conclude that the returns to education descend from 9.5 percent in 1995 to 7.4 percent in 2006. The decrease in these returns confirms the existence of over-education. Second, they use the ORU model to measure the effect of over-education on wages. Again, their results are consistent with the literature: there is a pay penalty for over-educated workers. The same result is found by Nieto (2014) using data from 2007 and 2011-2012. She estimated a pay penalty of 1.3 percentage points in 2007 and 4.2 percentage points in 2011-2012 for over-educated workers.

Other authors have measured the wage penalty (premium) for over-(under-) educated workers for Spain. Aguilar Ramos and García Crespo (2008) used panel data (from 1995 to 2001) to evaluate the pay penalty associated with education mismatch. Following Chevalier (2003), they used subjective methods to classify individuals into two different types of over-educated: those who are ‘really’ over-educated and those who are ‘apparently’ over-educated. They concluded that those ‘apparently’ over-educated have a higher pay penalty. This suggests that it is important to take into account heterogeneity among the over-educated, since, according to their results, those who are not actually over-educated (because they are in fact less skilled) are more punished in the labour market.

Research looking at the effect of over-education on labour mobility can also be found for Spain. However, the results are not as conclusive as the ones for the UK. For instance, García Serrano and Malo (1996) analysed the labour mobility of over-educated workers in the Spanish labour market in 1991. They stated that there is a positive relationship between being over-educated and inter-firm mobility. However, they cannot conclude that over-education was a temporary phenomenon since they do not find a relationship between intra-firm mobility and education mismatch. Similarly, Alba-Ramírez and Blázquez (2003) found that the likelihood of changing firms or occupation depends on the type of over-education (they do not have the skills, they have different skills, they have similar skills), being higher for those whose formal education skills are not related to the skills required in the job (area or field of study mismatch).

Both the international studies and those focusing on Spain agree that over-education is a phenomenon which affects individuals differently, depending on their personal and temporary characteristics. It seems that a young single woman, with low experience and low on-the-job training and holding a temporary contract was more likely to be over-educated in the 1990s (Alba-Ramírez, 1993; García Serrano and Malo, 1996). But what about recent years? Murillo, Rahona and Salinas (2010) estimated there is still over-education in Spain and that the over-educated suffer a pay-penalty. But, who are more likely to be over-educated? What are their characteristics? To what extent does being a man rather than a woman alter the probability of being over-educated? This thesis attemps to shed some light to these questions.

1.3. Concluding remarks

As has been explained in this chapter, the positive and causal relationship between education and the improvement of the individual's abilities was the mainstream thought of the first economists, from Adam Smith to Gary Becker, who showed some concern about the economic value of education. According to them, the labour market will return the individual effort (pecuniary or not) in improving their abilities and skills, accommodating them in a job with a better wage or/and better conditions. The first empirical studies seemed to corroborate this, since those with higher educational levels used to have higher returns, especially higher wages, although they also had better job conditions: more stable contracts or better conditions in the workplace. There were two main reasons: first, higher education workers are supposed to be more multi-purpose and easily adapt to technological or organisational changes; and second, the supply of highly skilled workers in the previous generations was relatively low.

However, the upgrade of the supply of skilled individuals due to the spread of first literacy, then compulsory education, and secondary and above all, tertiary education later on, increased the concerns about the positive and causal relationship between higher educational levels and the improvement of the individual’s abilities or their higher returns in the labour market (Freeman, 1976).

All in all, the main problem of this educational upgrading was that, apparently, labour markets were not creating skilled jobs for those more educated individuals at the same pace. As a result, an imbalance has been coming into sight in the labour market: a surplus of educated population (supply of skilled workers) or a deficit of skilled jobs (demand for skilled workers).

As a result, the greater the supply of higher educated workers, the stronger the competition to get a job in the labour market. They should compete to find a job, and then to find a job according to their educational level. Therefore, even if a skill-biased technical change occurs increasing the demand for highly skilled workers, the increase in the supply of these sort of workers will imply a higher competition between them, for instance, for finding a job according to the field of study of their degree. Besides, not only individuals compete with those from the same generation, but also, with those from other ones, even if they hold different educational levels. The reason: the oldest generations (50-65 years old) counterbalance part of their lower formal education level with more experience and on-the-job training.

Some theories, such as Human Capital Theory or even more recently, Theory of Career Mobility, state the temporary aspect of the over-education phenomenon. However, empirical studies and recent data have demonstrated that this is becoming a permanent situation for workers in most saturated labour markets. The excess of supply of highly skilled workers jointly with the continuous new university graduates entering the labour market and the little creation of highly skilled jobs have made the over-education into a sort of trap which is impossible to get out of.

Since the democratisation of tertiary education in most developed countries and the consequent increase in the number of students enrolled in these higher educational levels, the economic studies evaluating first the returns on education, and later the (im)balances in the labour market have mushroomed.

First, empirical studies have demonstrated that, at the beginning of the spread of tertiary education, the higher the educational level, the larger the returns on education and the better the conditions of the jobs they secure in the labour market. But, eventually, the continuous rise in the supply of higher educated workers, while the demand for them in the labour market grew less, made evident the difficulty the labour market had to accommodate all of the supply of these skilled workers. As a consequence, labour market mismatches occurred.

Studies focused on measuring the level of this mismatch have increased. They initially measure the consequences of the education mismatch, since there were not available data to measure individuals' skills. When data were finally available, they shifted to the analysis of the skill mismatch. Most of these studies have analysed the consequences that these education or skill mismatches have on wages, productivity, or even health, but the extent to which individuals are more likely to be over-educated has attracted less attention.

Despite the fact that there are few studies analysing the main characteristics of individuals more likely to be over-educated, identifying them is essential to reduce the proportion of these mismatched workers in the labour market. For instance, knowing whether there is a gender difference in the likelihood of being over-educated, or whether over-educated women have different characteristics from over-educated men would help policy makers to design and implement more efficient policies to reduce it.

To what extent the probability of being over-educated has changed over time has not been measured for the Spanish case for a long time. Besides, empirical studies have rarely analysed men and women as different subgroups. They have studied either only men or men and women together as a unique group including in the analysis a dummy covariant accounting for gender.

Furthermore, the datasets used in this thesis have been unused to measure the degree of over-educated workers in the Spanish labour market. The Living and

Working Conditions Survey, the European Community Household Panel or the Spanish Family Budget Survey have been the most analysed.

Regarding individuals' skill heterogeneity, the Survey of Adults Skills (SAS-PIAAC) conducted by the OECD between 2011 and 2012, has given an opportunity for researchers to evaluate the knowledge of cognitive skills as well as the computer skills of the adult population, among others. Besides, these data have been harmonised among countries, thereby allowing for comparisons among countries.

As has been shown, the literature measuring and analysing the education mismatch in the labour market is large. The aim of this thesis is to contribute to this important literature background with the analyses of three important aspects which have been less considered by the literature. First, the analysis of the characteristics of those over-educated workers in the Spanish labour market and their tendency over time. Second, what are the effects that accounting for observable cognitive skills has on the probability of being over-educated. And third, forecasting the number and proportion of over-educated university graduates in the Spanish labour in the next decade. Furthermore, data have been reported by gender in order to see whether there are significant differences between men and women.

“To a large extent, therefore, education has played a significant role as a buffer against unemployment. At the same time, this increase in the general level of education of population will have a positive influence in the long run, as productivity and ability to learn of future workers will have been enhanced by this extension of the educational system.”

Luis Toharia (1997, p. 181)

Chapter 2. Educational level and labour market indicators in Spain

Following the same pattern as other European countries, in the last decades of the 20th century, the educational level of the Spaniards increased. A generational substitution of the oldest less educated Spaniards by new generations of higher educated ones entering the labour market changed the characteristics of the supply of workers. It seems that the demand for these higher educated workers did not follow the same trend. This chapter introduces the main characteristics of the Spanish labour market to contextualise the thesis in the Spanish economic and historic domain and to better understand the current Spanish labour market.

2.1. Introduction

To contextualise the research in the Spanish economic and historic domain, this second chapter presents the Spanish historical context and the main indicators of the Spanish labour market, in absolute terms and in relation to other European countries. Knowing the political, social and economic context is important to understand the main trends of the labour market. Legislation, culture, but also the economic situation are factors that affect both sides of the labour market: demand and supply.

The demand for one particular type of worker (low skilled, skilled, highly skilled) will not only be influenced by the production system, but also by the economic moment. The level of technology on the production system will determine the kind of task needed by the job, and therefore the qualification required for doing this job. For instance, according to the Skill-Biased Technical Change theory (SBTC) or the Routine-Biased Technical Changes theory (RBTC), the new technologies will change the sort of workers demanded by the firms. However, their effect is undetermined, since they can follow two different directions. Researchers including Acemoglu (2002b) and Goldin and Katz (2008) explain how technical changes can have different effects on the demand for skilled workers, as for instance, happened with the technological change of the 19th and 20th century. According to Acemoglu (2002b), the demand for more or fewer skilled workers is related to their supply, and therefore, the direction of the new technology will depend on their cost-effectiveness. He explains that in the 19th century there was a higher supply of less-skilled workers which made it more cost-effective to introduce new technology which was more labour-intensive and required workers with lower abilities; whereas in the 20th century, the implementation of skill biased technologies was more profitable since the supply of highly skilled workers was larger (Acemoglu, 2002b).

Focusing on the end of the 20th century and the beginning of 21st, the effect that the constantly changing technology will have on the economy is not completely clear. On the one hand, firms may need and demand higher proportions of highly skilled workers, since they are supposed to be more versatile than the ones with lower skills. Therefore, highly skilled workers will be abler to adapt to job changes or to learn new techniques quickly. On the other hand, these new technologies might also increase routine-based jobs. The effect of these changes on developed economies is not clear, since it is also likely that these routine-based jobs will be moved to less developed countries with lower labour costs due to the economy globalisation.

The economic moment is the other main factor which determines the level of demand for workers. Whether there is an economic expansion or an economic crisis will not only determine the level of the demand for workers, but also the level of skills required for the jobholders. It is especially relevant in the Spanish case, since, in general terms, the demand for unskilled or low skilled workers increases with economic expansions (Muñoz de Bustillo Llorente and Antón Pérez, 2011; Torrejón, 2019). Traditionally, and especially in the last decades of the 20th century and the beginning of the 21st, the construction and services sectors have been the motors of the Spanish economic expansions, and these sectors are more likely to require low skilled workers. In fact, some studies have also found a positive relationship between economic expansions and the increase of the proportion of early-school-leavers who are supposed to be lower skilled (Aparicio, 2010; Choi de Mendizabal and Calero Martínez, 2018; Felgueroso, Gutiérrez Domènech and Jiménez-Martín, 2013).

Regarding the supply of workers and the characteristics of the job searchers, they will be determined in part by the individuals' abilities. As seen in the literature review chapter, individuals' abilities depend on their innate capabilities, but also their educational level and their experience. However, it is not easy to evaluate the individuals' innate abilities and their experience, therefore, the literature has usually used the formal educational level to evaluate them and to determine the individuals' skills level.

In turn, individuals' educational level is the result of the individuals' decisions but also of how easy it is for them to get a diploma. Therefore, opportunity to access the education system, for instance the supply of schools or universities, will also influence the educational level held by the workers. In this sense, the creation of new universities and new degrees in the 80s and the 90s in Spain propelled part of the growth of the educational level of the labour force. This chapter aims to cover all these trends.

The rest of the chapter is organised as follows. Using a historical perspective, Section 2.2 introduces the main political, social, and economic changes in Spain in the last years of the 20th century. Section 2.3 uses data from Eurostat to place the Spanish labour market in a European context. An analysis of the educational level of the labour force and their employability in the Spanish labour market from the last years of the 20th century until nowadays is presented in Section 2.4. Individual data from the Spanish Labour Force Survey (SLFS) from 1995 to 2018 is used in this latter section. Section 2.5 introduces the concept of education (im)balance and a first approximation to its measurement in the Spanish labour market. The chapter finishes with Section 2.6, where the main conclusions are presented.

2.2. Historical context

The Spanish case is unusual in the general European context because of the negative consequences of the political, social and economic legacy of Franco's dictatorship (1939-1975). Politically, Franco's dictatorship finished when he died on 20th November 1975 and the first democratic elections were called on 15th June 1977. However, the economic changes started just shortly before this. The end of national self-sufficiency (in 1959) and the introduction of the *Development Plans,* in the decade of 1960, helped the economy develop, and eventually the Spanish economy achieved the industrial development that had not been possible at the beginning of the 20th century.[28] In less than three decades (1959-

[28] It was the "Desarrollismo" period (1960s), the "golden period" of the Spanish economy. Industrialisation was the motor of economic development. The annual GNP growth-rate was greater than 4 percent, with two digit-rate increases in some of the years. In 1986 Spain joined the EEC. For a more detailed description of the Spanish economic development, see, for instance, Carreras and Tafunell (2006, pp. 331-364).

1986), Spain transformed itself from an autarchy (national self-sufficiency) under a dictatorship to a democratic country which later joined the EEC.

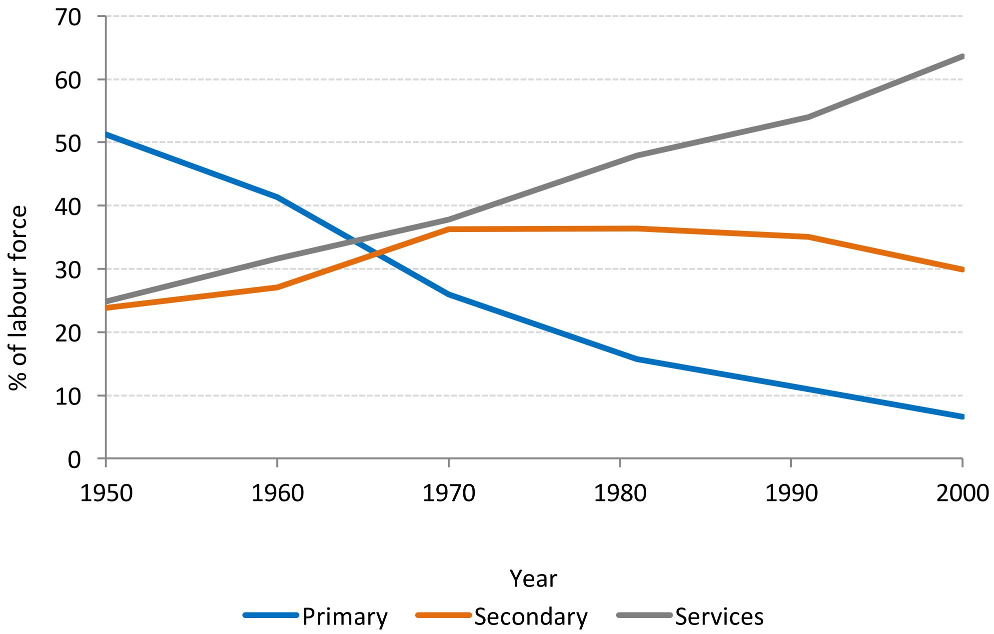
From the 1960s, several circumstances occurred simultaneously provoking a feedback process: industrialisation and the increase in the labour market demand for skilled workers due to the skill-biased technical change, rural exodus increasing urbanisation and population density, demographic change with a baby boom as a consequence of the increase in living standards and the increase in people´s expectations about their future, and the growth in educational levels among the population first in primary and secondary education, and later in tertiary education, and especially in university education. All these trends together contributed to the creation of a modern developed country in the last half of the 20th century.

Even though it was later in time, the Spanish industrialisation followed the same pattern as other European countries. The rural exodus led to a concentration of the population in cities and an increase in the active population working in industrial and services sectors. In 1960, 14.5 percent of the Spaniards were still living in towns with fewer than 2,000 inhabitants and less than 60 percent of the labour force were working in industrial and services sectors. By 2000, the percentage of Spaniards living in small towns had reduced to 7.5 percent, and the proportion of the labour force working in the secondary and tertiary sectors increased by 94 percent.[29]

Increased urbanisation and industrialisation led to skill-biased development, and consequently better-skilled workers were required in the labour market. In 1950, the GDP distribution by sector was 30.7 percent in the primary sector, 26.9 percent in the secondary sector and 42.4 percent in the services sector. In 1970 it changed to 11.9 percent, 36.9 percent and 51.2 percent, respectively (Carreras and Tafunell, 2006, p. 454).

[29] http://www.ine.es/inebaseweb/pdfDispacher.do?td=158004&ext=.pdf; http://www.ine.es/prodyser/pubweb/anuario01/01demog.pdf (p. 55) and Nicolau (2005, p. 149).

Figure 2.1 shows the labour force distribution by economic sector from 1950 to 2000. In 1950, the labour force distribution by economic sector was 51.3 percent in the primary sector, 23.9 percent in the secondary sector and 24.9 percent in the services sector. In 1970, the labour force working in the primary sector was 26 percent, whereas the workers in the secondary and the services sectors were 36 percent and 38 percent of the total labour force, respectively (Nicolau, 2005, p. 149).



Source: Nicolau (2005, p. 149)

Figure 2. 1. Spanish labour force distribution by economic sector. 1950-2000

All these factors together contribute to a structural change in the Spanish economy in the last third of the 20th century. According to different economic historians' contributions, the main social and economic characteristics of this modernisation process can be summarised in the following four points (Carreras and Tafunell, 2016; Maluquer de Motes, 2014; Myro, 2005):

1. Changes in the productive structure, with the decline of agriculture and the increase of the secondary sector and services, in terms of GDP and labour

force participation. Secondary and services sectors became more capital intensive and therefore there was a rise in labour productivity.

2. Commercial openness of the Spanish economy, stimulating a specialisation, in labour-intensive industries and tourism.
3. Rising importance of public administration. The constitution of a democratic government and the welfare state implied an increase in the public investment (infrastructures and social investment) and in the demand for civil servants. The demand for more educated people increased during the democratic transition (1976-85). Besides, at the end of the 1970s and beginning of the 1980s there was the creation of 17 regions (Autonomous Communities: NUTS-2). As a consequence, part of the policy-making was slowly decentralised in the new autonomic governments, increasing also the demand for administrative civil servants.
4. The income and wealth distribution has become more equal among individuals and territories. The reason was the introduction of the welfare state and the new fiscal system, based on the introduction of three new taxes: personal income taxes, company taxes, and VAT.

These structural changes propelled an increase in production and a considerable economic growth which raised the Spaniards' living standards. Regarding the characteristics of the labour force, the structural changes demanded workers with at least a minimum knowledge of literacy and numeracy, the kind of proficiency usually taught in primary or secondary education. It was therefore necessary for political actions to increase the literacy and numeracy levels of the general population. The political decision to improve people's literacy was reflected in the new legislation. In 1970, the General Education Law (Ley General de Educación) was passed, which declared full time education until 14 years old compulsory.[30] It was not until the 1990s that the implementation of the LOGSE (General Education System Organic Law) made second grade of secondary education

[30] The previous law was implemented in 1870 (Moyano Law – Ley Moyano) and despite the fact that it made full time education until ten years old compulsory, it was barely enforced.

compulsory, and thereby increasing the years of compulsory full-time education until the age of 16.[31]

As a result, in the last forty years of the 20th century the level of education in Spain increased spectacularly, first in the proportion of secondary education graduates, then, the main increase was in the proportion of students holding a university degree. In the academic year 1980-81 only 21 percent of the 18-year-old population entered the university system in Spain, whereas in 1999-2000 this percentage doubled and rose up to 55 percent (García-Montalvo, 2001). The increase is similar when looking at the proportion of the Spanish labour force holding intermediate and high education levels. While in 1980, only 17.8 percent of the labour force had finished at least intermediate level studies and only 2.5 percent had higher education; in 2000 these percentages were, respectively, 43 percent and 12 percent (Albert *et al.*, 2003). As it will be explained later in Section 2.4 of this chapter, the growth in the number of university graduates was not only due to an increase in the demand for this level of studies, but also to the increase in the supply.

2.3. The Spanish labour market in an European context

Comparisons of the structure of the labour markets are usually made between more regulated European regimes and the unregulated regime of the USA. However, this chapter focuses on five European countries. This analysis will be useful to reassert the heterogeneity amongst the labour market of the European countries. Regarding the characteristics and structure of their labour markets the EU countries can be classified in four main categories: Scandinavian countries, Continental Europe, Southern Europe, and Anglo-Saxon countries (Esping-Andersen and Regini, 1999; Vaughan-Whitehead, 2011).

[31] Before LOGSE, in Spain only primary school was compulsory (students finished this education level when they were 13-14 years old). Primary education had two different cycles: medium and superior, but both of them were considered as primary education.

Apart from Spain, which will be included within Southern Europe countries, and following the previous classification, the four countries selected would contain at least a sample of all the categories: Scandinavian countries (Sweden), Continental Europe (France), Southern Europe (Italy), and Anglo-Saxon (the United Kingdom).

For instance, a first difference between these four groups would be the degree of the state's participation in the regulation of the labour market: an unregulated regime in the Anglo-Saxon group, a more regulated one in Continental and Southern Europe, and a third regime in between, in the Scandinavian countries.

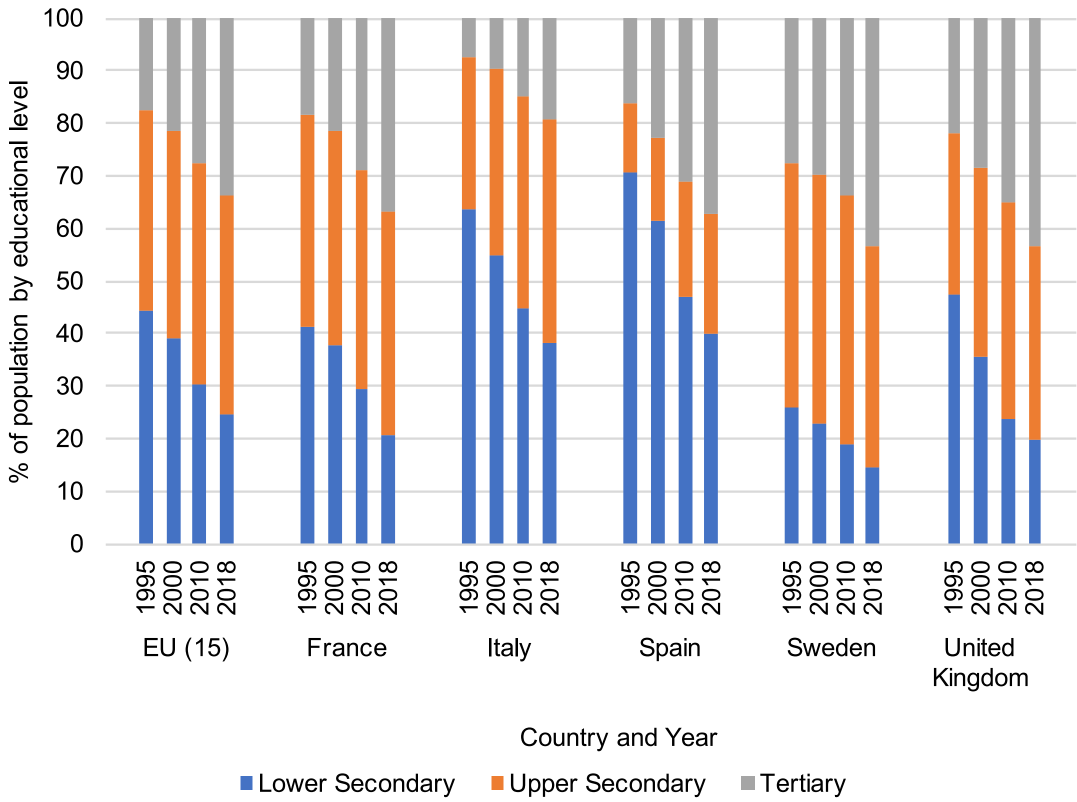
Despite the fact that recently, most European countries, following the EU directives, have made their labour markets more flexible, there are still some differences among them. The structure of the labour market is a consequence of the history, culture, society, policy, and economy of each country. However, going deeper into these differences is not the objective of this research and, therefore, the discussion will not extend further on this topic. Nevertheless, it is worth mentioning some of their differences, for example the lack of hire and dismissal regulations in the English labour market since the enactment of the neo-liberalism and anti-union legislation of Margaret Thatcher in the 1980s; whereas in Spain, despite the attempts to make the Spanish labour market more flexible, it is still tied to an important central collective bargain, as well as, a regional and sectoral ones, which makes the functioning of the labour market more rigid (Eichhorst and Marx, 2016; Esping-Andersen and Regini, 1999).

At the beginning of the 1980s, the main belief in the literature on labour market policies, but also among politicians, was that the rigidities of the Spanish labour market were the main cause of these unsuccessful outcomes of the Spanish labour market. However, the new legislation enacted during the 80s and 90s which was intended to make the labour market more flexible did not solve the problem. The new regulations allowed employers to use temporary contracts in a laxer way. A worker could, then, be hired for three years using consecutive fixed-term contracts. At the end of this three-year period, the contract had to become a permanent one or the worker had to be dismissed. In the case of dismissal, the redundancy pay for a fixed-term contract would be lower than for permanent contracts. As a result, a new troublesome consequence appeared in the Spanish labour market: the increase of these temporary contracts. Moreover, the increase of the uncertainty due to these fixed-term contracts had negative consequences for the individuals as well as for society.

Apparently, in most cases, the recent EU directives attempting to relax the labour market regulations have only worsened the worker's conditions or weakened the worker's rights, instead of making the labour market more flexible. And this deterioration has been even worse for specific groups such as the young, migrants, or women, those who usually already have the poorest indicators in the labour markets (Malo, 2015; Molina and López-Roldán, 2015; Muñoz de Bustillo Llorente and Antón Pérez, 2011).

To place Spain in a European perspective, Figure 2.2 shows the changes in the distribution of the population aged 25-64 by their educational level in 1995, 2000, 2010, and 2018 in five European countries, as well as for EU-15. It shows a convergent tendency between the proportion of people with tertiary education in Spain and in other European countries excluding Italy, which is still showing relatively low proportions of people who have finished tertiary education. However, unlike the rest of the European countries included in the chart, Spain still has a higher proportion of people aged 25-64 with educational levels lower than secondary education, and lower proportions of people aged 25-64 with upper-secondary education.

The change in the distribution of the population by educational level has been more noticeable in Spain and in Italy. In 1995, the proportion of the population with lower secondary education was around 20 percentage points (pp) higher in these countries than the average of the EU-15. In 2018, this difference reduced to only 10 pp.



Source: Eurostat (http://epp.eurostat.ec.europa.eu)

Figure 2. 2. Distribution of population (25-64 years old) by country and educational level. 1995-2018

Data from Figure 2.2 indicate that there is a polarisation of the population in relation to their educational level in the Spanish labour market. There are three main factors which contribute to this polarisation:

(a) a high proportion of school failure in the early educational stages. The percentage of early leavers from education and training in Spain in 2018 was 18 percent of the total population of the same age group (18-24-year-old), 14.5 percent in Italy, 10.7 percent in the United Kingdom, 8.9 percent in France, and 7.5 percent in Sweden.[32] Furthermore, in the Spanish case, this percentage growth with economic expansions due to the higher demand for low-skilled workers especially in the construction and services

[32] http://ec.europa.eu/eurostat/data/database.

sectors (Aparicio, 2010; Choi de Mendizabal and Calero Martínez, 2018; Felgueroso, Gutiérrez Domènech and Jiménez-Martín, 2013);

(b) a high proportion of people enrolled in university studies. In 2017, 38.9 percent of the 18-year-old individuals were enrolled in tertiary education in Spain, 0.8 percent in Sweden, 3.7 percent in Italy, 33.6 percent in the United Kingdom, and 43.2 percent in France;[33]

(c) there are few people enrolled in secondary education, and especially in vocational secondary education. In 2017 in Spain, 35.2 percent of the students were enrolled in the vocational upper secondary education; 34.1 percent in Sweden, 39.9 percent in France, 46.6 percent in the United Kingdom; and 55.3 percent in Italy.[34] It is worth noting here that the Spanish governments (regional as well as national) have attempted to promote this vocational secondary education in different occasions. However, until now, this has not been successful, and the supply of workers with vocational education is still being insufficient to cover the demand for such jobs in the Spanish labour market. Consequently they have to be covered by other educational groups and a higher investment in on-the-job training is necessary. In fact, in 2020, the regional government of Catalonia has tried to propel the vocational upper secondary education, and has launched an advertising campaign to inform the population about it.[35]

However, and despite the recent economic crisis initiated in the United States in 2007, the Spanish labour market indicators have improved. A part of this amelioration can be explained by the reduction of the labour force. The number of new migrants entering the Spanish labour market had reduced constantly. In 2003, 706,224 foreigners migrated to Spain; 599,074 in 2008; 280,772 in 2013.[36] Besides, the return to their countries or the migration to other European countries of part of the migrants who had lost their job in Spain during the crisis was one of

[33] Ibid.
[34] Ibid.
[35] http://queestudiar.gencat.cat/ca/estudis/fp/
[36] http://www.ine.es.

the causes. In 2008, 288,432 migrants left Spain, whereas in 2013 they were 532,303 (84.5 percent of increase).[37] But the migrants have not been the only ones who moved, there has also been an increase in the number of Spaniards who emigrated to other countries. In 2009, 35,372 Spaniards emigrated to another country, in 2016, 86,112.[38] These migrations explain, for instance, the increase in the employment rate and the decrease of the unemployment rate (Tables 2.1 to 2.3).

To put the Spanish statistics in a European perspective, the next three tables display the main labour market indicators for EU-15, France, Italy, Spain, Sweden, and the UK. Table 2.1 shows the total data, whereas Table 2.2 presents data for men, and Table 2.3 for women.

In Spain, the activity rate is the only index which has increased enough to converge with the European Union (EU-15) (Tables 2.1 to 2.3, and Fig. 2.3). As Table 2.1 shows, the Spanish activity rate for all the population in 2018 was 73.7 percent, 0.7 percentage points (pp) lower than EU-15 (74.4 percent), 4.2 pp lower than the UK (77.9 percent), and 9 pp lower than Sweden (82.7 percent), although it is 1.6 pp higher than France (71.9 percent), and 8.1 pp higher than Italy (65.6 percent).

---

[37] These data only include individuals with foreign nationality (not Spanish). Source: http://www.ine.es.
[38] http://www.ine.es.

Table 2. 1. Main labour market indicators by country. Total population, 1995, 2005, and 2018

		Activity rate	*Employment rate*	*Unemployment rate*	*Youth unemployment rate*	*Long-term unemployment rate*	*Part-time employment*	*Fixed-term contracts*
EU - 15	**1995**	67.2	59.9	10.8	17.3	47.8	15.6	11.5
	2005	71.0	65.2	8.2	13.6	41.7	19.6	14.4
	2018	74.4	68.8	7.6	12.8	44.1	22.5	14.5
France	**1995**	67.9	60.0	11.9	20.1	40.3	15.5	12.2
	2005	69.4	63.2	8.5	15.2	40.6	17.2	13.1
	2018	71.9	65.4	9.1	16.5	42.0	18.0	16.7
Italy	**1995**	57.7	51.0	11.8	25.2	54.3	6.3	7.2
	2005	62.5	57.6	7.8	17.7	49.7	12.7	12.2
	2018	65.6	58.5	10.8	24.8	59.0	18.4	17.1
Spain	**1995**	60.9	46.9	22.8	35.9	54.7	7.2	35.0
	2005	70.0	63.6	9.2	14.8	24.4	12.0	33.4
	2018	73.7	62.4	15.4	26.2	41.7	14.5	26.9
Sweden	**1995**	76.9[a]	69.4[a]	9.0	14.3	20.6	6.5[a]	13.0
	2005	78.2	72,3	7.9	17.3	17.8[b]	12.7	15.7
	2018	82.7	77.4	6.5	12.3	18.3	18.4	15.9
United Kingdom	**1995**	75.1	68.4	8.8	13.1	43.5	23.2	6.9
	2005	75.4	71.8	4.8	9.6	21.0	24.2	5.6
	2018	77.9	74.7	4.1	8.0	26.3	24.6	5.4

Note: ***Activity rate*** *(% population aged 15-64),* ***Employment rate*** *(% population aged 15-64),* ***Unemployment rate*** *(% labour force, all ages),* ***Youth unemployment rate*** *(% labour force aged 15-29),* ***Long-term unemployment rate*** *(% total unemployment aged 15-64),* ***Part-time employment*** *(% total employment aged 15-64),* ***Fixed-term contracts*** *(% total employees aged 15-64).*

[a] 1996; [b] 2004

Source: Eurostat (http://www.ec.europa.eu/eurostat/web/lfs/data/database)

Table 2. 2. Main labour market indicators by country. Men, 1995, 2005, and 2018

		Activity rate	*Employment rate*	*Unemployment rate*	*Youth unemployment rate*	*Long-term unemployment rate*	*Part-time employment*	*Fixed-term contracts*
EU - 15	**1995**	77.7	70.3	9.6	15.9	47.1	4.7	10.7
	2005	78.9	72.9	7.6	13.2	41.3	7.0	13.7
	2018	79.6	73.8	7.3	13.1	44.0	9.9	14.1
France	**1995**	75.1	67.9	10.0	17.2	39.3	4.8	11.3
	2005	75.0	68.8	7.8	14.9	38.9	5.8	12.5
	2018	75.8	68.9	9.1	16.9	43.3	7.8	16.1
Italy	**1995**	73.2	66.6	9.2	21.5	54.8	2.8	6.0
	2005	74.6	69.9	6.3	15.5	47.6	4.3	10.4
	2018	75.1	67.6	10.0	23.5	58.5	8.0	16.6
Spain	**1995**	75.8	62.1	18.0	30.3	48.9	2.5	33.3
	2005	80.9	75.1	7.2	12.5	20.5	4.4	31.8
	2018	78.8	67.9	13.8	26.2	39.5	6.7	26.0
Sweden	**1995**	78.7[a]	70.4[a]	10.1	15.3	23.6	8.1	11.0
	2005	80.5	74.3	8.0	17.2	19.4[b]	10.4	13.9
	2018	84.4	78.8	6.7	12.8	20.7	12.8	14.3
United Kingdom	**1995**	83.6	75.0	10.2	15.0	49.6	6.5	6.1
	2005	82.1	77.8	5.2	10.5	25.1	9.0	5.1
	2018	82.6	79.1	4.2	8.5	29.6	11.1	5.1

Note: ***Activity rate*** *(% men aged 15-64),* ***Employment rate*** *(% men aged 15-64),* ***Unemployment rate*** *(% men labour force, all ages),* ***Youth unemployment rate*** *(% men labour force aged 15-29),* ***Long-term unemployment rate*** *(% men unemployment aged 15-64),* ***Part-time employment*** *(% total men employed aged 15-64),* ***Fixed-term contracts*** *(% total men employed aged 15-64).*

[a] 1996; [b] 2004

Source: Eurostat (http://www.ec.europa.eu/eurostat/web/lfs/data/database)

Table 2. 3. Main labour market indicators by country. Women, 1995, 2005, and 2018

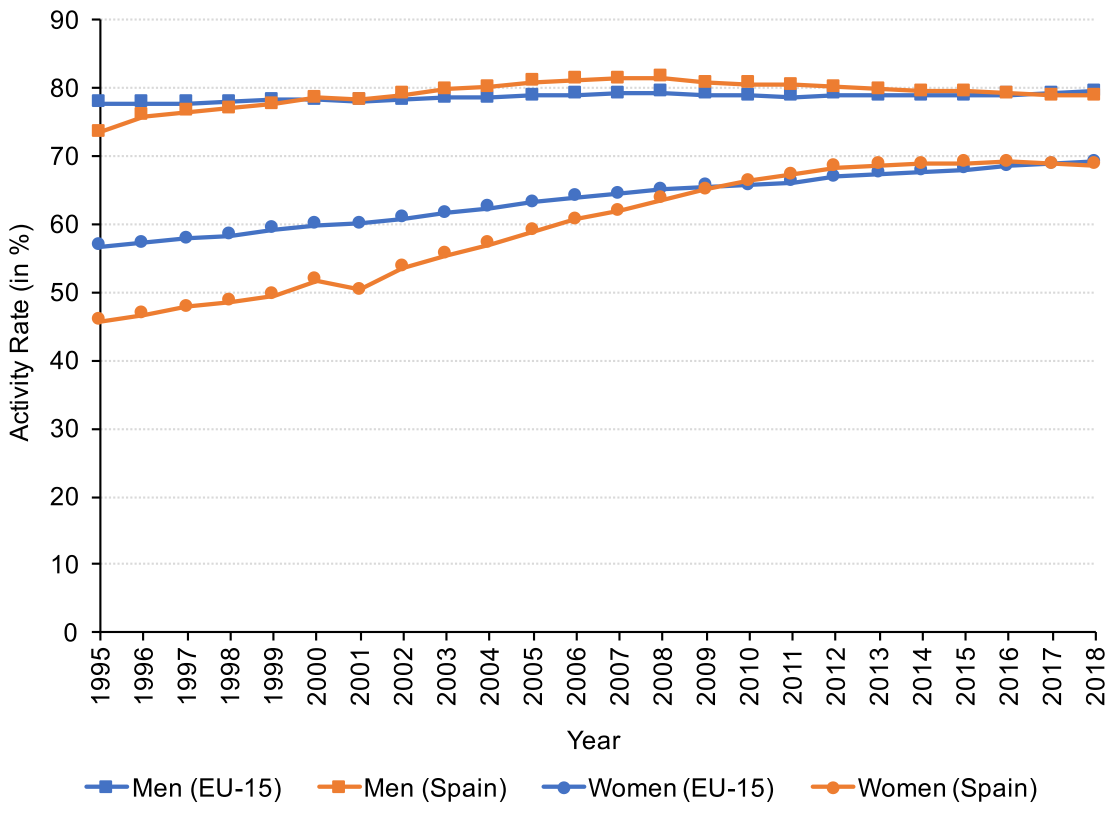
		Activity rate	*Employment rate*	*Unemployment rate*	*Youth unemployment rate*	*Long-term unemployment rate*	*Part-time employment*	*Fixed-term contracts*
EU - 15	**1995**	56.8	49.6	12.5	19.0	48.5	31.1	12.6
	2005	63.2	55.7	9.0	14.0	42.1	35.5	15.2
	2018	69.3	63.8	7.9	12.5	44.1	37.0	15.0
France	**1995**	60.9	52.5	14.1	23.3	41.2	28.8	13.4
	2005	64.1	57.9	9.3	15.5	41.6	30.3	13.6
	2018	68.2	61.9	9.2	16.0	40.6	28.8	17.3
Italy	**1995**	42.5	35.7	16.2	30.1	53.9	12.8	9.2
	2005	50.5	45.4	10.1	20.7	51.6	25.5	14.6
	2018	56.2	49.5	11.9	26.6	59.6	32.4	17.7
Spain	**1995**	46.1	31.8	30.8	42.7	60.3	16.3	38.3
	2005	57.9	51.8	12.0	17.6	27.6	23.4	35.6
	2018	68.6	56.9	17.1	26.2	43.7	23.9	27.8
Sweden	**1995**	75.0[a]	68.3[a]	7.9	13.4.	16.6	39.6	14.8
	2005	75.9	70.2	7.8	17.4	15.8[b]	37.7	17.6
	2018	81.0	75.9	6.3	11.8	15.5	33.2	17.6
United Kingdom	**1995**	66.5	61.7	7.0	10.6	32.1	43.7	7.8
	2005	68.8	65.8	4.3	8.5	15.1	41.8	6.2
	2018	73.2	70.3	4.0	7.5	22.5	39.7	5.8

Note: ***Activity rate*** *(% women aged 15-64),* ***Employment rate*** *(% women aged 15-64),* ***Unemployment rate*** *(% women labour force, all ages),* ***Youth unemployment rate*** *(% women labour force aged 15-29),* ***Long-term unemployment rate*** *(% women unemployment aged 15-64),* ***Part-time employment*** *(% total women employed aged 15-64),* ***Fixed-term contracts*** *(% total women employed aged 15-64).*

[a] 1996; [b] 2004

Source: Eurostat (http://www.ec.europa.eu/eurostat/web/lfs/data/database)

Figure 2.3 shows the convergence between the Spanish and the EU-15 activity rates from 1995 to 2018 by gender. In 1995, the activity rates for women were around 10pp lower in Spain than in the EU-15, and in 2010 they had already converged. The higher convergence between Spain and EU-15 in the women's activity rates compared to men's ones reflects the increase in the women's participation in the Spanish labour market during the last decade of the 20th century and the beginning of 21st. The higher educational levels held by women, and therefore the increase of their opportunity costs of being out of it is without doubt one of the main explanatory factors. In fact, literature related to the Spanish labour market stated the higher likelihood for university-educated women to participate in the labour market, although, they also held higher probabilities of being unemployed (Albert *et al.*, 2000).



Source: Eurostat (http://www.ec.europa.eu/eurostat/web/lfs/data/database)

Figure 2. 3. Activity rate by gender, 1995-2018

But the improvement of some of the indicators of the labour market during the last years and the upward tendency of the educational level of the labour force and their convergence in some of these with other European countries do not seem to be enough. Spain still diverges from the main European countries in terms of labour market indicators such as salaries, activity rates (particularly in women's activity), unemployment, productivity and proportion of fixed-term contracts.

For instance, Spain still has one of the highest unemployment rates, for both men (13.8 percent in 2018) and women (17.1 percent in 2018), and especially for the youth (26.2 percent in 2018, both for men and women) (Table 2.1 to 2.3). These differences are also important in terms of wages. The hourly wage in 2011-12 was around 13 Euros per hour in France and the UK (13.3 in France and 12.6 in the UK) and 12 Euros per hour in Italy. In Spain, the average hourly wage in 2011-12 was only of 9.7 Euros per hour (10.3 for men and 9.1 for women).[39]

These data show that the problem of the Spanish labour market is still structural, since it does not improve when the economy is growing. In fact, if the productivity of employment is considered, it has usually increased when there was a crisis and it deteriorated when there was economic growth. The main reason is that during the economic growth low productivity jobs such as the building and services sectors are created, and those are also the first jobs to disappear when there is an economic crisis (Muñoz de Bustillo Llorente and Antón Pérez, 2011; Torrejón, 2019).

The previous tables shed some light on the main problems of the Spanish labour market. Despite the improvement of the activity and employment rate in the last years, as it has been already commented, the rest of the labour market indicators show worse values in the Spanish case.

Unemployment, in general, and youth unemployment and long-term unemployment, in particular, are some of the main issues in the Spanish labour market. All the Spanish unemployment indicators have decreased from 1995 to

[39] PIAAC. Survey of Adults Skills (individual data from public use files).

2018, but they are still among the highest rates in the UE-15. The UK and Sweden have been able to maintain their unemployment rates in single figures, whereas the EU-15 and the rest of the countries listed in the table, including Spain, have been able to reach single-digit unemployment rates in 2005. However, in 2018, the unemployment rate in Italy and Spain has risen again to values in double digits. As expected, the results are even worse when focusing on the youth unemployment rate. In this latter case, only the UK has succeeded in reducing the rate to values below 10 percent in both 2005 and 2018. For Spain, the youth unemployment rate is the highest reaching 35.9 percent in 1995 and 26.2 percent in 2018. Therefore, unemployment and the difficulties for the Spanish labour force to find a job is one of the biggest problems of the Spanish labour market. It does not fall very much in expansionary periods but rises rapidly in recessions.

The long-term unemployment rate is also an important indicator to take into account since it can become a trap for those who suffer from it. From 1995 to 2018 the long-term unemployment rate decreased in Spain by 13 percentage points (pp), and as a consequence, this latter year, Spain managed to reduce the difference with the EU-15 to only 2.4 pp. In 2018 Spain, the aggregate long term unemployment rate was 41.7 percent (39.5 percent for men and 43.7 percent for women), 0.3 pp lower than France (42 percent) and 17.3 pp lower than Italy (59 percent). The United Kingdom (26.3 percent) and above all Sweden (18.3 percent) have a lower long-term unemployment rate (Tables 2.1 to 2.3).

The reduction of the long-term unemployment rate does not have an easy interpretation. Initially it could be interpreted as a better labour market situation, but if we consider the high levels of the other unemployment rates, unemployment rate and youth unemployment rate, this results interpretation seems implausible. Both the creation of new jobs and the increase of employment do not explain it since the employment rate and the long-term unemployment rate are following the same trend. A high turnover is not possible either, since there is also a reduction in the proportion of fixed-term contracts. Therefore, the most likely reasons for this decrease are first, the discouragement of those who have spent long periods unemployed and who are finally resigned to not finding a job; and second, the return of migrants to their original countries, their migration to a third country, and the emigration of some Spaniards to other countries.

Regarding those unemployed who are discouraged, data from the second quarter of 2018 show that more than 40 percent of those economically inactive who can be considered potentially economically active were not working because of their difficulties in finding a job. From 623,100 individuals considered economically inactive but with possibilities of being economically active, 44 percent has given up on finding a job because of discouragement. By gender, women are the most discouraged (47.8 percent versus 37.6 percent for men).[40]

As seen before, since 2008 there has been an increase in the number of migrants leaving Spain (288,432 in 2008 and 532,303 in 2013).[41] Furthermore, there has also been an increase in the number of Spaniards who emigrated to other countries from 35,372 in 2009 to 86,112 in 2016.[42] As a result, the labour force has been reduced, and therefore, the unemployment rates and the long-term unemployment rates have also decreased, since the potential individuals who can be unemployed have reduced.

There are two other labour market indicators that deserve particular attention in the Spanish case: part-time employment and fixed-term contracts. Despite the increase in the proportion of part-time employment, from 7.2 percent in 1995 to 14.5 percent in 2018, Spain was still in 2018 one of the European countries with the lowest proportion of workers in part-time employment (24.6 percent in the United Kingdom, 18 percent in France, 18.4 percent in Italy and Sweden, and 22.5 percent in the EU-15) (Table 2.1). Women were the most affected by this type of employment in all the countries considered (Tables 2.3). Probably the few part-time employment opportunities and the difficulty in combining family duties and a professional career are the main obstacles for Spanish women to completely enter the labour market, which would help to explain the relatively low

[40] https://www.ine.es/jaxiT3/Tabla.htm?t=4164&L=1
[41] These data only include individuals with foreign nationality (not Spanish). Source: http://www.ine.es.
[42] http://www.ine.es.

female participation rate (68.6 percent in 2018) (Pérez Infante, 2008, pp. 40-41). This is not surprising, since women are the ones who are still carrying out most of the family and home duties, such as, child care, elder care or housework.

Cultural and social history, which has traditionally placed the burden of family and home duties on women's shoulders, is one of the main causes. In Spain, women spend almost twice as much time as men doing unpaid work (housework, child and adult care, etc.) and according to the latest data available from the OECD, Spain is still one of the countries where women do most of the unpaid work.[43]

Although having a part-time job in Spain is not always a possibility, it does not seem to be the solution either. First, more than half of the workers securing a part-time job in Spain declare their desire to have a full-time job. In 2017 in Spain 61.1 percent of the workers declared they were working in a part-time job because they could not find a full-time one (68.5 percent for men, and 58.2 percent for women). In the same year, for EU-15 the proportion was 23.2 percent for women and 36 percent for men (26.2 percent for all the population).[44]

Second, the minimum wage in Spain is relatively low, therefore, working part-time reduces the monthly wage, making it difficult to make ends meet. The average monthly earnings for Spanish part-time workers in 2014 in the industry, construction and services sectors were 823 Euros, compared to 1,285 Euros in the Euro Area-18 (EA-18), for full-time workers it was 2,102 Euros (2,832 Euros in the EA-18). For females, the full-time monthly earnings in Spain were 1,943 Euros (2,433 Euros in the EA-18), and 822 Euros for part-time workers. And for males, 2,213 Euros for full-time workers in Spain (3,072 in the EA-18), and 827 Euros for part-time workers (1,238 in the EA-18).[45]

Another good indicator to account for the labour market conditions is the proportion of fixed-term contracts. Traditionally, the Spanish labour market has had a very rigid legislation (Pérez Infante, 2008; Polavieja, 2005). In the 80s the government tried to make it more flexible, aiming at a labour market which could

---

[43] https://stats.oecd.org/Index.aspx?datasetcode=TIME_USE#.
[44] http:// http://ec.europa.eu/eurostat/data/database.
[45] Ibid.

easily adjust to the economic changes, increasing the number of contracts when there was an economic expansion and decreasing the number of contracts when there was an economic retraction (Toharia, 1998). That was the main objective of the labour legislation approved in 1984 (Law 32/1984, 2nd August).[46] This new legislation seems to have worked during the recent economic crisis, since 90 percent of the employment losses in Spain were of workers securing a fixed-term contract (Vaughan-Whitehead, 2011).

In spite of this, the strict conditions of the permanent contracts, with better conditions for the employee as well as a higher financial compensation in case of dismissal (eight days per year worked in the case of fixed-term contracts and 45 days per year worked in the case of permanent contracts) and the more permissive labour legislation contributed to the expansion of fixed-term contracts as an instrument to reduce the labour costs for the employer (Pérez Infante, 2008), rather than to make the labour market more flexible before economic fluctuations (Recio, 1994). In 2012, while the Conservative Party was in power, they introduced a new labour reform (*Real Decreto-ley 3/2012, de 10 de febrero, de 2012* – Royal Decree-Law 3/2012 of 10th February 2012). The objective was to reduce the difference between permanent and fixed-term contracts, and especially, among their financial compensations. However, instead of improving the conditions for the fixed-term contracts, this new legislation worsened the conditions of the permanent contracts, reducing their financial compensation in case of dismissal (33 days per year worked instead of 45).

In fact, in some cases, fixed-term contracts are used as an entrance door to a company in the Spanish labour market. The law allows the employers to use this sort of contract when they hire a new worker for the first time. And in most cases, given the difference in the financial compensation in the case of dismissal, the fixed-term contract is renewed as many times as possible within a maximum of three years, with the possibility to extend it for one more year when the collective sector agreement allows it. After this period, the contract should become

[46] The complete text of the law can be found in http://www.boe.es/diario_boe/txt.php ?id=BOE-A-1984-17436.

permanent or the employee should be dismissed.[47] However, in the Spanish cases, most of these fixed-term contracts are a dead-end road and not a first step to finally getting a permanent contract.

Consequently, in the Spanish labour market, even though the proportion of fixed-term contracts has been reduced during the last decades (from 35 percent in 1995 to 26.9 percent in 2018), it is still higher than in other European countries (15.9 percent in Sweden, 16.7 percent in France, 17.1 percent in Italy, and 5.4 percent in the United Kingdom in 2018) (Table 2.1).

The literature has shown that fixed-term contracts have negative consequences for both, worker and society. For the worker, holding a fixed-term contract means a lower labour stability, and therefore, a higher difficulty to commit oneself with the job, to plan one's life or to carry on with personal and professional projects. For society in general, and for economy in particular, higher proportions of fixed-term contracts mean high job turnover, low on-the-job training investment, and therefore a waste of time and resources for the companies, and a loss of productivity for the economy of the country (Dolado, Jansen and Jimeno, 2002).

Furthermore, higher rates of fixed-term contracts do not necessarily have to be considered as negative. All depends on how they are used by the employers. In fact, fixed-term contracts are used in different ways amongst the countries. For instance, in the United Kingdom it usually acts as a step to get a permanent job, whereas in other European countries, as for instance France or Spain, as it has already been stated, fixed-term contracts can become the unique sort of contract that individuals can get (Cabrales, Dolado and Mora, 2014).

Using data from the Spanish Labour Force Survey, García-Pérez and Muñoz Bullón (2007) analysed the same young individuals (16-25 years old) from 1996 to 2003 and found that less than 10 percent of the fixed-term contracts became permanent after the first contract (9.7 percent for workers with high qualifications, 7.8 percent for medium qualifications, and 6.5 percent for low qualifications). This

---

[47] Article 15.1 of the Real Decreto Legislativo 2/2015 de 23 de Octubre (Spanish Royal Legislative Decre of 23th October 2/2015). https://www.boe.es/buscar/act.php?id=BOE-A-2015-11430&tn=1&p=20180704#a15.

proportion increases after several fixed-term contracts (until 17.9 percent for high educated workers, 15.3 percent for medium qualifications, and 10.7 percent for low qualifications) (García-Pérez and Muñoz Bullón, 2007, pp. 23-25). Besides, they stated that the labour market reforms of the 1990s have intensified the transitions from fixed-term contracts to permanent ones, since the likelihood of securing a permanent contract increased after 1997 (García-Pérez and Muñoz Bullón, 2007).

As regards gender differences, Tables 2.2 and 2.3 display the main labour market indicators separately for men and women, respectively. Tables show that in all European countries women have worse labour market indicators than men, especially when employment and unemployment rates are considered. Nevertheless, the participation of women has improved during the last two decades in all the European countries considered in the table, and especially in Spain. For instance, considering the Spanish activity rate for all the population, it increases by 12.8 percentage points (pp) from 1995 (60.9 percent) to 2018 (73.7 percent) (Table 2.1). In contrast, Men's participation rates have only increased by 3 pp, from 75.8 percent in 1995 and 78.8 percent in 2018 (Table 2.2); whereas women's activity rates have increased 22.5 pp (from 46.1 percent in 1995 to 68.6 percent in 2018) (Table 2.3). Therefore, the gender gap has been reduced by almost 20 points (from 29.7pp in 1995 to 10.2 pp in 2018). Women's activity rate increase has been basically caused by the rise in the proportion of women holding higher education levels and in their higher likelihood to enter the labour market (Garrido Medina, 2004; Toharia, 2008). In 1995 in Spain, 10.7 percent of the women labour force had a university degree, in 2018 women represented 20 percent (22 percent and 38 percent, respectively when only employed women are considered).[48]

The increase in the women's employment rate –by 25.1 percentage points (pp) from 1995 to 2018– and the decrease in the women's unemployment rate –by 13.7 pp from 1995 to 2018– are explained by the growing demand for labour, but

[48] http://www.ine.es.

it is also due to the higher proportion of part-time contracts for women –in 1995 Spain, 16.3 percent of the women's labour contracts were part-time, and in 2018, this proportion rose to 23.9 percent.

In conclusion, data show an improvement of the educational level of the labour force and of the main indicators of the Spanish labour market, although these latter are still worse than in other European countries. The main problems of the Spanish labour market are still unemployment, precariousness and the difficulty to balance work and family life.[49] These problems are especially relevant for women who are still coping with the main part of the housework. Whether this is despite the successive legal reforms, or because of them, is debatable.

2.4. Changes on the Spanish labour market over time

Focusing on the Spanish labour market, this section analyses data from the Spanish Labour Force Survey from 1995 to 2018 to sort out the supply of and the demand for workers regarding their educational level. Before basing the discussion on the university graduates in the following chapters, this section first discusses trends in the educational level of the Spanish population, to focus then on working out to what extent the occupational distribution of workers by educational level has suffered some changes during this period. Data are depicted separately by gender.

2.4.1. Educational level and changes over time

This subsection shows the upswing on the formal education level of the Spanish labour force due to the generational change. The retirement of older generations less formally educated and the entrance of younger generations more formally educated (basically due to the generalisation of compulsory education) have

[49] In fact, precariousness has become one of the main milestone for the Spanish labour market, especially after the global economic crisis which began in 2007.

changed the distribution of the population by education level, especially when it is analysed by age-group and gender.

This section aims to present the educational structure of the Spanish labour market and its more recent trends. The rest of the chapters focuses on university graduates, therefore, the objective of this chapter in general, and this section in particular, is to explain why it is important for the Spanish case to concentrate on the study of university graduates. That is, how the growth of the university graduates has modified the educational structure of the Spanish labour market.

Using data from the Spanish Labour Force Survey (SLFS) from 1995 to 2018, this section analyses how the changes in the education legislation and in the demand for higher educational levels helped the younger generation of the labour force to increase their educational levels. SLFS is a quarterly household survey carried out by the National Statistics Institute of Spain (Instituto Nacional de Estadística, INE), where interviewees are a stratified sample of the total Spanish population with about 150,000 observations for each quarter. Every quarter, one sixth of the sample is changed, so every interviewee answers the survey in six different quarters, although they cannot be identified in the general data.[50]

The higher education level of younger generations entering the labour market was propelled by three main factors: the enlargement of the supply of university studies from the 1980s (University Reform Act –LRU, 1983–); the increase in the population's living standards which facilitated the entrance of the middle class to the university system (demand side); and the increase in the age for compulsory education to secondary level from 14 to 16 years old in the 1990s. Besides, the retirement of the older generations with lower education levels also helped to reduce the absolute number and the proportion of individuals with compulsory education or even lower levels of education.

The period of analysis in this section is from 1995 to 2018. 2018 is the last year with available data from the SLFS. The analysis starts in 1995 to see the effect

[50] There is the possibility to follow households during four periods when longitudinal data from the SLFS is used, but this is not the case here. The sample used here is based on the cross-section description of the data.

of the different education acts (General Education System Organic Law –LOGSE, 1990– and University Reform Act –LRU, 1983–) on the supply of higher educated labour force. Besides, this is the same period which will be used in Chapter Three to estimate the ability of the Spanish labour market to accommodate the expansion of the university graduates. Using the same data in both chapters will facilitate the comparison between the results and therefore their coherence.

The Spanish educational system has been quite steady during all the 20th century and beginning of the 21st century. Furthermore, these modifications prevent classifying these together over the entire period. Trying to maintain the distinction between vocational and academic education as well as the best classification to compare it with the labour market, the final and general grouping that will be used for the educational level considers four categories: compulsory, vocational, academic and university (Table 2.4).

As regards university graduates, a short note is required before the analysis of the data starts. As can be seen in Table 2.4, the university group includes short degrees (three years), degrees (five years), masters, and doctorate. Until the homogenisation of the European University System, which started with the university reforms of 1992 and continued with the application of the Bologna Process, the Spanish university system was divided into three different levels. (a) A three-year degree or a short-cycle degree (‘Diplomatura’) usually included more vocational studies. This could be followed by people with academic secondary education, as well as, by people with higher vocational education. (b) A five-year degree or a long-cycle degree (‘Licenciatura’) used to last between four and six years, although five years were the most common in length. Only people with academic secondary education were able to enrol on a long-cycle degree. But they could enrol in this educational level immediately after the academic secondary education, or after finishing a short-cycle degree. (c) Postgraduates and doctorates.[51]

[51] Until 2000 data include only people with a doctorate. Since 2000 people with more than one university degree are divided between those who hold a postgraduate and

Table 2. 4. Classification of educational levels groups

Own groups	National Statistics Institute (INE) groups
Compulsory or less	Illiterates, without education, primary education and compulsory education (until 14 years old until 1990 and until 16 years old afterwards)
Vocational	Vocational medium and high post-compulsory education
Academic	Academic post-compulsory education (before university)
University	All university graduates (short degree (3 years), degree (5 years), masters, and doctorate)

Source: INE (http://www.ine.es)

Using the classification of the previous table, Figure 2.4 shows the upswing in the educational levels of the Spanish workforce. The labour force survey includes all individuals aged between 16 and 64 years old, however the sample used here has been restricted to the 25-64-year-old Spaniards. Data only include individuals with Spanish nationality. At the end of the 90s there was a flood of immigration, most of it from non-European countries, with substantially different educational backgrounds, who were employed in the Spanish labour market sometimes in jobs not related to their formal education (Albert, 2008). In 1998, 1.6 percent of the population in Spain was from other countries. In 2003, the proportion of migrants in Spain rose to 6.2 percent of the total population, composed of 38.7 South Americans, 22.1 percent from the EU, 19.6 percent from Africa, 13.1 percent from the rest of the Europe, 4.8 percent from Asia, 1.6 percent from the North America, and 0.1 percent from the rest of the world.[52]

The lower threshold is 25 years old because this is the age at which around 70 percent of the university students in Spain finish their degree. Taking into account that in the academic year 2017-2018 about 70 percent of the students enrolled at university for the first time were under 25 years old (57 percent were even under 22 years old), it seems reasonable to restrict the sample to the 25-64-year-old

those with a doctorate. For the pre-Bologna Spanish university system see Mora et al (2000).

[52] http://www.ine.es/ss/Satellite?L=&c=INECifrasINE_C&cid=1259925137681&p=1254735116567&pagename=ProductosYServicios%2FPYSLayout.

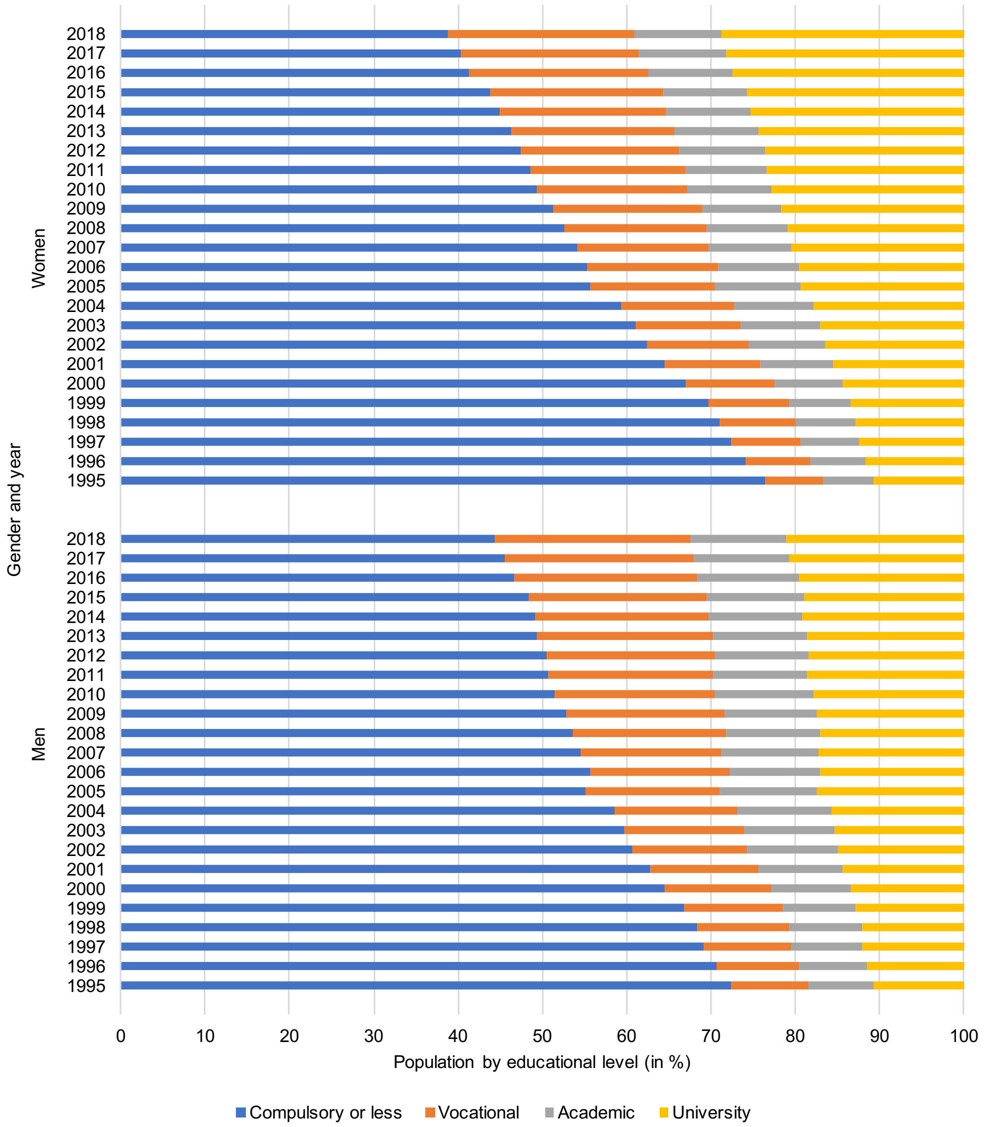
population. It is likely that those who finished the degree when they were older than 25 (30 percent), also started the degree later, and probably after they had spent some time in the labour market. Actually, it is also likely that they have been working while studying.[53] Besides, this low threshold of 25 years old is also used in the literature (Alba-Ramírez, 1993; Albert, 2008; García-Montalvo, 2001).

In the last years of the 20th century and the beginning of the 21st, the higher living standards of Spaniards along with the growth of the supply of post-compulsory education (fields of study but also number of centres offering these grades) encouraged the individuals' participation in the education system. The result has been the improvement of the educational level of the workforce. This education upgrade has been important for both sexes, but, it has been much greater for women (Fig. 2.4).

As Figure 2.4 shows, the most important increase in post-compulsory education (Vocational, Academic and University categories) has occurred in the labour force with university studies. This group has grown by 14 percentage points (pp), which means an increase of more than 1.2 million people in absolute terms. This difference is much higher for women with an increase in their representation at university level of more than 18 pp (from 10.7 percent in 1995 to 28.7 percent in 2018). As for men, only 21 percent were university graduates in 2018. Therefore, the generational change in the composition of the labour force market by educational level has had a more significant impact for women.

---

[53] University statistics from the *Ministerio de Ciencia, Innovación y Universidades* (http://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/universitaria/estadisticas.html).



Note: Only individuals with Spanish nationality.
Source: SLFS (2nd quarters), INE

Figure 2. 4. 25-64-year-old individuals by educational level and gender. Spain, 1995 and 2018

The displacement of less formally educated individuals by the new more formally educated generations becomes clearer if we focus on the age-group changes over time. The cohort effect is noticeable in 1995 as well as in 2018. For both men (Table 2.5) and women (Table 2.6), there is a reduction in the proportion of individuals with compulsory education at the most and an increase in the proportion of individuals with higher educational levels.

Both tables (2.5 and 2.6) reflect the two main changes previously described. First, proportions of individuals with compulsory education at the most is higher in older generations; second, for each age group, there is a reduction in the proportion of individuals holding only compulsory or less education levels, especially among the younger groups. Therefore, there is a generation effect, but there is also a time effect. In both cases, the education legislation and the supply of and demand for post-compulsory education are the main explanatory factors. On the one hand, legislation changes in the 1980-90s enabled the creation of new universities and new degrees. The number of private and state universities in Spain was 32 in 1978 and 70 in 2000 (Reques Velasco, 2007). On the other hand, there was an economic growth tendency increasing the population's living standards and families' expectations. All these factors together propelled both sides of the educational market, the supply of and the demand for post-compulsory education, and especially, for tertiary education.

Nevertheless, the proportion of youth holding only compulsory or less education levels is still significant. In 2018 Spain, 31.7 percent of 25-29-year-old men achieved compulsory or lower education levels. For 25-29-year-old women this proportion was only 19.6 percent. One should bear in mind that here only Spaniards are considered. Thus, there does not seem to be a problem of supply of education since other youth had the same opportunities and were able to finish post-compulsory education. On the contrary, it seems to be a problem of early-school-leavers. It is worth remembering that Spain is still one of the European countries with a high proportion of early leavers from education and training. In 2018 Spain, 18 percent of the total population of the same age group left their education, more than twice the proportion of Sweden (7.5 percent), twice the proportion of France (8.9 percent), and between 7 and 11 percentage points that of the UK (10.7 percent) and Italy (14.5 percent).[54]

[54] http://ec.europa.eu/eurostat/data/database.

Table 2. 5. 25-64-year-old men by age group and education level. Spain, 1995 and 2018 (in percentages)

		Compulsory or less	Vocational	Academic	University	Total
1995	25-29	53.8	18.9	11.8	15.5	**100.0**
	30-34	61.5	15.0	11.0	12.5	**100.0**
	35-39	64.5	10.3	12.2	13.0	**100.0**
	40-44	71.4	8.3	8.9	11.5	**100.0**
	45-49	77.8	5.9	5.5	10.7	**100.0**
	50-54	83.5	4.4	3.6	8.5	**100.0**
	55-59	87.6	3.3	2.7	6.4	**100.0**
	60-64	90.5	2.4	2.2	4.9	**100.0**
	Total	**72.3**	**9.2**	**7.7**	**10.8**	**100.0**
2018	25-29	31.7	31.5	10.2	26.7	**100.0**
	30-34	39.6	30.8	6.7	22.9	**100.0**
	35-39	37.8	28.9	10.2	23.1	**100.0**
	40-44	36.5	26.0	11.4	26.1	**100.0**
	45-49	42.9	24.2	11.0	21.8	**100.0**
	50-54	48.1	21.8	11.3	18.8	**100.0**
	55-59	52.4	17.4	13.8	16.3	**100.0**
	60-64	57.1	12.6	13.9	16.3	**100,0**
	Total	**44.4**	**23.1**	**11.4**	**21.1**	**100.0**

Note: Only individuals with Spanish nationality
Source: SLFS (2nd quarters), INE

Table 2.6 focuses on women and displays slight differences from the men's table (Table 2.5). First, the lower proportion of women holding compulsory education at the most among the younger women (12.1 percentage points (pp) less than for men), and second, the higher proportion of university graduate women than men, also among the younger women (15 pp higher, 41.7 for women and 26.7 percent for men). In both cases, the data show that fewer women leave early from education and training compared with men. In 2018, 21.7 percent of 18-24-year-old men left education or training in Spain, but only 14 percent of women.[55]

[55] http://ec.europa.eu/eurostat/data/database

Table 2. 6. 25-64-year-old women by age group and educational level. Spain, 1995 and 2018 (in percentages)

		Compulsory or less	Vocational	Academic	University	Total
1995	25-29	50.1	17.8	10.8	21.3	**100.0**
	30-34	58.6	14.3	10.6	16.5	**100.0**
	35-39	68.6	7.9	9.0	14.5	**100.0**
	40-44	78.7	5.0	5.6	10.6	**100.0**
	45-49	86.2	3.1	3.6	7.2	**100.0**
	50-54	91.0	1.6	2.2	5.1	**100.0**
	55-59	94.7	0.0	1.6	3.7	**100.0**
	60-64	96.3	0.0	1.2	2.4	**100.0**
	Total	**76.6**	**6.8**	**5.9**	**10.7**	**100.0**
2018	25-29	19.6	29.3	9.5	41.7	**100.0**
	30-34	25.5	27.6	7.4	39.5	**100.0**
	35-39	25.1	27.9	9.0	38.0	**100.0**
	40-44	25.9	26.9	9.7	37.4	**100.0**
	45-49	36.8	24.0	10.7	28.5	**100.0**
	50-54	45.1	20.6	11.3	22.9	**100.0**
	55-59	51.3	17.5	11.5	19.8	**100.0**
	60-64	62.5	10.4	11.0	16.1	**100.0**
	Total	**38.8**	**22.2**	**10.3**	**28.8**	**100.0**

Note: Only individuals with Spanish nationality
Source: SLFS (2nd quarters), INE

This subsection has clearly described the improvement in the educational level of the Spanish population and its causes. This upgrade has been especially important for women in general, and for women with higher education levels in particular, who have increased the absolute and the relative participation of women in the Spanish labour market. However, these data do not account for the occupations they are holding, and therefore, do not clarify whether gender segregation has been reduced. It does display the trend of the youngest women's participation in the university system. Indeed, as can be seen in the comparison of the data of Tables 2.5 and 2.6, in 2018 and for the younger generations, the proportion of university-educated women out of all women in the age range 25 to 64 is higher than the percentage of university-educated men out of all men in this same age range (around 20 percentage points).

Holding higher educational levels also affects other spheres of women's life, and especially those related to raising a family. Empirical literature on demography found that higher educated women got married later, usually because they postponed their marriage until they finished their degree; and they also have a different behaviour pattern regarding children. They have their first child later, but they also have the second child sooner, reducing the time they spent out of the labour market (Gutiérrez Domènech, 2008).

An increase in the population's knowledge is important for the individual as well as for the general development of the country. For an individual, a knowledge upgrade is important not only from an economic perspective (regarding the labour market returns), but also in a more general perspective. It will be easier for a person to deal with other people or even with administration if they are literate and are able to search, understand and interpret information. While studying in the education system, especially in tertiary education, individuals learn some specific skills, but also some general skills, such as, how to cope with new concepts and procedures. Education provides individuals with knowhow which will be very useful in their future life to solve possible problems or complicated situations. But, when we focus on the economic gains, the literature has broadly demonstrated the positive relationship between educational level and economic returns (direct and indirect, pecuniary and not pecuniary), especially for those with tertiary education.[56]

The education upgrading of the population also affects the general economy of the country. History has demonstrated that countries with a more highly educated population have had higher economic growth (Goldin and Katz, 2008; Hanushek and Kimko, 2000; Hanushek and Woessmann, 2007; Hanushek and Woessmann, 2015). It has been possible due to their knowhow to introduce new technologies and even to create them. Besides, as more educated workers usually receive higher wages, they also pay higher taxes (direct and even indirect, because they earn and spend more money). Therefore, they contribute to the increase of the public budget and the financial standing of the nation. Another

[56] In the 1990s numerous studies measuring the returns of education for individuals and for society appeared. Focusing on Spain, there is, for instance, the work of Alba-Ramírez and San Segundo (1995); Barceinas *et al.* (2000); San Segundo (1997); Vila and Mora (1996); Vila and Mora (1998).

matter is how the state distributes this public budget, whether it is in welfare or in defence or elsewhere, but this exceeds the objective of this study.

2.4.2. Occupational levels and changes over time

In this subsection the distribution of occupations over time is analysed to deepen the description of the Spanish labour market. The previous subsection has focused on the characteristics of the supply of workers. This subsection sheds some light on the demand for these workers to outline whether there has been a structural change in the distribution of jobs by occupational category. As in the previous subsection, data has been analysed by gender.

Using the same data as in the previous subsection, SLFS from 1995 to 2018, and in order to obtain a better indicator of the outputs of the Spanish education system, I still consider in this subsection only the workers with Spanish nationality. Despite the fact that excluding non-Spanish people's data is less representative of the real working age population, only Spanish people have been included because this gives a better indicator of the outputs of the Spanish education system (Albert, Toharia and Davia, 2008).

In the original data of the SLFS, occupations are classified using the two-digit break-up of the National Classification of Occupations (CNO-1994 for 1995-2010 data, and CNO-2011 for 2011-2018 data). These are national classifications valid only for the Spanish labour market, but their categories have a correspondence in the 2008 International Standard Classification of Occupations (ISCO-08).[57] To facilitate interpretation, I have used the main groups of the International Classification for Occupations ISCO-08 instead of the National Classification used by the survey (CNO-94 and CNO-2011).[58] I omit the population from the

[57] ISCO is the International Standard Classification of Occupations used by the International Labour Organization (ILO): http://www.ilo.org/public/english/bureau/stat/isco/index.htm.

[58] I have followed INE's equivalence table which shows the correspondence between ISCO and CNO-94 classifications. This table can be found in http://www.ine.es/clasifi/cno94_ciuo88.htm.

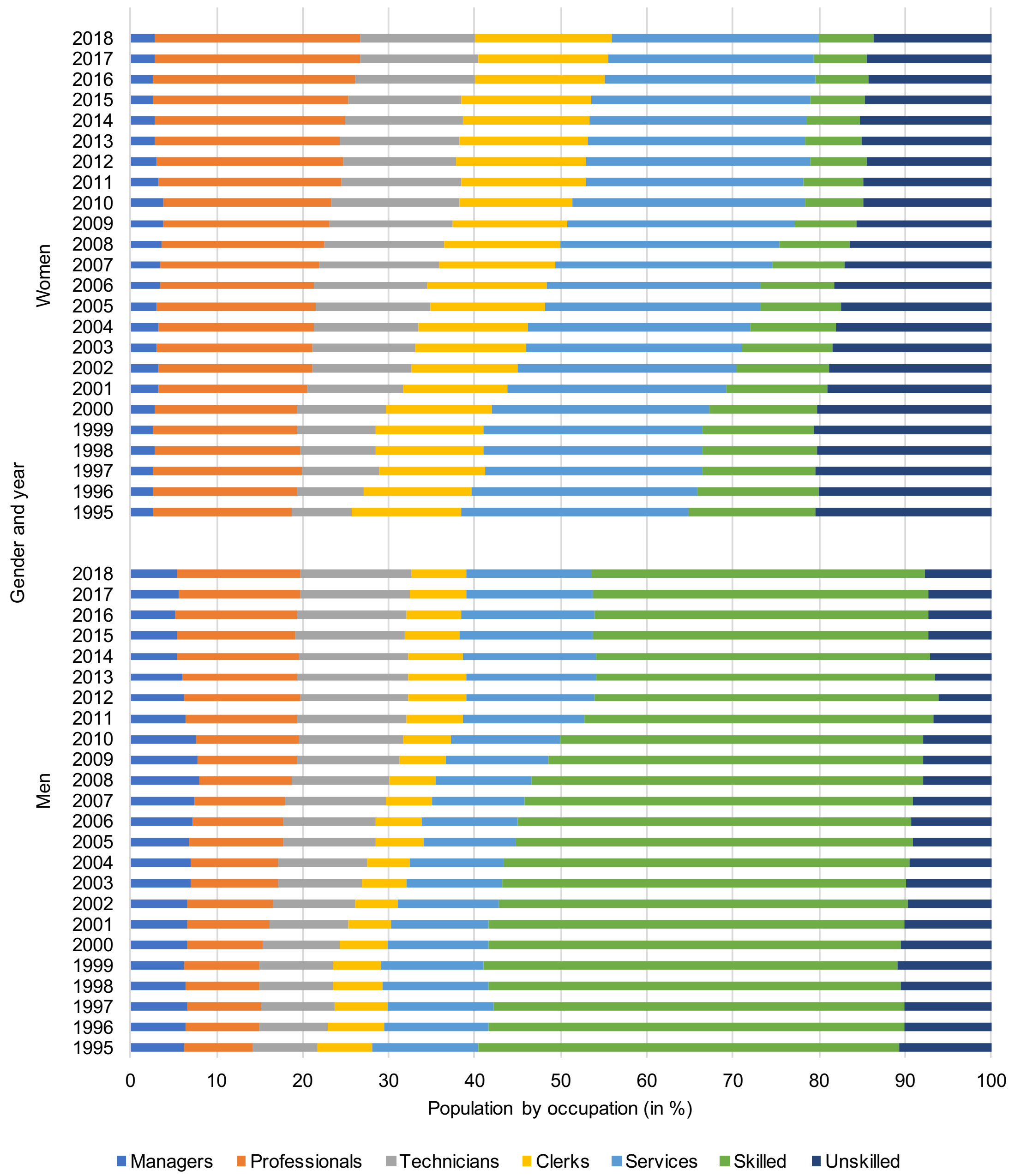
army, since they are not part of the labour market, but their exclusion does not affect the final results since they only represent 0.31 percent of the total sample (6,509 observations for all the period from 1995 to 2018, less than 300 observations per year in a set of data of around 150.000 observations per year). Table 2.7 describes the occupational groups used in this chapter and their equivalence to the international classification (ISCO-08).

Table 2. 7. Description of the classification of professional groups

Groups	Major groups of the ISCO 2008
(Omitted)	Group 0 – Army
Managers	*Group 1* – Legislators, senior officials and managers
Professionals	*Group 2* – Professionals
Technicians	*Group 3* – Technicians and associate professionals
Clerks	*Group 4* – Clerks
Services	*Group 5* – Service workers and shop and market sales workers
Skilled	*Group 6* – Skilled agricultural and fishery workers *Group 7* – Craft and related trade workers *Group 8* – Plant and Machine operators and assemblers
Unskilled	Group 9 – Elementary occupations

Source: ISCO-08

Figure 2.5 summarises the main trends of the demand for workers in the Spanish labour market from 1995 to 2018. Data are displayed by gender making the gender comparison easier. The gender disparity on the dimension of each professional group is relevant, especially in two occupational groups. By difference the most important occupational group for men is ‘Skilled’, whereas for women it is ‘Services’ and, in the last years, ‘Professionals’ is also growing. Therefore, ‘Skilled’ for men and ‘Services’ for women could be defined as the two Spanish occupational sectors where the workers with less formal education are concentrated. This result suggests a high level of job segmentation in the Spanish labour market, at least for those jobs requiring lower education levels.



Note: Only individuals with Spanish nationality.
Source: SLFS (2nd quarters), INE

Figure 2. 5. 25-64-year-old employees, by professional group and gender. Spain, 1995-2018.

'Skilled' group includes all qualified workers in the primary, secondary and building sectors. All these jobs are generally male dominated, especially in the building sector, which explains why almost 50 percent of men's jobs are concentrated in this occupational group in all years, whereas only 15 percent of women were working in this occupational group in 1995, and less than 10 percent in 2018 (Fig. 2.5).

'Services' usually concentrates a higher number of women because it includes those jobs related to personal services such as home-based personal services, hairdressers, tour guides and similar, which are highly feminised.

Indeed, the dissimilarity index of segregation (Duncan and Duncan, 1955) is higher than 0.5500, indicating that more than half of 25-64-year-old Spanish women should change their occupational category to achieve a gender-proportional distribution in the Spanish labour market. Even so, a reduction in the occupation segregation in the Spanish labour market has taken place in the last years. The Duncan dissimilarity index of segregation has been reduced from 0.6817 in 1995 to 0.6165 in 2005, and to 0.5583 in 2018. Nevertheless, the gender concentration in the Spanish labour market in some occupational groups persists.

The services sector is the principal sector of the Spanish economy, followed by the industry, building sector, and agriculture. In 2000, the services sector accounted for 65.8 percent of the gross domestic product (GDP) and 66 percent in 2018.[59] Furthermore, in the second quarter of 2018, 75.4 percent of the employed were working in the service sector (64.4 percent when men are considered, and 88.5 for women).[60]

As expected, a higher proportion of men than women are employed as 'Managers', although they represent less than 7 percent of the total men workers (3 percent in the case of women) (Fig. 2.5). On the contrary, a great proportion of women are working as 'Professionals' (almost 20 percent, 9 pp more than men).

Women with higher education levels are more willing to participate in the labour market than those who have not studied, since the opportunity cost of being out

[59] https://www.fbbva.es/wp-content/uploads/2003/09/DE_2003_LPrados_El_progreso_economico.pdf
[60] http://www.ine.es

of the labour market is higher for those who hold higher educational levels, especially for women. The reason is that women are still carrying out most of the family and home duties, both, housework and child care or elder care.

The proportion of workers in the ‘Technicians’ and ‘Clerks’ group is also higher for women than for men (2 percentage points (pp) on average for all the years for ‘Technicians’ and almost 8 pp for ‘Clerks’). This confirms the previous affirmation. Women holding higher education studies had higher opportunity costs of being out of the labour market, and therefore, they are more likely to enter the labour market than women with lower educational levels. The distribution of education levels in the female population showed that, on average, 19.7 percent of women hold higher education degrees (Fig. 2.4). When we reduced the sample to those who are employed, the proportion rose up to 27.7 percent. In contrast, the proportion of women with compulsory education at the most reduces by 11 percentage points.[61]

To sum up, the previous figure suggests the occupational gender segregation and the education polarisation in the Spanish labour market. However, it does not clarify whether job tasks suit the skills and abilities of the workers.

Table 2.8 and 2.9 present the different gender patterns in the occupational distribution by age groups. As in the previous figure, in the case of men (Table 2.8), the professions grouped as ‘Skilled’ (Table 2.7) are the largest, in all age groups and in both years under consideration (1995 and 2018). This demonstrates that in this case, differently from the education level considered before, there is not an important generation or time effect.

Between 1995 and 2018, the biggest change in male occupation distribution has been in ‘Technicians’ and ‘Professionals’. In all age groups the proportion of ‘Technicians’ has at least doubled, while the ‘Professionals’ have increased but to a lesser extent. Therefore, it appears that the changes have followed a similar pattern in all age groups, maintaining the initial differences between cohorts.

[61] The latter data is not shown here, but it could be calculated easily using the SLFS.

Table 2. 8. 25-64-year-old men by age group and occupation. Spain, 1995 and 2018 (in percentages)

		1	2	3	4	5	6	7	Total
1995	25-29	2.4	7.2	8.1	6.6	15.6	44.7	15.3	**100.0**
	30-34	4.8	8.5	8.8	6.9	14.5	44.7	11.8	**100.0**
	35-39	6.5	9.3	8.3	7.0	13.1	46.9	9.0	**100.0**
	40-44	7.4	9.1	7.3	6.9	11.8	48.7	8.9	**100.0**
	45-49	7.4	8.2	7.1	7.1	9.8	51.5	8.9	**100.0**
	50-54	7.4	7.0	7.4	5.5	11.0	52.0	9.7	**100.0**
	55-59	7.3	5.3	5.9	4.8	10.7	54.9	11.2	**100.0**
	60-64	8.4	7.1	4.8	4.0	10.6	53.2	11.9	**100.0**
	Total	**6.2**	**8.0**	**7.5**	**6.4**	**12.4**	**48.8**	**10.7**	**100.0**
2018	25-29	0.0	16.5	13.5	6.1	15.6	34.2	14.1	**100.0**
	30-34	0.0	14.4	13.8	5.1	15.0	41.8	9.9	**100.0**
	35-39	3.8	14.0	13.5	6.0	13.6	41.9	7.4	**100.0**
	40-44	6.0	15.5	13.4	6.7	14.1	37.3	7.0	**100.0**
	45-49	7.0	14.5	12.3	6.8	14.1	38.1	7.2	**100.0**
	50-54	5.6	13.0	13.2	6.7	14.7	39.5	7.3	**100.0**
	55-59	6.8	12.1	13.2	7.5	14.8	39.4	6.2	**100.0**
	60-64	7.6	16.0	10.5	5.5	15.7	37.9	6.7	**100.0**
	Total	**5.2**	**14.0**	**13.0**	**6.5**	**14.5**	**38.9**	**7.7**	**100.0**

Note: Only men with Spanish nationality between 25-64 years old are included. Both public and private sectors. Categories: following the ISCO88 (Army omitted)
1 – Managers 3 – Technicians 5 – Services 7 – Unskilled
2 – Professionals 4 – Clerks 6 – Skilled
Source: SLFS (2nd quarters), INE

The women's case (Table 2.9) is different from the men's in both perspectives, within time and age groups. There are two main shifts: (a) a reduction in the proportion of women employed in elementary occupations, especially in older generations, women more than 50 years old; and (b) an increase in the proportion of 'Technicians' and 'Services' through all the generations, but especially in the youngest ones, showing again an educational-demographic change.

The younger generations of women are more educated and therefore they are more able to work in more skilled jobs than the older generations. Table 2.9 demonstrates the progressive incorporation of younger generations of more educated Spanish women to the labour market, partially, because of the increase in their opportunity cost of being inactive (Alba-Ramírez and San Segundo, 1995; Vila and Mora, 1998).

Table 2. 9. 25-64-year-old women by age group and occupation. Spain, 1995 and 2018 (in percentages)

		1	2	3	4	5	6	7	Total
1995	25-29	0.0	15.7	8.6	20.2	26.3	11.8	17.4	**100.0**
	30-34	0.0	20.1	8.8	16.5	24.5	11.9	18.2	**100.0**
	35-39	2.8	19.9	8.0	13.9	24.9	12.9	17.6	**100.0**
	40-44	0.0	18.2	7.6	11.4	27.6	13.8	21.4	**100.0**
	45-49	0.0	14.9	5.9	10.1	30.4	15.5	23.2	**100.0**
	50-54	0.0	12.6	0.0	7.1	30.9	21.1	28.3	**100.0**
	55-59	0.0	10.6	0.0	0.0	35.2	24.8	29.4	**100.0**
	60-64	0.0	0.0	0.0	0.0	31.0	33.0	36.0	**100.0**
	Total	**0.5**	**16.3**	**6.4**	**12.7**	**27.6**	**15.3**	**21.3**	**100.0**
2018	25-29	0.0	28.3	13.4	14.0	31.9	4.7	7.7	**100.0**
	30-34	0.0	30.1	11.6	15.5	30.5	5.2	7.1	**100.0**
	35-39	2.8	27.8	12.5	17.9	26.3	5.3	7.4	**100.0**
	40-44	4.0	26.2	13.8	17.3	22.8	5.2	10.7	**100.0**
	45-49	3.4	20.8	12.4	18.8	22.8	6.5	15.3	**100.0**
	50-54	3.0	20.0	14.1	14.3	23.1	6.9	18.6	**100.0**
	55-59	2.4	21.5	14.2	14.5	19.9	7.0	20.5	**100.0**
	60-64	0.0	21.2	14.7	12.5	20.9	9.2	21.5	**100.0**
	Total	**2.4**	**24.1**	**13.4**	**16.0**	**24.1**	**6.2**	**13.8**	**100.0**

Note: Only women with Spanish nationality between 25-64 years old are included. Both public and private sectors. Categories: following the ISCO88 (Army omitted)
1 – Managers 3 – Technicians 5 – Services 7 – Unskilled
2 – Professionals 4 – Clerks 6 – Skilled
Source: SLFS (2nd quarters), INE

These data suggest a slightly increase in the proportion of higher skilled jobs in both men and women. In 1995, 14.5 percent of men and 22.7 percent of women were working in 'Technicians' and 'Professionals' jobs, whereas in 2018 these proportions were 27 percent for men and 37.5 percent for women. Therefore, in the case of women, this tendency to increase the proportion of higher skilled jobs is fairly evident, especially in the last ten years of the dataset. But, as it has already been pointed out, this could be the result of a selection bias in women's decision to enter the Spanish labour market.

For both genders and years, the highest proportion of workers is in the occupational category of 'Skilled', even if their proportion is fairly different: 48.8 percent of men and 27.6 of women in 1995; and 38.9 percent of men and 24.1 percent of women in 2018.

Altogether, there is a displacement of workers from the lower skilled jobs to those which required higher educational levels. Therefore, the occupational distribution has apparently followed the same pattern as the educational level. Now, it should be disentangled whether their intensity has been to the same degree, and the Spanish labour market has been able to balance their supply of and demand for workers according to their educational level.

2.5. Education match in the Spanish labour market

Another important point is whether there is a match between worker's education level and the tasks required by the workplace. The results of a first attempt to analyse whether in the Spanish labour market there has been a match between education level, as a proxy of worker's abilities, and their job are shown in Table 2.10. A deeper analysis of the ability of the Spanish labour market to accommodate the labour force will be carried out in the next chapters.

The aim of this section is to outline whether there has been an educational level balance in the Spanish labour market. In the previous sections of this chapter the analysis of the main characteristics of some European countries has stated the difficulties of the Spanish labour market to accommodate all of their labour force. As a consequence, the Spanish unemployment rates are higher than the ones for other European countries during this period.

The ability of a labour market to accommodate all their workers is important, but it is also fundamental for the labour market to be able to accommodate its labour force in the appropriate jobs.

The previous section showed the increase in the educational level of the population and the industrialisation of its productive system in Spain. However, the improvement of the educational level of the labour force would need some years to be fully reflected in the labour force supply. Similarly, changes in the demand for labour force due to industrialisation (new production processes, technology shifts, etc.) might have taken some years to develop. The next table

(Table 2.10) shows whether these two upward movements occurred at the same time.

Table 2. 10. Population by educational level, gender and occupation. Spain, 1995 and 2018 (in percentages)

	Compulsory or less		Vocational		Academic		University	
	1995	2018	1995	2018	1995	2018	1995	2018
Men								
Managers	4.4	3.0	5.6	3.6	12.9	7.5	11.7	10.5
Professionals	0.1	0.2	2.5	2.3	1.6	2.2	60.3	54.5
Technicians	3.9	6.0	14.3	19.6	21.8	18.8	11.8	14.4
Clerks	4.0	2.8	9.6	7.2	20.7	13.2	7.0	8.6
Services	13.7	16.1	11.4	14.6	16.7	26.0	3.4	6.4
Skilled	60.0	58.1	50.4	47.1	21.5	27.2	4.7	4.5
Unskilled	14.0	13.8	5.9	5.8	4.8	5.2	1.1	1.1
Matched	**87.7**	**88.1**	**85.7**	**88.4**	**20.7**	**13.2**	**60.3**	**54.5**
Mismatched	**12.3**	**11.9**	**14.3**	**11.6**	**79.3**	**86.8**	**39.7**	**45.5**
Women								
Managers	2.3	1.4	2.2	1.8	4.8	3.1	3.3	4.3
Professionals	0.1	0.2	4.0	3.1	2.7	1.7	70.2	59.9
Technicians	3.1	3.9	15.9	25.5	17.6	16.8	9.0	11.0
Clerks	7.1	5.2	26.8	23.8	38.3	27.2	11.4	15.4
Services	32.7	38.1	37.9	30.2	25.4	34.9	4.9	7.4
Skilled	22.7	14.2	6.2	5.8	5.1	6.1	0.7	0.9
Unskilled	32.1	37.0	8.0	9.9	6.2	10.2	0.6	1.0
Matched	**87.4**	**89.3**	**86.8**	**85.2**	**38.3**	**27.2**	**70.2**	**59.9**
Mismatched	**12.6**	**10.7**	**13.2**	**14.8**	**61.7**	**72.8**	**29.8**	**40.1**

Note: Only Spanish between 25-64 years old are included. Both sectors public and private. Categories: following the ISOC88 (army omitted)
In blue those workers whose educational level matched with their job tasks.
Source: SLFS (2nd quarters), INE

As shown in the literature review (Chapter One), the lack of a concrete definition as well as a lack of an adequate measurement of the educational level required by each job makes the task more difficult (Green, McIntosh and Vignoles, 1999). Neither individuals nor workplaces are homogeneous. A given education level does not ensure that individuals will have the same skills or abilities.

Real skills would depend on the innate (potential) personal abilities, on the kind or content of the education (grade, subject, etc.) received, and on the experience received during on-the-job training. Equally, different workplaces, even though they are classified in the same occupational category, can require (need) different skills/abilities to do the tasks. However, in this preliminary analysis, where the aim is a broad overview, I ignore these differences and assume they do not have an important effect on the match between the worker and the workplace.

Therefore, for this preliminary analysis I have considered the following classification: 'Unskilled' includes all the jobs that do not need any qualification to be done; 'Skilled' and 'Services' those which need compulsory education, although these two professional groups can need in some cases vocational studies; 'Clerks' includes jobs whose tasks need academic or vocational education; higher vocational education is required by 'Technicians'; and 'Professionals' includes the professions that need at least university-graduate workers. Following this classification, Table 2.10 shows the distribution of Spanish individuals in professional categories by gender, year, and educational level.[62]

The level of education match in the Spanish labour market was quite important both in 1995 and 2018 in all education levels, but especially in the academic one, i.e. post-compulsory education below university level. Data from Table 2.10 show that in 1995, 20.7 percent of men with academic education matched the appropriate job in the labour market, 13.2 percent in 2018. Match indicators are slightly higher for women holding academic education: 38.3 percent in 1995 and 27.2 percent in 2018. It is not surprising since academic education is theoretically a step to accessing higher education, and therefore, it is difficult to find appropriate jobs for individuals with this educational level. On the contrary, the level of matching for individuals with other education levels is higher than 85 percent for compulsory or less and for vocational studies, and lower than 70 percent for university-graduates.

Nevertheless, data show an upswing in the education mismatch between education level and profession between 1995 and 2018 for those men with academic and university studies, and those women holding vocational, academic

[62] The occupation categories follow the main groups of the International Standard Classification of Occupations (ISCO-08: http://www.ilo.org/public/english/bureau/stat/isco/isco08/index.htm).

and university studies. The total in the table shows the level of education match and mismatch for each educational level in the Spanish labour market in 1995 and 2018 by gender. The table shows that the mismatch increase was especially relevant for women whose education mismatch increased on average 7.7 percentage points (pp), but with a big dispersion among educational levels: more than 10 pp for those educational levels higher than academic and less than 2 pp for those individuals with compulsory or vocational education. In particular, the increase in the educational mismatch for women has been: 11.1 pp for academic education (from 61.7 to 72.8 percent), 10.3 pp for university education (from 29.8 to 40.1 percent), and 1.6 pp for vocational studies (from 13.2 to 14.8 percent). Only women holding compulsory or lower educational levels have reduced their levels of mismatch from 12.6 to 10.7 percent.

For men, the mean of the education imbalance growth was 6.7 pp. Men holding academic education increase their educational mismatch by 7.6 pp (from 79.3 to 86.8 percent), and those holding a university degree by 5.8 pp (from 39.7 to 45.5 percent). Men holding other educational levels, compulsory and vocational, have increased their educational balance in the Spanish labour market, and therefore they are generally employed in a job which does not require either higher or lower educational levels (Table 2.10).

Some other gender differences come up from the education match data of the table for both years, 1995 and 2018. When analysing the data by educational level, gender differences became more evident since the main gender differences are within each group and not in the total amount of match per education group.

Both in 1995 and 2018, most women with compulsory education (35 percent on average) were working in the services sector (‘Services’). In contrast, an average of 67 percent of the education-matched men with compulsory education is employed in ‘Skilled’ jobs, while for women the proportions are 26 percent in 1995 and 16 percent in 2018. These data reflect the gender segregation of the Spanish labour market, even in those jobs where lower education levels are required, where the ‘Skilled’ professions are mainly taken by men (of all the ‘Skilled’ jobs, 86 percent are taken by men, and only 51 percent of the ‘Unskilled’).[63]

In the case of vocational education, the gender distribution among the professional groups also follows the masculinisation and feminisation of either the education system and the demand of the labour market. In vocational studies, there is a significant gender segregation between the administrative studies and other mechanical or more technical studies. Traditionally vocational studies have been more attended by men. For instance, in the academic year of 2017-18 there were 51,078 men and 21,102 women. Therefore, even though the number of individuals following administrative studies are similar by gender (6,425 men and 6,525 women in the academic year 2017-18), the proportion that they represent out of the total by gender is higher for women (31 percent) than for men (12.6 percent). As a result, this gender imbalance is also reflected in the labour market. The principal job opportunities for administrative studies are in the ‘Clerks’ group, which explains the higher proportion of women working in this sector, whereas the other more technical vocational studies are more masculinised and their principal job opportunities are in the ‘Skilled’ group. The data of the previous table confirm this.

Academic education is intended to be a step to university studies. Therefore, those who hold academic education had only received general education and supposedly they had not acquired the necessary skills for any profession, appart from ‘Clerks’. They usually know how to do clerical tasks and therefore, it is the most suitable profession for them. However, data in Table 2.10 indicate that in 1995 the majority of women with academic studies were working in this occupational group (38.3 percent). In 2018 for women and in both years for men, most of the individuals with this education level were working in ‘Skilled’ jobs or in the ‘Service’ sector. This distribution of the proportions indicates that in the Spanish labour market sometimes holding a diploma in academic studies is equivalent to holding a diploma in vocational studies, even if higher specific

[63] These data are not shown in the table, but it can be calculated using the original data.

education is supposed to have been done in vocational studies. The main reason is probably the general idea that the education system (either vocational or academic) only gives general skills, and the experience gained from daily work on the job are what gives the specific skills required in each workplace. Therefore, an individual with academic studies has demonstrated s/he is able to learn, and would be a good candidate also for on-the-job training.

The group with fewer gender differences related to the education match is university graduates. In both cases, around 70 percent or even more of the graduates are matched. Nevertheless, a slightly difference can be found in the distribution of the mismatched ones. More than 10 percent of men were working as ‘Managers’, whereas a similar percentage of women were working as ‘Clerks’, therefore, it is likely that the latter were over-educated.

The ‘Managers’ group is likely to require university studies, since this sample only includes employees. Another important feature for this occupational group is the difference in the gender distribution. The male dominance of ‘Managers’ is clearly seen in the data (Table 2.10).

Therefore, according to this preliminary analysis, it appears that, especially in the last years, the structure by occupational groups of the Spanish labour market had some difficulty accommodating the continuous entrance of highly skilled workers. The result points out an increase in the education mismatches in the labour market and, therefore, a waste of individual and social resources. Further research on the probability for university graduates being over-educated are discussed in the following chapters.

2.6. Spanish labour market at a glance

The main goal of this chapter is to contextualise the Spanish labour market, in general, and the employability of the university-graduates in the Spanish labour market, in particular. Using different datasets, each one of the three main sections contributes to the main goal of the chapter.

After the introduction, the historical perspective is useful to put the reader in the Spanish context. In the third section of this chapter, the analysis of the indicators of the Spanish labour market in relation to other European Countries (the United Kingdom, France, Italy, and Sweden) set Spain in the context of some of its neighbouring countries. The analysis of the data shows that the gentle convergence between the Spanish and the European labour market's indicators has not been enough to fully solve some of the structural problems the Spanish labour market is facing. Higher unemployment rates, and especially, higher youth and female unemployment rate, or a higher proportion of fixed-term contracts are some of the features of the Spanish labour market. The discussion of data demonstrated that an increase in the supply of jobs, and above all of quality jobs (jobs with better conditions, higher tenure contracts, and higher wages) is necessary to improve the Spanish labour market's indicators.

In the fourth section of this chapter, the main characteristics of the Spanish labour force in the last years of the 20th century and beginning of the 21st have been described. Using individual data from the Spanish Labour Force Survey of 1995 and 2018, I have homogenised and reclassified the data to make them comparable. The resulting figures have allowed me to identify the main characteristics of this labour force, their educational levels and their main occupations, and how these two variables have changed over time. Data have been analysed separately for men and women to account for the gender differences.

Three main changes have characterised the Spanish labour market in the last decades. First, an increase in the general educational level of the labour force caused by two main factors: an increase in the supply of and demand for compulsory and post-compulsory education; and, the replacement of the older generations by younger generations who have more formal education. On the one hand, the implementation of a new law, *General Education System Organic Law (LOGSE)* which made full-time education until 16 years old compulsory, and the creation of new universities and new degrees, which almost doubled from 1970s to 2000, propelled the supply of both compulsory and post-compulsory education, respectively. On the other hand, the higher expectations of the individuals regarding their living standards helped to stimulate the demand for education. This propelled this generational replacement.

Second, there have been gender differences in the movement of the occupational distribution over this period. Changes are more intense in the case of women. Data show an important reduction in the proportion of women in unskilled occupations amongst older women above 45-50 years old, and an increase in the proportion of technical occupations, at least amongst younger women under 40 years old. Traditionally, women have had higher levels of illiteracy and lower educational levels than men. As a consequence, their participation in the labour market has been less active and in unskilled occupations. However, women's participation in the labour market has been growing steadily and they are now employed in more skilled occupations due to their higher participation in the educational system and the improvement of their educational level.

Third, the occupations held by people with an academic diploma or university degree have markedly changed over this period. There has been a reduction in the proportion of workers with diplomas and degrees in clerical jobs ('Clerks') for men and women. This trend has been offset by an increase in the proportion of those with higher education working as 'Technicians' and 'Services'. Analysed by gender, there has been an increase in the proportion of men and women graduates working in 'Technicians' and 'Legislators' occupations. This new occupational distribution responds to the skill-biased technological change of the Spanish productive system which has increased the labour market demand for non-manual workers, and especially, for those more qualified ones.

All in all, this chapter has shed some light on the absolute and relative numbers of the Spanish labour market and its main disadvantages. First, the low ability to take advantage of the number of workers, skilled and unskilled, that are not able to find a job. And second, the low ability to create the adequate job for the adequate worker. In order to solve these matters, changes in the education as well as in the productive system are necessary.

In terms of education, new policies helping to reduce the proportion of early school leavers, especially among men, and promoting the vocational training are needed. Encouraging the current vocational training making the complementarity between education and on-the-job training easier will help to increase the individuals' skills and, therefore their possibilities to find a job in the labour market.

Regarding the productive system, an increase in demand for university graduates will help to reduce their difficulties to find a proper job. This will reduce the proportion of over-educated university graduates, but also the likelihood for a process of job's upgrading.

This preliminary data analysis suggests that the Spanish labour market has experienced some difficulties to completely accommodate its university graduates, or its highly skilled workers as they have been defined here, in the first decade of the twenty-first century. However, an econometric analysis and a better measurement of over-education are needed. Using the same data as here (SLFS), Chapter Three applies the statistical mean approach to measure the degree of over-education and to estimate the predicted probability of being over-educated using a probit and a variation of the Heckman two-step model.

"Overeducated relates in this context to the connection between years and type of college studies and labor market earnings and opportunities, not to the broader cultural aspects of higher education"

Richard B. Freeman (1976, p. 4)

Chapter 3. To what extent has the Spanish labour market been able to accommodate the expansion of highly skilled workers?

In the last half of the twentieth century there was a substantial increase in the number of university graduates in most developed countries. In some cases, the labour market showed some difficulty accommodating them. This has reinforced researchers' interest in the qualification balance of the labour market. Focusing on Spain, this chapter analyses the ability of the Spanish labour market to absorb the increasing number of university graduates. Using data from the Spanish Labour Force Survey, specifically the second quarters from 1995 to 2018, I model the likelihood of a university graduate being over-educated to identify whether the Spanish labour market has been able to match them.

3.1. Introduction

The composition of the labour market completely changed during the last half of the 20th century in most developed countries. Generation after generation there has been a replacement of less educated workers by more educated and professional ones. This process has occurred in all developed countries, at different times and speeds, e.g. in the USA the big increase in the number of university students was in the 1950s-1960s; whereas in most European countries it was in the 1980s-1990s.[64] Chapter Two has demonstrated that Spain was not an exception, and the supply of and demand for university studies has been increasing, especially since the 1980s-90s. As a consequence, the proportion of the population holding a university degree more than doubled (it was 16.4 percent in 1995 and 37.3 percent in 2016) (Figure 2.2).

According to Hartog and Oosterbeek (1988, p. 187) over-education is a situation in the labour market where *"the individual is simply supposed to have been at school longer than necessary for this job*". Using this definition as a starting point, *the goal of this chapter is to disentangle whether there has been a saturation of the Spanish labour market during the last years of the 20th century and the beginning of the 21st*. In particular, this chapter aims at: first, measuring the level of education mismatch in the Spanish labour market during this period. Second, exploring the changes in the probability for an individual to be over-educated in the Spanish labour market during this period. Third, examining if those university graduates who are over-educated share a set of characteristics. And fourth, accounting for possible gender differences.

Statistical methods are used in this chapter to measure education mismatch. In particular, and following Verdugo and Verdugo (1989), the mean plus/minus one standard deviation of the years of education for the workers for a specific occupation (using the two-digit break-up occupations classification) in the first

---

[64] See Freeman (1976) for the U.S.A., Tight (2009) for the U.K. or De Miguel (1998) for the Spanish case.

year of the sample is used as a baseline to measure the degree of education mismatch for all the individuals of the same two-digit occupational category for all the period. Leaving the definition of the required education for each occupation constant in the first year of the period permits one to identify the effect that changes in education have on the likelihood of being over-educated. Once the proportion of over-educated university graduates has been calculated, a variant of the Heckman selection model is used to measure the likelihood for university graduates to be over-educated in the Spanish labour market and whether they share a set of characteristics. Finally, the predicted probabilities for those university graduates employed and over-educated in different years, regions, nationality, etc. are calculated by gender in order to account for possible differences between men and women.

The rest of the chapter is divided as follows. The next section summarises the main factors which have contributed to the upward trend on the supply of and demand for university studies in Spain since the decades of the 1980 and 90s. Section 3.3 describes the source. I use data from the second quarters of the Spanish Labour Force Survey (SLFS), for the period 1995-2018. Section 3.4 describes the sample used; employees between 25 and 64 years old who hold at least one university degree. The probit model used as a baseline model and the variation of the Heckman two-step model (Heckman, 1979) applied to account for a possible sample selection bias are presented in Section 3.5. The main findings are shown in Section 3.6. Finally, Section 3.7 presents the conclusions of the chapter.

3.2. The supply of and demand for university studies in Spain

In the last decades of the 20th century the increase in the number of university graduates was noticed in the Spanish labour market with the replacement of older generations. In 1977 only 4.4 percent of the active population aged between 25 and 64 held a university degree, whereas in 1995 this proportion rose to 11.8

percent, to 15.9 percent in 2000, and to 25 percent in 2018.[65] This upward trend on the supply of a higher educated labour force was the result of the combination of two factors: the expansion of the supply of higher education after the creation of new universities and new degrees, and the greater individuals' demand for this educational level (Bordas and Montané, 1993; San Segundo, 1997).

The restoration of democracy (1977) and the general administrative decentralisation (1979-1983) boosted the number of state universities. Technical universities were created and universities were expanded to demographically lower density areas in order to facilitate the access to university studies (Reques Velasco, 2007). The increase in the number of private universities at the end of the 1980s and beginning of the 1990s was due to the higher administrative independence of the universities after the enactment of the new legislation in 1983 (University Reform Act –Ley de Reforma Universitaria, LRU–), which was in part developed by a Royal Decree-Law in 1991 (Real Decreto 557/1991) (Egido Gálvez, 2006; Rahona López, 2008).

As a consequence, the number of universities and their supply of degrees more than doubled from the 1970s. Table 3.1 shows the expansion of the university system, first with the creation of new state universities in the 1970s and 1980s, and then with the creation of private universities from 2000s.

In spite of the increase in private universities, the proportion of state universities has always been quite important in Spain. They represented around 72 percent of the universities in the 1960s, 86 percent in the 1970s, and 93 percent in the 1980s and 1990s. Only in the first decades of the 21st century did the proportion of state universities decrease to 64 percent (50 out of 78 in 2019). The main reason was the expansion of the private universities (in 2019, 28 out 78 were private universities).

[65] For 1995 and 2000, data from García-Montalvo (2001); and data from the National Statistics Institute (http://www.ine.es) for 2018.

Table 3. 1. Universities in Spain by public-private ownership, 1960-2019

Year	State universities	Private universities	Total
1960	12	5	17
1968	14	5	19
1970	22	4	26
1978	28	4	32
1987	39	3	42
1994	42	7	49
2000	49	21	70
2006	49	23	72
2019	50	28	78

Source: Reques Velasco (2007) for data from 1960 to 2006; http://www.crue.org for 2019

In the Spanish case, the growth of private universities is as important as the increase of the state universities. The employability of those university graduates coming from private universities is equal or even higher than those graduated from state universities. In fact, in most of the cases, the proportion of university graduates from these private universities who were working in the Spanish labour market the year after their graduation is more significant than those university students who have graduated from a state university. Table 3.2 shows the proportion of university students who graduated from a Spanish university in the academic year 2012-2013 who were paying the official insurance in the first year after graduation.[66] Data are displayed by gender, field of study and university ownership (private or state).

Table 3.2 shows that, for both genders, the proportion of university graduates in the academic year 2012-2013 and working in the Spanish labour market in the following year (2013-2014) is higher for those who graduated from a private university. In the case of men, there was a difference of 12.2 percentage points (pp) (59.3 percent of university graduate men working in the Spanish labour market had graduated from a private university the year before, 47.1 percent had graduated from a state university), and 15.3 pp for women (58.3 percent of

[66] In Spain, contributing to the national insurance (*Seguridad Social*) is compulsory for all workers whether they are working as employees or as employers. Therefore, the proportion of people, in this case of university graduates, paying this national insurance is a good proxy of the proportion of actual workers.

university graduate women working in the Spanish labour market had graduated from a private university the year before, and 43 percent had graduated from a state university). Analysing by field of study, the highest activity rates are concentrated on those private university graduates on 'Social Sciences and Law' and 'Arts and Humanities', for both sexes (Table 3.2). But this is not a surprising result, the higher employability of specific degrees is well known in the literature (Dolton and Vignoles, 2000; Oliver, Ramos and Raymond, 1999). Indeed, the higher rates of students enrolled in these more demanded degrees are also a reflection of the higher probabilities of being employed in the future.

Table 3. 2. University graduates contributing to the national insurance in the first year after graduation by gender, field of study, and university ownership, 2013-2014 (in percentages).

	Field of study	State U.	Private U.	Total U.
Men	Social Sciences and Law	45.2	60.7	48.1
	Engineering and Architecture	58.6	66.0	59.5
	Art and Humanities	26.5	42.7	27.3
	Health Science	38.8	47.0	41.0
	Sciences	27.2	30.5	27.3
	Total	**47.1**	**59.3**	**49.0**
	Field of study	**State U.**	**Private U.**	**Total U.**
Women	Social Sciences and Law	45.7	62.5	49.0
	Engineering and Architecture	51.6	55.5	52.1
	Art and Humanities	29.1	42.9	29.8
	Health Science	41.8	48.8	43.1
	Sciences	30.1	35.0	30.3
	Total	**43.0**	**58.3**	**45.5**

Source: University statistics http://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/universitaria/estadisticas.html

In Spain, the workers' contributions to the national insurance depend on the professional category required by the workplace. Therefore, they may be divided according to the level of their contribution to the national insurance. In this sense, data from the university statistics are divided into three different groups: university graduate, middle and low group. As a consequence, the proportion of university graduates contributing to the national insurance as university graduates may be considered as a proxy of those university graduates being educationally matched. And following the same criteria, the proportion of university graduates contributing to the national insurance as middle or low groups could be considered as the proportion of over-educated university graduates in the Spanish labour market. Using these data, Table 3.3 shows that the highest proportion of university graduates matched in a university-graduate job in the Spanish labour market are those graduates from state universities. In 2013-2014, 64 percent of university graduate men from a state university were working in a university-graduate job. For women, the proportion of university graduate matched rose to 65.4 percent. In contrast, these percentages reduced to 47 percent for men and 41.8 percent for women when focusing on university students who graduated from a private university.

Therefore, data show that workers graduated from a state university depicts higher proportions of education match in the Spanish labour market. These results remain for all fields of study and for both sexes, although it is more significant for women who graduated with ‘Social Science and Law’ degrees, with a difference of more than 30 pp (62.5 percent from state universities and 31.8 percent from private universities); and for men holding a degree in ‘Social Science and Law’ (22.7 pp: 56.8 percent from state universities and 34.1 percent from private universities) and ‘Sciences’ (25.6 pp: 66.7 percent from state universities and 41.1 percent from private universities).

Table 3. 3. University graduates paying as working on university graduate jobs in the national insurance in the first year after graduation by gender, field of study, and university ownership, 2013-2014 (in percentages).

	Field of study	State U.	Private U.	Total U.
Men	Social Sciences and Law	56.8	34.1	39.6
	Engineering and Architecture	69.2	55.6	57.5
	Art and Humanities	55.5	37.8	39.1
	Health Science	80.6	70.2	73.5
	Sciences	66.7	41.1	41.8
	Total	**64.0**	**47.0**	**50.2**

	Field of study	State U.	Private U.	Total U.
Women	Social Sciences and Law	62.5	31.8	39.4
	Engineering and Architecture	66.9	50.8	52.8
	Art and Humanities	39.4	35.5	35.8
	Health Science	79.5	71.2	73.0
	Sciences	50.6	38.2	38.6
	Total	**65.4**	**41.8**	**46.7**

Source: University statistics http://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/universitaria/estadisticas.html

Hence, those individuals who graduated from a private university have a greater participation in the Spanish labour market, even if a lower proportion of them are working in a university-graduate job. All in all, these data demonstrate that in the Spanish case it is important to account for the supply of private university studies as well as of the state ones.

During this period, the pattern followed by the demand for university studies was parallel to the supply. Table 3.4 displays the number of undergraduate students enrolled in a degree course at a Spanish university from 1985-86 to 2015-16. Data show an increase in the number of university students until the academic year 2000-01. This upward trend in the demand for university studies was fed on the one hand, by the entrance into university of the weighty baby boom cohort; and, on the other hand, by the intensification of women's participation in university studies (Albert, 2008). Since the academic year 2000-01 there has been a slight decrease in the number of university students that is due basically to a decrease in the students enrolled in state universities, since the number of students enrolled in private universities was still increasing (Table 3.4).

Table 3. 4. University students by university ownership

Year	State universities		Private universities		Total universities	
	Students enrolled	Women (%)	Students enrolled	Women (%)	Students enrolled	Women (%)
1985-86	797,596	49.3	56,593	51.8	854,189	49.5
1990-91	1,080,386	50.8	60,186	54.4	1,140,572	51.0
1995-96	1,449,601	52.7	58,875	54.2	1,508,476	52.7
2000-01	1,437,722	53.5	117,452	53.3	1,555,174	53.5
2005-06	1,304,597	54.6	137,856	51.5	1,442,453	54.3
2010-11	1,252,832	54.2	172,186	55.2	1,425,018	54.4
2015-16	1,143,223	54.3	178,475	55.5	1,321,698	54.4

Source: University statistics http://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/universitaria/estadisticas.html

As a result of this superior supply of and demand for university studies, the proportion of individuals aged between 25 and 64 and holding a university degree has increased in the Spanish labour market. As was seen in Chapter Two, in the Spanish case, this percentage grew from 10.7 percent in 1995 to 13.9 percent in 2000, 20.3 percent in 2010, and 24.9 percent in 2018 (average of data from Tables 2.5 and 2.6). By gender, this upward trend has been more noticeable for women than for men (the proportion of university graduate women increased by 18.1 percentage points from 10.7 to 28.8 percent, and the proportion of university graduates men grew in 10.3 percentage points from 10.8 to 21.1 percent).

In most developed countries, including Spain, the labour market found some difficulties responding to these changes in the educational composition of the workforce, and some qualification imbalances appeared.

Measuring the degree of qualification imbalance in a determined labour market is important to evaluate to what extent the educational and labour market policies are functioning properly. As is known, in economic terms, a qualification imbalance produces a waste not only of personal, but also, of social resources, especially when this imbalance is due to the over-education of the workers (García-Montalvo, 2009). In this latter case, there may also be a decrease in the worker productivity or job satisfaction (Battu, Belfield and Sloane, 1999; García-Montalvo, 2009; Henseke and Green, 2016).

Furthermore, knowing which the main characteristics of the individuals suffering from over-education are is crucial to help the policy makers to design good policies to reduce this qualification imbalance or their effect on productivity.

The objective of this chapter is to estimate the changes on the conditional probability for a university graduate to be wrongly accommodated in a less skill demanding job by the Spanish labour market in the period 1995-2018. The analysis focuses on how the individual characteristics (age, sex, etc.), as well as the region, labour market characteristics (activity sector), or period influenced (positively or negatively) the probability for a university graduate to be over-educated in the Spanish labour market.

In fact, since the publication of *The Overeducated American* by Freeman (1976), the production of academic articles measuring and analysing over-education has increased in all developed countries. A summary including the main literature references can be found in Hartog (2000).

In the Spanish case, the first article measuring and analysing over-education was published in 1993 by Alba-Ramírez (1993). Applying subjective methods to data from a 1985 survey (Living and Working Conditions Survey), he found that 17 percent of the workers were over-educated.[67] Since then, using cross-sectional data as well as panel data, different authors have measured the level of over-education.[68]

For instance, using data from the SLFS and taking into account all the educational levels, García-Montalvo (1995) analysed the education mismatch between education and employment, and the tendency of over-qualification for the period 1985-1993 (before the changes on the university-studies structure in 1992 and the implementation of the Bologna Reform)[69] He concluded that over-qualification levels in Spain increased between 1985 and 1993, although he found a relatively small proportion of over-educated, around 7 percent for all the years considered.[70]

The low proportions of over-educated workers found by these first researchers were likely to be due to the period they focused on. They did so in the period before there was an increase in the Spanish supply of and demand for university studies and, therefore, of university graduates.

Just like the international studies, most studies on Spain focus on the effect of over-education on earnings (Aguilar Ramos and García Crespo, 2008; Alba-Ramírez, 1993; Alba-Ramírez and Blázquez, 2003; Murillo, Rahona and Salinas, 2010; Nieto, 2014), although there are also some studies about labour mobility (Alba-Ramírez, 1993; Alba-Ramírez and Blázquez, 2003; García Serrano and Malo, 1996) or related to job satisfaction (Badillo Amador, López Nicolás and Vila, 2008). Some of them have also centred on university graduates (Alba-Ramírez

[67] His results were smaller than the ones obtained for other European countries. The reason is that he used data from 1985 when the increase in the number of the university graduates in Spain was only in its beginnings.
[68] A summary including some of the most important studies which have measured the education mismatch in Spain and its results can be found, for instance, in Aguilar Ramos and García Crespo (2008); Fernández (2004); Nieto (2014); or Sanromá and Ramos (2004).
[69] In fact, the Bologna Reform was implemented in Spain in the academic course 2008-2009. The main objective of this reform was to create a more common university system in all the European countries.
[70] He uses the term ‘over-qualification’ instead of ‘over-education’, since he presumes that “it is not the level of education which is inadequate but employment or occupation.” (*my own translation*) (García-Montalvo, 1995, p. 11).

and Blázquez, 2003; Mora, García-Montalvo and García-Aracil, 2000; Parellada, Duch and Alvarez, 2009). Nevertheless, few studies focused on the probability of being over-educated (García-Montalvo, 1995).

3.3. The source: The Spanish Labour Force Survey

The dataset used in this chapter is based on individual data from the Spanish Labour Force Survey (SLFS). There are other Spanish statistics that can be used to study the labour market, such as Statistics of Employment (done by the *National Institute of Employment – Instituto Nacional de Empleo*), Annual Labour Cost Survey (done by *National Statistics Institute – Instituto Nacional de Estadistica*) or Survey of Work Situation (done by the *State Department of Employment and Social Affairs – Ministerio de Trabajo y Asuntos Sociales*).[71] However, SLFS is the only survey which, first, has been taken periodically since 1964 and over a long period of time. And second, it contains information about both the formal educational level once finished and the labour market status of the population. Consequently, it is one of the best sources for analysing the changes of the Spanish labour market, although it does not contain any information about wages, real individual skills, or the skills required for the job.[72]

Despite its methodological and classification changes (Pérez Infante, 2006), the most important characteristics of the population and their occupations appear in all their waves and, they are comparable in most cases. For instance, between 1995 and 2018, the period considered here, there were two changes in the classification of the educational level (2000 and 2014), and one in the classification of the occupations, in 2011.

[71] A good description of the available statistics, their sources, objectives and period covered can be found in the Spanish National Statistical Institute web page (http://www.ine.es). A more detailed description can be found in Pérez Infante (2006).
[72] Only the SLFS carried out in 2005 contained wages data. Since SLFS does not contain data on wages it does not allow calculating returns to education. This is one of the main reasons why this source is not usually used to analyse over-education, using Mincerian equations.

Regarding the educational classification, there is a correspondence table which allows the standardisation of the categories of both series.[73] Therefore, it is possible to obtain a time series for the variables and period of this study, creating some common definitions for these variables (García and Gómez, 1994). Besides, in this chapter each educational level has been turned into years of education to apply the mean statistical approach for measuring the degree of over-education. Furthermore, converting the educational level to years of education also makes the comparison between the different educational classifications easier. A detailed table including the different educational classifications and their equivalences on years can be found in Appendix 2.

The comparison between classifications in the case of occupational classifications is slightly more complicated. There is only one change in the occupational classification, and it was in 2011. Data from 1995 to 2010 is classified using the National Classification of Occupations of 1994, and data from 2011 to 2018 is classified using the National Classification of Occupations of 2011.[74] They are national classifications valid only for the Spanish labour market, but its categories have a correspondence in the 2008 International Standard Classification of Occupations (ISCO-08).[75] There is also a correspondence table between the two national classifications. Unfortunately, the correspondence table is applicable for a three-digit and four-digit break-up of the occupational classification and the data used in this chapter is broken-up for a two-digit occupational classification.[76] In fact, as the National Statistical Institute (INE)

[73] A correspondence table can be found on INE's webpage: http://www.ine.es.
[74]The complete structure and the descriptive report for the National Classification of Occupations can be found on https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177033&menu=ultiDatos&idp=1254735976614
[75] ISCO is the International Standard Classification of Occupations used by the International Labour Organization (ILO): http://www.ilo.org/public/english/bureau/stat/isco/index.htm.
[76] The correspondence between the National Classification of Occupation of 1994 and 2011 can be found on https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177033&menu=ultiDatos&idp=1254735976614

states in its web page, a direct correspondence between both series is impossible.[77]

Therefore, in order to have consistent results, in this chapter the series are split into two time periods: 1995-2010 and 2011-2018. As a consequence, here there are two different baseline levels of over-education: 1995 and 2011. The definition of required education for each occupation in 1995 will be used to analyse the probability of being over-educated from the period 1995-2010; and the required education in 2011 will be used for the period 2011-2018.

3.4. Data

Following the main literature that used the SLFS, in this section, I only use data from the second quarters because they are usually the least seasonal ones (Albert, 1998; Toharia, 2003). Moreover, individuals residing in Ceuta and Melilla have been excluded from the sample (Ceuta and Melilla are territories, independent cities, situated in Northern Africa and they are out of the Spanish labour market).[78] Individuals working in the army have also been excluded from the sample, since they do not participate in the Spanish labour market.[79]

The analysis focuses on the period between 1995 and 2018. This analysis finishes in 2018 because this is the latest year of the SLFS with individual data available. It starts in 1995 for two main reasons. First, the middle of the 1990s is when two of the most important factors which contributed to the Spanish economy and society modernisation occurred: first, an expansion of the supply of a higher

[77] "*Up until 2010 inclusive, the classification of occupations used was CNO-1994. As of the first quarter of 2011, was introduced in the EAPS the new CNO-2011.(…). This change in classification assumed that it is not possible to directly link the information classified by type of occupation carried out at work without specific instruments which would make it possible to go from one classification to another.*" https://www.ine.es/en/daco/daco42/daco4211/nota_epa_retro0011_en.pdf

[78] In fact, Ceuta and Melilla were considered by the SLFS for the first time in 1989. Data from Ceuta and Melilla represents only 0.7 percent of the sample, therefore their exclusion does not affect the results.

[79] Population employed in the army represents only 0.2 percent of the sample, therefore their exclusion does not affect the results.

educated labour force, which sharply increased in less than three decades. And second, the literature considers 1995 as a pivotal point for the Spanish labour market, in the sense that the demand for more skilled labour force increased (Oliver, Raymond and Sala, 2001). As a matter of fact, in an analysis of the returns to education in Spain in the decade 1981-1991, using the Family Budget Survey, Vila and Mora (1998) concluded that their results

> *"suggest that the labor market has been able to absorb an increased supply of higher education graduates by shifting from industry to service production" since "Many jobs have disappeared in the agriculture and industrial sectors, while new white collar jobs requiring at least an academic secondary certificate have been created in the modern service sector. In particular, higher education has become a prerequisite for entry to the upper segment of the white collar labour market, which is associated with high wages and job mobility"* Vila and Mora (1998, p. 177)

Barceinas *et al.* (2001) analysing data from the same period (before 1995) also concluded that

> *"demand for more educated people has even overtaken supply. Probably, both the technological change and the tertiarisation that the Spanish economy has undergone in the last 20 years could explain this match between supply and demand"* Barceinas *et al.* (2001, p. 240)

The analysis of this chapter focuses on individuals between 25 and 64 years old who hold at least a university degree. I follow here the same criteria as in the rest of this thesis and situate the lower age threshold at 25 years old.[80] Likewise, the upper threshold is at 64 because 65 years old is the legal age of retirement and therefore most people finish their careers at that age. The sample only includes people who hold at least one university degree. The reasons for the selection are, first, university graduates are the focus of the research question; and second, they are more likely to be over-educated than people holding lower educational

[80] As was explained in Chapter 2, at the age of 25 years old most of the students who enrolled for the first time at university at the age of 18-20 already graduated. Besides, this low threshold has been also used in the literature. See, for example, Alba-Ramírez (1993), García-Montalvo (2001) and Albert (2008).

levels. People with higher educational levels will be over-educated unless they secure a highly skilled job, whereas people with lower educational levels can hold a low skill demanding job and they are still not over-educated, although they will also be mismatched (under-educated).[81]

It is worth remembering here who is included in the group "people who hold at least one university degree", particularly in the Spanish case. As seen in Chapter Two, until 1992, there was two types of university degrees in the Spanish university system: three-year degrees ('diplomaturas') and five-year degrees ('licenciaturas').[82] This chapter focuses on those university students with at least a five-year degree, that is a long-cycle degree ('licenciatura'). I had considered including all university graduates, independently of whether the degree is a short-cycle ('diplomatura') or a long-cycle ('licenciatura') because previous literature suggests that holding a university degree is what really makes the difference, even if it is a short-degree (Barceinas *et al.*, 2000; Consejo de Universidades, 1987; Riviére Gómez, 2002). However, data analysis demonstrates that the model works better when individuals holding a three-year university degree are excluded. A three-year degree is seen as a more professional degree and, moreover, people with higher vocational training were able to enrol to this education grade. Despite this, the results of the model including individuals holding the three-year university degree have also been included in Appendix 2. Furthermore, these latter results are also discussed in the text in order to contrast them with the results when they are excluded from the sample.

All in all, the final sample used in this chapter for the period 1995-2018 includes employees aged between 25 and 64, whose educational level is at least a five-year university degree. Note that the variable used accounts for the educational level finished. Therefore, only those people who have actually finished the degree

[81] It does not mean that the only group that can be over-educated are university graduates. As García-Serrano and Malo (1996) found, other educational levels can also be over-educated (García Serrano and Malo, 1996, p. 113).
[82] For the pre-Bologna Spanish university system see Mora *et al*. (2000).

and hold the degree are included, while people who have attended university studies but who have not finished the degree are not included.

Following the main literature, employers, self-employed, and unpaid family workers have been excluded from the data, since it is difficult to determine the real level of education required for the job when someone is self-employed (García-Mainar and Montuenga-Gómez, 2005; Verdugo and Verdugo, 1989).

The sample includes other variables related to personal characteristics: gender, age, and nationality; related to the labour market as activity sector; and the place of residence within the country (NUTS-1). All these variables are included in the model as control variables.

3.5. Methodology

Statistical methods are applied to measure the degree of education mismatch. As explained in the literature review chapter, there are three main ways to measure the qualification (im)balance of the labour market: subjective, objective, and statistical methods. None of them is perfect, and usually the characteristics of the data are the ones which define which method will be the best to measure the mismatch.

The Spanish Labour Force Survey does not provide individual's assessment related to the educational level required by the workplace. Therefore, subjective methods cannot be used to analyse it. But SLFS does allow researchers to use the other two methods: objective and statistical. In fact, both have been already used by some researchers. For instance, García-Montalvo (1995) and Oliver, Ramos and Raymond (2001) used statistical methods and García-Montalvo (1995) used objective methods to analyse over-education in Spain at the end of the 80s beginning of the 90s.

Objective methods are difficult to apply to this survey since they need an 'objective' definition of the educational level required by the workplace, a sort of Dictionary of Occupational Titles where the type and the degree of skills required by the workplace are defined. See, for instance, Rumberger (1981). And these definitions likely change over time, making the longitudinal analysis more complicated. Statistical methods can be calculated using only the individual's years of education. External information is not necessary here.

In this chapter, I use statistical methods to measure the degree of education mismatch in the Spanish labour market. To do so, I first calculate the number of years needed to finish each of the educational levels (Appendix 2).

Second, following Verdugo and Verdugo (1989), I measure the education mismatch using a variation of the statistical mean approach. Verdugo and Verdugo (1989) used the mean of the number of years of formal education for each occupation of the three-digit occupation classification. In contrast, in this chapter I calculate the mean of the number of years of formal education for each occupation of the two-digit occupation classification.[83] The objective is to determine the threshold for each occupation in the first year of the sample: 1995 for the first period (1995-2010) and 2011 for the second period (2011-2018). The threshold is on the mean plus/minus one standard deviation. In order to have an actual measurement of the qualification imbalance in the Spanish labour market, all individuals working as employees and aged between 25 and 64 years old have been included in the calculations. Individuals working in the army and those living in Ceuta or Melilla have been excluded.

As a result, the population working in a specific occupation is classified as *over-educated* if they have been at school for more years than the mean plus one standard deviation of the years of the other individuals working in that occupation. They have been classified as *under-educated* if they have been at school for fewer years than the mean minus one standard deviation compared to the rest of the counterworkers of the same occupation. And they are classified as *adequately educated* for the occupation they are holding when the number of years they have been in the educational system is within the range between the mean plus/minus one standard deviation.

[83] Occupations' classification can be found in Appendix 2.

I considered the option of calculating the mean accounting also for the activity sector. The reason was the difference in the required educational level that the same occupation can have in a different activity sector. For instance, it is possible that one university graduate working as a clerk with the same years of education is classified as over-educated in one activity sector (e.g. Manufacturing) and as education-matched in another one (e.g. Professional, scientific and technical activities). However, due to the limited data (only one observation) in some of the subgroups, it is not advisable to use these more disaggregated classifications to calculate the mean of the years of education. Therefore, the upper- and lower-limit to determine whether an individual is over-educated or not is related only to the occupation classification.

Furthermore, in order to avoid mistaking the results, the cut-off point between those who are mismatched and those who are matched is only measured in the first year of the samples: 1995 for the first sample (1995-2010), and 2011 for the second sample (2011-2018). According to Hartog (2000, p. 132), *"[using statistical approach] measures allocation, actual assignment practice as determined by hiring standard and labour market conditions*". Therefore, if the level of education mismatch is calculated for each year of the sample it is difficult to disentangle which part of the changes on the proportion of mismatch is due to the changes on the educational level of the workers and which part is due to changes in demand for them.

Another possible problem using the statistical approach to measure the level of education mismatch is using the arbitrary cut off point of one standard deviation. As a consequence, it can underestimate the matching level since it increases the cut-off point with two standard deviations (one plus and one less). However, given the important proportion of over-educated workers found in the Spanish case, this does not seem to be the case.

Besides, this method *"usually finds symmetry in under- and overeducation"* (Hartog, 2000, p. 133). Table 3.5 and Figure 3.1 will demonstrate that the latter is not a problem when measuring the education mismatch among wage-earning university graduates in the Spanish labour market.

Therefore, as was seen in the literature review chapter and earlier in this chapter, the statistical approach is not a perfect measure of education mismatch either. However, it allows the longitudinal analysis for cross-sectional data when the level of mismatch is maintained constant in the first year of the sample.

The main goal of the chapter is to estimate the changes in the conditional probability for a university graduate not to be accommodated in a proper skilled job. The analysis particularly focuses on the effect of the following control variables on the likelihood of a university graduate being over-educated: individual characteristics, place of residence in the country, labour market characteristics, and period of time.

To calculate this probability, there are two factors which should be taken into account. First, the dependent variable is dichotomous (being or not being over-educated), which means that a logit or a probit model should be used. Second, only those individuals who are employed are susceptible of being over-educated. In fact, the probability of being employed is a sort of pre-selection, and some people may prefer being unemployed than working in a low qualified job; those people likely share some characteristics. Therefore, it seems sensible to suppose that the likelihood of securing a job is not done at random. If so, some personal characteristics, as well as the area of residence or period of time, may affect the fact that an individual is employed. In other words, the estimation model should account for this possible sample selection bias. The variation of the Heckman two-step model used here fulfils both conditions: it estimates the likelihood of being over-educated using a probit and accounts for the probability of being currently employed.

According to Trevis Certo *et al.* (2016) "*sample selection bias is the result of a special case of endogeneity*", which makes it easy to confuse a simple case of endogeneity and a real sample selection problem. The Heckman two-step model explores whether a sample selection is biasing the core result.

The Heckman two-step model uses two different equations: the outcome equation and the selection equation. The selection equation is run first and, here, it estimates the probability of being employed. The outcome equation used here takes into account the results of the selection equation to estimate the probability of being over-educated given that the individual is employed in a job.

However, the Heckman two-step model should be used cautiously since it makes additional assumptions about the data. First, the model can cause some problems when working with a small size sample or if the sample contains few observed data. The Spanish Labour Force Survey is a large sample with more than 70 percent of observed data (72.8 percent for the period 1995-2010, and 78.6 percent for the period 2011-2018), therefore sample size is not an issue.

But the model can also have some problems when

> *"the errors are not distributed normally", either if "the correlation between the errors of the regression and the selection equations are small", or when "the degree of collinearity between the explanatory variables in the regression and selection equations is high"* Kennedy (2014, p. 271).

To test whether there is a correlation between the errors of the two equations (outcome and selection) one can test the significance level of the *athrho* and the Wald test of independence equations (*rho*). The values and significance level of the *athrho* and *rho* displayed in the next section (Section 3.6) will demonstrate that the null hypothesis of no correlation between error terms of the outcome and selection equations is rejected for the data from the second period 2011-2018 when the three-year degree university graduates are excluded, and for both periods when all university graduates are included in the sample.

To avoid the collinearity between the explanatory variables of both equations, the selection equation should have at least a covariant which does not appear in the outcome equation. In fact, this is one of the most important factors to consider when a Heckman two-step model is designed. For the model considered here, the selection equation of the model includes a set of explanatory variables which is not covariate in the outcome equation and that is accounting for the *relationship with the person of reference*. It is a good candidate to consider as exclusion variable. It is expected that the probability of being employed (selection equation) would be affected by the position of the individual within the household (for instance, the head of the family is more likely to be employed than the children).

But the probability of being over-educated (the outcome equation) would not be affected by the individuals' relationship with the person of reference. Moreover, a post-estimation test and the relatively small standard errors of the estimation coefficients of both equations have confirmed its validity.

Nevertheless, in order to be sure of the appropriateness of using the Heckman two-step model I also follow the advice of Trevis Certo *et al.* (2016), who state some steps should be considered when the Heckman two-step model is applied. First, it is important to ensure that the dependent variables of the outcome equation are only observed in a subsample in order to be able to consider a potential sample selection. The data used in this chapter accomplish this condition, since only around 50 percent of the workers are over-educated.

Second, the possible effect of other omitted variables in the correlation between the errors of both equations should be considered. For this model, there are two main variables which can affect the probability of being over-educated and that have not been considered here, since the data does not include them for all the years. One is the *size of the company* where the individual is working, understanding as a 'size' the number of workers employed in it, although the *activity sector* variable can indeed be considered as a proxy of the *size of the company* since one of the categories includes all individuals working in the public sector. It is expected to find a positive relationship between the size of the company and the probability of being employed (selection equation [2]), but a negative relationship between the size of the company and the probability of being over-educated (outcome equation [1]) (Büchel and Mertens, 2000). Therefore, it is expected that including this new variable would reinforce the correlation among the errors of both equations (the significance level of *athrho* and *rho*).

The *field of study* of the university degree is another potentially omitted variable. It is known that the field of study determines the likelihood of both being employed and being over-educated (Battu, Belfield and Sloane, 1999; Dolado, Felgueroso and Jimeno, 2000; Dolton and Vignoles, 2000; Lindley and McIntosh, 2015; Meroni and Vera-Toscano, 2017). In this case, the expected effect that the introduction of these new independent variables could have on the model is not straightforward. It will depend on the classification of the field of study, and their category reference.

Third, the significance of the explanatory variables in the selection equation estimated in the first stage indicates the existence of sample selection. Finally, although not essential, the significance of the Wald test of independent equations (*rho*) should reject the null hypothesis of no correlation between the error terms. The results of the model in Section 3.6 will show when the model also accomplishes these latter conditions.

Summarising, the Heckman two-step model used here estimates the conditional probability of being over-educated, using a selection model that accounts for the likelihood of securing a job. It can be formalised as follows:

$$\Pr(\text{over-educated}=1|X)=\beta_0+\beta_j X_{ij}+\varepsilon_i \qquad \text{Outcome equation} \quad [3.1]$$

$$Z^* = \alpha_j W_{ij}+e_i \qquad Z=\begin{cases} 1 \text{ if } Z^*>0; \\ 0 \text{ otherwise} \end{cases}$$

$$\Pr(Z=1)=\phi(\alpha_j W_{ij}) \qquad \text{Selection equation} \quad [3.2]$$

where X_{ij} is a matrix with all the explanatory variables (j), which varies depending on the specification and the sample (i), and W_{ij} are the independent variables of the selection function, as follows. ‘Outcome equation’ [1] accounts for the probability of being over-educated (probability of a university graduate being over-educated), whereas ‘selection equation’ [2] accounts for the probability of being employed as an employee.

More precisely, the variables included in the model are:

Outcome equation

- **Dependent variable**: *whether the individual is over-educated* (dummy variable: 1 if s/he is over-educated, 0 if s/he is not over-educated).

- **Independent variables:** *personal characteristics* (woman, age group, and nationality), *a spatial variable* (region of residence), *a variable to define the job characteristics* (activity sector), and *period* (single years).
  Where,
  - *Woman* (dummy variable: 1 if woman, 0 if man).
  - *Age group* (dummy variable: 4 categories: 25-34, 35-44, 45-54; and 54-64; reference category: 25-34).
  - *Not Spanish* (dummy variable: 1 if Not Spanish, 0 if Spanish or double nationality).
  - Activity sector (dummy variables: 10 categories; reference category: Agriculture, mining and fishing)
  - *Region* (dummy variable based on NUTS-1: 7 categories; reference category: Madrid region).
  - *Year* (dummy variable: single years; reference categories: 1995 for the period 1995-2010, and 2011 for the period 2011-2018).

Selection equation

- **Dependent variable:** *whether the individual is employed* (dummy variable: 1 if s/he is working as employee, 0 if s/he is an employee unemployed).
- **Independent variables:** *personal characteristics* (woman, age group, nationality, and the relationship with the person of reference), *a spatial variable* (region of residence), and *period* (single years).
  Where,
  - *Woman* (dummy variable: 1 if woman, 0 if man).
  - *Age group* (dummy variable: 4 categories: 25-34, 35-44, 45-54 and 55-64; reference category: 25-34).
  - *Not Spanish* (dummy variable: 1 if Not Spanish, 0 if Spanish or double nationality).
  - *Relationship with the person of reference* (dummy variable: 9 categories: person of reference, partner, children, children-in-law, grand-children, parents, other relatives, people working in the housework, other people; reference category: person of reference).

- *Region* (dummy variable based on NUTS-1: 7 categories; reference category: Madrid region).
- *Year* (dummy variable: single years; reference categories: 1995 for the period 1995-2010, and 2011 for the period 2011-2018).

In the original SLFS data set (from 1995 to 2018) age is a continuous variable. It has been recoded into four categories (25-34; 35-44; 45-54; 55-64) and transformed into dummies using the youngest group (25-34) as the reference category. This categorisation aims at splitting the individuals by the main life stages (25-34: first stages in the labour market; 35-44: independence and creation of a new family; 45-54: career consolidation; and 55-64: pre-retirement). Besides, the estimated model has higher statistically significant levels when these age groups are used in comparison with other age-group categories.

Gender is defined by the variable *Woman* which accounts for the gender differences in the probability of being over-educated. *Woman* is defined as a dichotomous variable with value '1' when the individual is a woman and '0' when he is a man. Hence, man is the reference category.

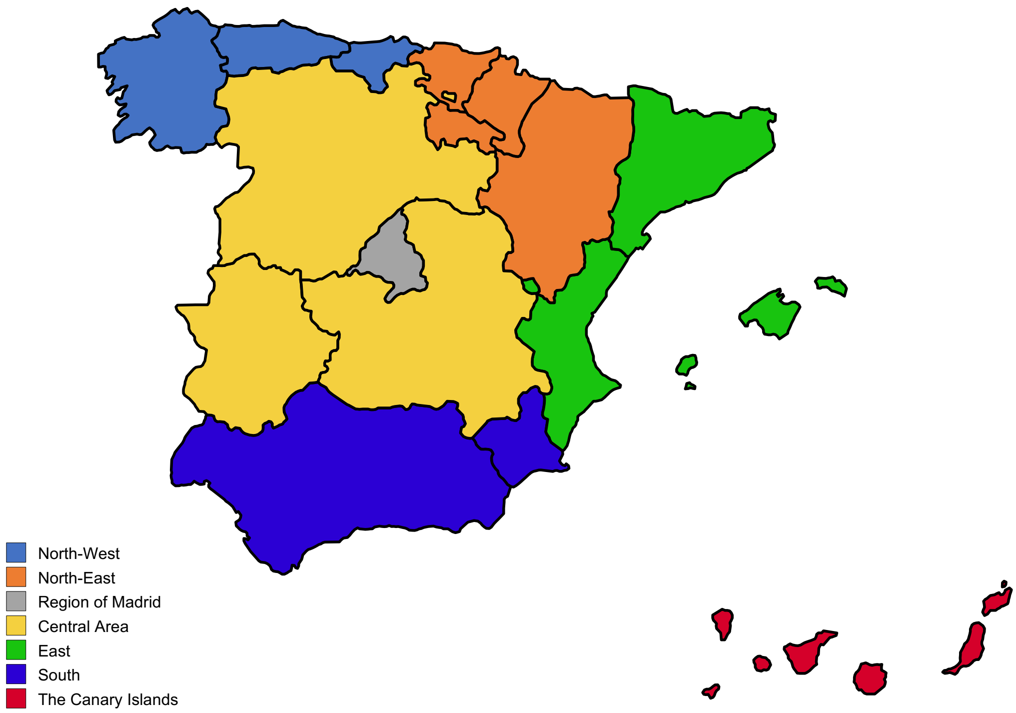
A nationality dummy variable is used to account for any discrimination by nationality in the Spanish labour market. Taking Spanish as the reference category the variable *Not Spanish* has value '0' when the individual has the Spanish nationality or dual nationality and value '1' for those individuals with 'foreign nationality'.[84]

A range of dummies for administrative regions is included to control the area effect (Madrid is the reference category). As in the rest of the chapters of this thesis, I am using NUTS-1 to account for the *region*. A map representing these territorial regions is drawn in Figure 3.1.

There are some differences in the productive structures among provinces (Pablo-Romero and Gómez-Calero, 2008), even within the same region of NUTS-2.

[84] Originally the Spanish Labour Force Survey divided the nationalities into three different categories: 'Spanish', 'dual nationality' or 'foreign'. Individuals with dual nationality are those who have Spanish nationality and another one.

However, the communication and transportation systems were substantially improved at the end of the 20th century and beginning of 21st. As a consequence, recently, workers have been more used to daily commuting among provinces, within the same region of NUTS-1. This justifies using NUTS-1 for the area effect analysis (Sanromá and Ramos, 2004). In fact, some studies have revealed the negative effect that moving regions has on the probability of being over-educated (Battu, Belfield and Sloane, 1999). The likelihood of being matched increases in parallel to the extension of the job-research area, because job offers are also larger.



Source: https://ec.europa.eu/eurostat/web/nuts/national-structures-eu

Figure 3. 1. Map of the Spanish regions (NUTS-1)

Regarding the labour market situation, there is a set of independent variables controlling for *activity sector*. Following the high-level SNA/ISIC aggregation, eight dummy variables have been included in the model to account for each of the activity sectors (Appendix 2).[85] These variables aim at considering the impact that being employed in a particular activity sector has on the probability of being over-educated. Each activity sector has its own productive structure, and therefore, they differ on the demand for highly skilled workers. As a consequence, the introduction of a variable accounting for these differences is recommended when the probability of being over-educated is considered.

According to previous literature, variables related to the labour market situation are potentially controversial since they may be endogenous (Greenwood, Jenkins and Vignoles, 2007). They are likely to be related to securing a highly skilled job as well as to the educational level. However, the particular characteristics of the Spanish labour market, with high unemployment rates and different opportunities depending on the activity sector justify controlling here this labour market variable.

Control variables indicating the period (*year*) when the survey was carried out are included to explain the period effect, that is, the changes in the probability of being over-educated due to the economic situation. Dummy variables have been constructed using single years.

Given the important differences in the labour market conditions between those workers securing a permanent contract and those securing a fixed-term contract, I also considered accounting for the effect of tenure. As was seen in the previous chapter, fixed-term contracts became very common in the Spanish labour market, especially for low-skilled jobs or for entering the labour market, particularly after the 1984 legal reform.[86] However, the post-estimation test suggested excluding it since it reduced the power of the model.

---

[85] https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF

[86] In 1984 the general Spanish law which regulates the rights of the workers changed (Ley 32/1984 Reforma del Estatuto de los Trabajadores). The most important change was the flexibility in the possibility of contracting workers for a fixed-term. A detailed analysis of the effects of this new legislation in the increase of the temporary contracts in the Spanish labour market can be found in Pérez Infante (2008). This legal modification was controversial, since not everybody agreed on the idea that the Spanish labour market needed a higher level of flexibility. On the contrary, some researchers

Furthermore, I considered including a variable to control whether the individual was working in the private or public sector. Indeed, this variable could also be used as a proxy to account for the size of the company, since private firms are more likely to be smaller than those from the public sector. This is particularly important in the Spanish industrial network where more than 90 percent of the private firms are small and midsize enterprises (SMEs).[87] However the inclusion of the activity sector is in fact already considering this since one of the dummy variables ('Public administration, defence, education, human health and social work activities') accounts for the effect that working in public sector has on the probability of being over-educated.

In order to be able to contrast the estimated coefficients of the Heckman two-step model, I also run a baseline model. Using the same sample I run a probit model to estimate the probability for a university graduate to be over-educated controlling for the same independent variables. An over-estimation of the coefficients is expected when the probit model is applied since it does not account for the effect of the probability of being already employed which is expected to be lower than 1.

For both models, the probit and the Heckman two-step, the individual sample weights have been considered. Furthermore, the standard errors have been calculated taking into account the possibility for the observations to be clustered among households. In particular, I have considered the possible independence of individuals across groups (household), but the likely dependency across them within the groups (household), to obtain robust standard errors.

The interpretation of the coefficients estimated by a probit model is not entirely straightforward. The accumulative function of a probit model is based on a standard normal distribution, therefore, the estimated coefficients account for the change of the dependent variable in z-score units provoked by a one-unit change

---

thought that the main problem of the Spanish labour market was an excess of flexibility. See Muñoz de Bustillo Llorente and Antón Pérez (2011) or Recio (1994).

[87] https://industria.gob.es/es-es/Servicios/MarcoEstrategicoPYME/Marco%20Estratégico%20PYME.pdf

in the independent variable. However, the coefficients estimated by the probit model are useful to anticipate the positive or negative effect of the indicator variable, *caeteris paribus*, on the dependent variable (in this case, being over-educated). The comparison of the estimated coefficients for the probit and the Heckman models is also used here to check the appropiateness of using the selection model. Predicted conditional probabilities for a university graduate to be over-educated have also been calculated. The results are displayed in the following section.

3.6. Results

This section is divided into three subsections in order to facilitate the analyses of the results. The following subsection includes the distribution of the educational mismatch in the Spanish labour market. Subsection 3.6.2 depicts the estimated coefficients for both models (probit and the Heckman two-step) and both periods (1995-2010 and 2011-2018). The third and last subsection shows the conditional predicted probabilities of success, that is, the probability for a 25-64-year-old university graduate to be employed and over-educated conditional to the probability of being already employed.

3.6.1. Education mismatch in the Spanish labour market (1995-2018)

Using the sample excluding wage-earning three-year degree university graduates, this subsection depicts the distribution of the education mismatch in the Spanish labour market when the statistical mean approach is applied to measure it (Table 3.5 and Fig. 3.1). The same results, but for the sample including all university graduates (including three-year degrees) can be found in Appendix 2.

According to the data, the average proportion of over-educated university graduates for the first period (1995-2010) is 51.6 percent when the three-year degree university graduates are excluded. This reduces to 47 percent for the second period (2011-2018) (Table 3.5). For the sample which also includes the three-year degree university graduates (Appendix 2), the mean of the proportion of over-educated is 47.3 percent in the period 1995 to 2010, and 35.5 percent for the second period (2011-2018).

These data suggest a reduction in the proportion of over-educated university graduates from the first period (1995-2010) to the second one (2011-2018). However, the use of different baseline education mismatch levels in each of the periods (1995 for the first period and 2011 for the second period) imply changes in the education mismatch thresholds and therefore it includes variations in the supply of and demand for this sort of workers. Given that the proportion of over-educated university graduates in the previous period, but also in the following years of this second period, indicates an upward trend when the definition of the thresholds are constant, this suggests an education upgrading of the jobs or a crowding-out or bumping down in the Spanish labour market from 1995 to 2011. In other words, the demand for workers holding higher educational levels (more years of education) has increased in the Spanish labour market between 1995 and 2011.

Table 3. 5. Education mismatch in the Spanish labour market by gender. Excluding three-year degree university graduates (in percentages)

		Men	Women	Total
1995-2010	Under-educated	0.0	0.0	0.0
	Adequately educated	49.6	47.0	48.4
	Over-educated	50.4	53.0	51.6
2011-2018	Under-educated	0.0	0.0	0.0
	Adequately educated	54.8	51.6	53.0
	Over-educated	45.2	48.4	47.0

Source: SLFS (2^{nd} quarters), INE

Furthermore, for both periods, the proportion of over-educated workers is lower when all the university graduates are considered (Appendix 2). Given that the educational level has been converted into a number of years of education, these results can be explained by the lower number of years needed to finish a three-year degree. As a consequence, the education match between years of education held and years of education required by the job is more likely.

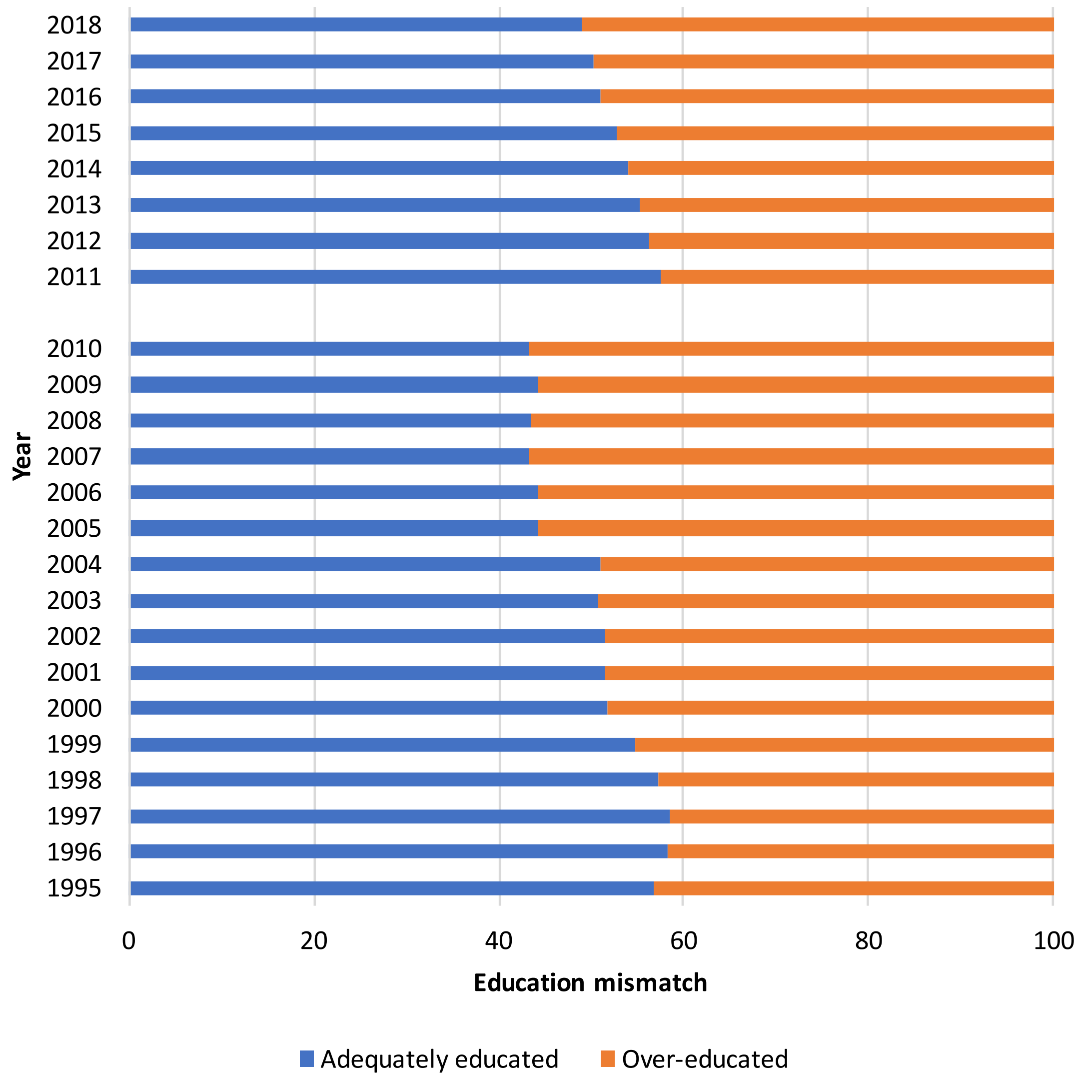
Table 3. 6. Gender distribution of university graduates by type of degree (in percentages)

		Men	Women
1995-2010	Three-year degree	42.5	57.5
	Five-year degree	50.6	49.4
	Posgraduate and Doctorate	57.4	42.6
2011-2018	Three-year degree	38.8	61.2
	Five-year degree	45.9	54.1
	Postgraduate and Doctorate	46.7	53.3

Source: SLFS (2nd quarters), INE

It is worth highlighting here the gender differences and the disparity in the results over the different samples. The proportion of over-educated university graduate women is higher than their men counterparts when the three-year degree university graduates are excluded, whereas it is lower when all the university graduates are considered. These results can be explained by the distribution of the sample (Table 3.6). For the first period, 1995-2010, the proportion of women is higher among the three-year degree university graduates (57.5 percent), whereas the proportion of men is higher among the higher educational levels (five-year degree: 50.6 percent, and postgraduate and doctorate: 57.4 percent). Data from 2011-2018, show a different gender distribution, and the proportion of women is higher in all university educational levels. However gender differences are even higher when the three-year degree is considered (61.2 percent of women and 38.8 percent of men).

Figure 3.2 shows the annual tendency in the distribution of the education mismatch in the sample excluding three-year degree university graduates. A figure using the sample including all university graduates can be found in Appendix 2. The first difference between the results of the two samples is on the under-educated. There is no under-educated when the three-year degree university graduates are excluded (Fig. 3.2), whereas between 2 and 3 percent of the university graduates are under-educated when all the university degrees are included (Appendix 2).



Note: When the three-year degree university graduates are excluded from the sample, the proportion of under-educated university graduates is 0 for all the period (the average result by gender and for each of the periods can be seen in Table 3.5)
Source: SLFS (2nd quarters), INE

Figure 3. 2. Education mismatch among university graduates in the Spanish labour market, 1995-2018. Three-year degree university graduates excluded (in percentages)

For both samples (including and excluding three-year degree university graduates) and for both periods (1995-2010 and 2011-2018), there is an upward tendency of the proportion of over-educated wage-earning university graduates compared with the baseline year (1995 for the first period and 2011 for the second period). This result makes us expect an upward trend in the probability for wage-earning university graduates to be over-educated in the Spanish labour market during this period.

Besides, in 2005 there is an increase in the proportion of over-educated workers compared with the previous years. This growth persists in the following years, from 2006 to 2010 and it appears in both samples when the three-year degree university graduates are included (Appendix 2) and excluded (Fig. 3.2). It could be explained by the increase in the number, and therefore, in the proportion of university graduates holding postgraduates degrees or even doctorates. It is lower than 2.5 percent until 2004 and more than 4.3 percent from 2005 when all the university graduates are considered; and from 4.4 percent in 2004 to 7.5 percent in 2005 when the three-year degree university graduates are excluded from the sample. The fact that people holding these higher educational levels have been more years in the educational system increases their likelihood of being over-educated since the number of years of education needed for doing the job are fixed at the levels of 1995 for each occupational category.

Focusing on the sample excluding the three-year degree university graduates (Fig. 3.2), in both periods there is an increase in the proportion of over-educated university graduates. However the starting point is lower for the second period (2011-2018) in comparison with the period from 1995 to 2010. On average, there is a difference of 5.5 percentage points between periods. As has already been stated, part of the differences between the results of each period are showing not only changes in the educational level of the supply of workers but changes in the educational level of the demand for them. Since the statistical mean approach has been used to measure the degree of education mismatch, this reduction in the proportion of over-educated workers from 2010 to 2011 suggests a possible education upgrading of the jobs or crowding-out. A similar tendency in the proportion of university graduates who are over-educated can also be found when all the university graduates are included in the sample (Appendix 2).

3.6.2. Probability for a university graduate to be over-educated in the Spanish labour market (1995-2018)

This section summarises the results of the maximum likelihood estimations for the probit and the Heckman two-step models for the sample excluding the three-year degree university graduates. The estimated coefficients for the sample including all university graduates can be found in Appendix 2.

Table 3.7 depicts the estimated coefficients for the probit model for the period 1995-2010, and Table 3.8 the ones for the period 2011-2018. For both periods, the estimated coefficients are positive for the variable *woman* indicating a higher probability for women to be over-educated than men. According to the data, these gender differences have reduced in the second period. In the first period (1995-2010) the estimated coefficient is 0.1301 whereas it reduces to 0.0797 in the second period (2011-2018). Some variations in the results appear when all the university graduates are considered (Appendix 2). In this latter case, the estimated coefficient is positive only for the period 1995-2010. Moreover, the coefficient is lower when the three-year degree university graduates are excluded. For the second period, results depicted in Appendix 2 show a lower probability for university graduate women to be over-educated than men.

Regarding age-group, results show differences among periods when the three-year degree university graduates are excluded (Tables 3.7 and 3.8). Data suggest a lower probability of being over-educated for the reference category 25-34 in the first period (1995-2010) and a higher likelihood for the second period (2011-2018). However, when all the university graduates are considered (Appendix 2), the probability of being over-educated is higher for 25-34-year-old university graduates in both periods. Furthermore, the differences in the probability of being over-educated between the reference category (25-34-year-old) and the other generations rises when the age of the individuals increases. Therefore, except for the first period with the sample excluding three-year degree university graduates (Table 3.7), the older the individuals, the lower the probability of being over-educated, meaning the higher the difference with the reference category.

Table 3. 7. Estimated coefficients for the probit model. Three-year degree university graduates excluded. Spain, 1995-2010

(Y=over-educated)

Independent variables (Reference category)		Coef.	Robust Std. Err.	Sig. Level
Woman (Man)		0.1301	*0.0110*	***
Age group (25-34)	**35-44**	0.0549	*0.0138*	***
	45-54	0.0342	*0.0158*	*
	55-64	0.0414	*0.0220*	n.s.
Not Spanish (Spanish)		0.4568	*0.0351*	***
Activity sector (Agriculture, forestry and fishing)	**Manufacturing**	-0.1823	*0.0665*	**
	Construction	-0.1717	*0.0723*	*
	Sales	0.2871	*0.0667*	***
	Information and communication	-0.5524	*0.0748*	***
	Financial and insurance activities	0.3564	*0.0686*	***
	Real estate activities	0.3584	*0.1000*	***
	Professional, scientific, technical	-0.5564	*0.0669*	***
	Public administration	-0.6874	*0.0647*	***
	Other services	-0.2646	*0.0676*	***
Region (Madrid)	**North-West**	0.1119	*0.0209*	***
	North-East	0.1439	*0.0190*	***
	Centre	0.1524	*0.0188*	***
	East	0.0297	*0.0186*	n.s.
	South	0.0815	*0.0188*	***
	Canary Islands	0.1173	*0.0277*	***
Year (1995)	**1996**	-0.0408	*0.0393*	n.s.
	1997	-0.0577	*0.0390*	n.s.
	1998	-0.0448	*0.0382*	n.s.
	1999	0.0209	*0.0373*	n.s.
	2000	0.0749	*0.0368*	*
	2001	0.0618	*0.0366*	n.s.
	2002	0.0698	*0.0365*	n.s.
	2003	0.0760	*0.0358*	*
	2004	0.0656	*0.0359*	n.s.
	2005	0.2207	*0.0365*	****
	2006	0.2047	*0.0353*	***
	2007	0.2364	*0.0347*	***
	2008	0.2470	*0.0350*	***
	2009	0.2396	*0.0349*	***
	2010	0.2551	*0.0348*	***
Constant		0.1164	*0.0717*	n.s.
Number of observations		**88,785**		
Pseudo R2		**0.0753**		
Wald Chi-squared (35)		**4834.19**	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Note: Individual sample weights considered.
Source: SLFS (2nd quarters), INE

These results might suggest either that the Spanish labour market is having more difficulties to accommodate the new university graduates entering the labour market or that, as the Career Mobility Theory states, over-education is a temporary phenomenon which especially affects those starting their career. However, the positive estimated coefficients for the dummy variables accounting for year and their upper tendency in both samples and periods suggests that the first option, the saturation of the Spanish labour market, seems to be the most plausible one.

As expected, Spaniards are less likely to be over-educated than those holding another nationality. Administration difficulties for foreigners to validate their official studies or legislation differences to practise their profession among countries are usually the main causes.

By regions, results show a lower probability of being over-educated for those living in the region around Madrid. As the administrative capital of the country, Madrid concentrates not only a big bulk of public administration, but also a huge number of the headquarters of multinational companies working in Spain. Both of them usually demand highly skilled workers.

‘Sales’ and ‘Real estate activities’ are the two activity sectors whose results show higher probability for university graduates to be over-educated than the reference category, the ones working in the ‘Agriculture, forestry, and fishing sector’. This is true for both periods when the three-year degree university graduates are excluded from the sample. Besides, for the first period, 1995-2010, those working in ‘Financial and insurance activities’ are also more likely to be over-educated than those working in the reference category sector. The higher participation of university graduates in these sectors can explain this effect.

Summarising, the analysis of the estimated coefficients of the probit model suggest higher probabilities of being over-educated for women. Besides, an upward tendency over time is found in the first and second period, and for the two samples (including and excluding three-year degree university graduates).

These results treat the sample as if data was randomly distributed. However, it is well known that only those individuals securing a job are susceptible to being mismatched. And it is likely that those individuals securing a job share a set of characteristics. Therefore, some sample selection might be expected to exist. If so, estimated coefficients from the probit model are likely to be over-estimated.

Table 3. 8. Estimated coefficients for the probit model. Three-year degree university graduates excluded. Spain, 2011-2018

(Y=over-educated)

Independent variables (Reference category)		Coef.	Robust Std. Err.	Sig. Level
Woman (Man)		0.0797	*0.0124*	***
Age group (25-34)	**35-44**	-0.0452	*0.0164*	**
	45-54	-0.1049	*0.0179*	***
	55-64	-0.1456	*0.0222*	***
Not Spanish (Spanish)		0.4016	*0.0309*	***
Activity sector (Agriculture, forestry and fishing)	**Manufacturing**	-0.4934	*0.0891*	***
	Construction	-0.4460	*0.0967*	***
	Sales	0.2161	*0.0892*	*
	Information and communication	-0.8303	*0.0931*	***
	Financial and insurance activities	-0.4462	*0.0907*	***
	Real estate activities	0.2584	*0.1177*	*
	Professional, scientific, technical	-0.4989	*0.0893*	***
	Public administration	-0.6839	*0.0876*	***
	Other services	-0.3943	*0.0905*	***
Region (Madrid)	**North-West**	0.1333	*0.0217*	***
	North-East	0.1044	*0.0221*	***
	Centre	0.0732	*0.0220*	**
	East	-0.0476	*0.0213*	*
	South	0.0531	*0.0222*	*
	Canary Islands	0.1735	*0.0335*	***
Year (2011)	**2012**	0.0390	*0.0267*	n.s.
	2013	0.0592	*0.0270*	*
	2014	0.0875	*0.0269*	**
	2015	0.1278	*0.0270*	***
	2016	0.1669	*0.0257*	***
	2017	0.1930	*0.0252*	***
	2018	0.2205	*0.0251*	***
Constant		0.2638	*0.0909*	**
Number of observations		**68,611**		
Pseudo R2		**0.0507**		
Wald Chi-squared (27)		**2662.82**	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Note: Individual sample weights considered.
Source: SLFS (2nd quarters), INE

In order to account for this possible sample selection, I use here a variation of the Heckman two-step model. Following Trevis Certo *et al.* (2016), Tables 3.9 and 3.10 display the estimated coefficients, the standard errors, and the significance levels of the probit model for the selection equations (Equation [3.2]) of the

Heckman two-step model for the first and second period, respectively. These tables aim at evaluating the role of the independent variables of the selection equation (first stage) to confirm the existence of sample selection. The statistical significance of the estimated coefficients of the covariates of these second equations suggest that sample selection is a potential problem. These results are also confirmed by the fact that only a subsample of the workers is over-educated (51 percent for the sample '1995-2010', and 47 percent for the sample '2011-2018'). The same coefficients significance can be found when the three-year degree university graduates are included in the sample. These latter results can be found in Appendix 2.

A set of dummy variables accounting for the relationship with the person of reference is included as exclusion variables in the selection equation. It is expected that being or not the person of reference of the family affects the probability of being employed, but not the likelihood of being over-educated. The survey does not specify who should be the person of reference, but it is expected that the person who will answer the survey will be someone with the highest influence in the daily functioning of the household, and therefore, that this will also determine their likelihood of being employed in the labour market. Besides, as there are not any determinants related to the characteristics of the person who will be identified as the person of reference (gender, age, marital status, educational level, etc.), it seems sensible to think that their probability of being over-educated will not be affected. Besides, the statistical postestimation test also validates these exclusion variables.

The estimated coefficients for these exclusion variables are negative and statistically significant showing that those individuals who are not the person of reference are less likely to be employed than the ones who have been classified in the survey as the person of reference of the household. Only those individuals living in the household and without any familiar relationship with the person of reference depict higher probabilities of being employed than the person of reference. The highly statistically significance of the post-estimation test (Chi-squared (8)= 8305.64 for the first period and Chi-squared (8)=2396.97 for the second period) and the relatively small standard errors of the estimation coefficients of both equations have confirmed its validity as exclusion variable.

Table 3. 9. Estimated coefficients for the selection equation of the two-step Heckman model. Three-year degree university graduates excluded. Spain, 1995-2010

(Y=employed)

Independent variables (Reference category)		Coef.	Robust Std. Err.	Sig. Level
Woman (Man)		-0.3379	*0.0121*	***
Age group (25-34)	**35-44**	0.1559	*0.0155*	***
	45-54	0.0986	*0.0193*	***
	55-64	-0.6750	*0.0212*	***
Not Spanish (Spanish)		-0.7830	*0.0273*	***
Person of reference	**Partner**	-0.2385	*0.0157*	***
	Children	-0.6597	*0.0164*	***
	Children-in-law	-0.4794	*0.0799*	***
	Grand-children	-1.0357	*0.0811*	***
	Parents	-0.3086	*0.1358*	***
	Other relatives	-0.3650	*0.0511*	***
	People working in the housework	6.4882	.	.
	Other people	0.0611	*0.0749*	n.s.
Region (Madrid)	**North-West**	-0.1973	*0.0200*	***
	North-East	-0.0803	*0.0192*	***
	Centre	-0.1778	*0.0184*	***
	East	0.0289	*0.0190*	n.s.
	South	-0.2352	*0.0183*	***
	Canary Islands	-0.1091	*0.0281*	***
Year (1995)	**1996**	0.0471	*0.0336*	n.s.
	1997	0.0495	*0.0331*	n.s.
	1998	0.0891	*0.0328*	**
	1999	0.1490	*0.0329*	***
	2000	0.2421	*0.0328*	***
	2001	0.3009	*0.0326*	***
	2002	0.2908	*0.0329*	***
	2003	0.2970	*0.0325*	***
	2004	0.3254	*0.0319*	***
	2005	0.3583	*0.0331*	***
	2006	0.3549	*0.0333*	***
	2007	0.4615	*0.0328*	***
	2008	0.4514	*0.0330*	***
	2009	0.3700	*0.0322*	***
	2010	0.3012	*0.0319*	***
Constant		1.1820	*0.0313*	***
Number of observations		**111,705**		
Pseudo R2		**0.0901**		

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Note: Individual sample weights considered.
Source: SLFS (2nd quarters), INE

Table 3. 10. Estimated coefficients for the selection equation of the two-step Heckman model. Three-year degree university graduates excluded. Spain, 2011-2018

(Y=employed)

Independent variables (Reference category)		Coef.	Robust Std. Err.	Sig. Level
Woman (Man)		-0.2640	*0.0128*	***
Age group (25-34)	**35-44**	0.0669	*0.0179*	***
	45-54	-0.0252	*0.0201*	n.s.
	55-64	-0.6093	*0.0219*	***
Not Spanish (Spanish)		-0.7134	*0.0247*	***
Person of reference	**Partner**	-0.0152	*0.0150*	n.s.
	Children	-0.7124	*0.0188*	***
	Children-in-law	-0.2765	*0.0861*	**
	Grand-children	-0.8054	*0.1489*	***
	Parents	-0.4345	*0.1117*	***
	Other relatives	-0.5223	*0.0543*	***
	People working in the housework	6.7571	*.*	.
	Other people	0.0214	*0.0819*	n.s.
Region (Madrid)	**North-West**	-0.1683	*0.0214*	***
	North-East	-0.0003	*0.0227*	n.s.
	Centre	-0.1435	*0.0216*	***
	East	-0.0524	*0.0213*	*
	South	-0.2696	*0.0216*	***
	Canary Islands	-0.2524	*0.032*	***
Year (2011)	**2012**	-0.0549	*0.0267*	*
	2013	-0.1314	*0.0266*	***
	2014	-0.1480	*0.0265*	***
	2015	-0.0825	*0.0261*	**
	2016	-0.0405	*0.0255*	n.s.
	2017	0.0206	*0.0253*	n.s.
	2018	0.0662	*0.0252*	**
Constant		1.3985	*0.0299*	***
Number of observations		**87,454**		
Pseudo R2		**0.0770**		

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Note: Individual sample weights considered.
Source: SLFS (2nd quarters), INE

The key findings of the Heckman two-step model performed are presented below (Table 3.11 and 3.12). Only for the second period (2011-2018), the estimated coefficients for *athrho* are statistically significant, as well as the Wald test of independence equations (*rho*). Therefore, at lest for the second period, the null hypothesis of no correlation between the error terms of the outcome [3.1] and the selection [3.2] equations should be rejected. This result suggests that the error terms of both equations of the Heckman two-step model are not independent. Overall, and even if the coefficients for *athrho* are statistically not significant for the first period, it seems sensible to apply the Heckman two-step model in order to obtain adjusted estimates, since the other verification test confirms it. When all university graduates are included (Appendix 2), the coefficients for *athrho* and the Wald test of independence equation (*rho*) are also statistically significant, as in the model using data for the second period (2011-2018) when the three-year degree university graduates are excluded.

Some differences appeared in the estimated coefficients between periods and samples when the Heckman two-step model is applied. For instance, estimated coefficients for *woman* are positive when the sample excluding three-year degree university graduate is used and negative when all university graduates are included in the sample (Appendix 2), even if for the latter results, the coefficients are statistically not significant for the first period (1995-2010). In other words, the results suggest a change in the gender differences in the probability of being over-educated depending on the sample used. In this sense, the results estimate a higher probability for a woman to be over-educated when only those with education levels equal or higher than five-year degree are included, whereas they are expected to be less likely to be more over-educated than men when all the university graduates are included. This result can be related to gender distribution among educational levels, as it has already been pointed out.

A positive coefficient indicates that, *caeteris paribus,* a university graduate woman employed in the Spanish labour market is more likely to be over-educated than a university graduate man. This is true when the three-year degree university graduates are excluded. In contrast, when all the university graduates are included in the calculations this sign becomes negative. These latter results indicate lower probabilities for university graduate women to be over-educated (Appendix 2).

Regarding age-group, different results appear for each of the periods. For the first period (1995-2010), the estimated coefficients suggest lower probabilities of being over-educated for the youngest generations (25-34-year-old), being the ones aged between 35 and 44 the most likely to be over-educated. For the second period (2011-2018), the estimated coefficients reveal higher probabilities of being over-educated for the younger generations (35-44-year-old more likely to be over-educated than those aged between 25-34). The oldest generation (55-64-year-old) is expected to be less likely over-educated (the estimated coefficient is negative and its value is the highest one). The same disparity in the sign of the estimated coefficients between periods was also found when the probit model were used (Tables 3.7 and 3.8). The robustness of these results may, therefore, show a change in the ability of the Spanish labour market for accommodating the youngest generations of university graduates. It may be related to the use of new information and communication technologies required in the workplaces.

The rest of the dependent variables of the outcome equation, those accounting for nationality, activity sector, region, and year have the same sign as the ones from the probit model. However, as expected, in most of the cases their value is lower since the likelihood of being employed is not considered by the probit model.

Therefore, the consistency of the results for both models (probit and the Heckman two-step) and for both samples (1995-2010 and 2011-2018) when three-year university graduates are excluded points to the robustness of the models. For the two models (probit and the Heckman two-step) and periods (1995-2010 and 2011-2018), the estimated coefficients are very similar. The sign for most of the explanatory variables is the same in both models and periods, and their values reveal few differences between models. The higher differences in the values of the estimated coefficients between the probit and the Heckman two-step model appear in the first period, 1995-2010. For instance, for the variable *woman*, the estimated coefficient for the first period (1995-2010) is 0.1301 for the probit model and 0.1183 for the Heckman two-step model (Table 3.7 and 3.11, respectively). For the second period (2011-2018), they are 0.0797 and 0.0514, respectively (Table 3.8 and 3.12, respectively).

Table 3. 11. Estimated coefficients for the outcome equation of the Heckman two-step model. Three-year degree university graduates excluded. Spain, 1995-2010

(Y=over-educated |employed)

Independent variables (Reference category)		Coef.	Robust Std. Err.	Sig. Level
Woman (Man)		0.1183	*0.0161*	***
Age group (25-34)	**35-44**	0.0658	*0.0180*	***
	45-54	0.0449	*0.0196*	*
	55-64	0.0289	*0.0245*	n.s.
Not Spanish (Spanish)		0.4353	*0.0415*	***
Activity sector (Agriculture, forestry, and fishing)	**Manufacturing**	-0.1812	*0.0664*	**
	Construction	-0.1714	*0.0722*	*
	Sales	0.2870	*0.0666*	***
	Information and communication	-0.5508	*0.0747*	***
	Financial and insurance activities	0.3568	*0.0685*	***
	Real estate activities	0.3582	*0.0999*	***
	Professional, scientific, technical	-0.5560	*0.0668*	***
	Public administration	-0.6853	*0.0646*	***
	Other services	-0.2631	*0.0676*	***
Region (Madrid)	**North-West**	0.1049	*0.0218*	***
	North-East	0.1410	*0.0191*	***
	Centre	0.1470	*0.0194*	***
	East	0.0307	*0.0186*	***
	South	0.0740	*0.0199*	***
	Canary Islands	0.1137	*0.0278*	***
Year (1995)	**1996**	-0.0390	*0.0394*	n.s
	1997	-0.0561	*0.0390*	n.s.
	1998	-0.0420	*0.0383*	n.s.
	1999	0.0251	*0.0375*	n.s.
	2000	0.0818	*0.0373*	*
	2001	0.0704	*0.0375*	n.s.
	2002	0.0783	*0.0373*	*
	2003	0.0848	*0.0367*	*
	2004	0.0754	*0.0370*	*
	2005	0.2316	*0.0378*	***
	2006	0.2156	*0.0368*	***
	2007	0.2503	*0.0370*	***
	2008	0.2607	*0.0372*	***
	2009	0.2513	*0.0365*	***
	2010	0.2651	*0.0360*	***
Constant		0.0857	*0.0778*	n.s.
Number of observations		**111,705**		
Selected observations		**88,785**		
Wald Chi-squared (35)		**4743.96**	***	
athrho		**0.0786**	**n.s.**	
rho		**0.0784**	**n.s.**	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Note: Individual sample weights considered.
Source: SLFS (2nd quarters), INE

Table 3. 12. Estimated coefficients for the outcome equation of the Heckman two-step model. Three-year degree university graduates excluded. Spain, 2011-2018

(Y=over-educated | employed)

Independent variables (Reference category)		Coef.	Robust Std. Err.	Sig. Level
Woman (Man)		0.0514	*0.0149*	**
Age group (25-34)	**35-44**	-0.0114	*0.0201*	n.s.
	45-54	-0.0754	*0.0209*	***
	55-64	-0.1836	*0.0231*	***
Not Spanish (Spanish)		0.3157	*0.0409*	***
Activity sector (Agriculture, forestry and fishing)	**Manufacturing**	-0.4834	*0.0880*	***
	Construction	-0.4391	*0.0955*	***
	Sales	0.2123	*0.0881*	*
	Information and communication	-0.8154	*0.0921*	***
	Financial and insurance activities	-0.4338	*0.0896*	***
	Real estate activities	0.2536	*0.1160*	*
	Professional, scientific, technical	-0.4902	*0.0882*	***
	Public administration	-0.6706	*0.0866*	***
	Other services	-0.3863	*0.0894*	***
Region (Madrid)	**North-West**	0.1095	*0.0225*	***
	North-East	0.1013	*0.0219*	***
	Centre	0.0547	*0.0223*	*
	East	-0.0517	*0.0211*	*
	South	0.0189	*0.0239*	n.s.
	Canary Islands	0.1360	*0.0350*	***
Year (2011)	**2012**	0.0329	*0.0265*	n.s.
	2013	0.0446	*0.0271*	n.s.
	2014	0.0702	*0.0272*	*
	2015	0.1163	*0.0272*	***
	2016	0.1581	*0.0257*	***
	2017	0.1904	*0.0250*	***
	2018	0.2217	*0.0248*	***
Constant		0.1847	*0.0932*	*
Number of observations		**87,454**		
Selected observations		**68,611**		
Wald Chi-squared (27)		**2326.13**	***	
athrho		**0.2916**	**	
rho		**0.2836**	**	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Note: Individual sample weights considered
Source: SLFS (2nd quarters), INE

In outline, the main findings of the estimation model performed are: first, university-graduate women, youth (less than 44 years old), foreigners, and those living out of the Madrid region are more likely to be over-educated; second, the probability of being over-educated increases over time. Since 1999 for the first period and during all the second period, the estimated coefficients for each single year has been positive, and its value increases year after year in most of the years of the first sample. In fact, the data reflect the effects of the economic crisis. According to previous studies, in the Spanish case, there is a positive relationship between the economic crisis and the proportion of over-educated workers (Fernandez, 2004). Tables from 3.7 to 3.12 show that the highest values for the estimated coefficients are in the period 2007-2010, the years of the economic crisis. These coefficients rose to values higher than 0.27 for the Heckman two-step model and higher than 0.23 for the probit model.

The results of the sample excluding the three-year degree university graduates are, in part, consistent with the main literature which usually includes population with different educational levels. The effect of gender, nationality, and age are in broad agreement with the general results of the literature (Aguilar Ramos and García Crespo, 2008; García-Montalvo, 1995; Murillo, Rahona and Salinas, 2010).

Some differences appeared when all the university graduates are included in the sample. For instance, the sign of the variable *woman* became negative, indicating a lower likelihood for women to be over-educated. This may be the consequence of their fewer years of education, that is the higher proportion of women among those holding a three-year university degree and lower proportion of women among those with higher educational levels: postgraduates and doctorates.

3.6.3. Predicted probabilities for a university-graduate to be over-educated in the Spanish labour market (1995-2018).

In order to see the changes in the probability for a 25-64-year-old university graduate to be over-educated in the Spanish labour market, Table 3.13 and Figures from 3.2 to 3.9 display the results of the predicted probabilities by time, age-group and nationality, activity sector, and region respectively. The same tables and figures with data for the samples including all university graduates can be found in Appendix 2.

Here, the predicted probabilities of success conditional on the selection equation have been calculated. This is, each of the predicted probabilities accounts for the probability for a 25-64-year-old university graduate to be employed and over-educated in the Spanish labour market, conditional on the probability for him/her to be already working as an employee.

Table 3.13 point out first, the higher probabilities for a university graduate to be over-educated in the Spanish labour market (more than 40 percent), and second, the increasing difficulties for university graduates to find a proper job in the Spanish labour market from 1995 to 2018. This result is in consonance with other studies analysing the effect of over-education on wages which had found an increase in the effect of over-education in the Spanish labour market (García-Montalvo, 1995; García-Montalvo, 2009; Murillo, Rahona and Salinas, 2010). Despite the higher predicted probabilities for university graduate women to be over-educated, the same general trend for these probabilities is followed by both genders.

As has already been stated, the lower predicted probabilities in the second period (2011-2018) for both samples (including and excluding three-year university graduates) suggests an upgrading in the years of education demanded by the labour market. The number of university graduates has been increasing. However, the baseline for the education mismatch has changed from 1995 to 2011, leading to an increase in the educational balance in the labour market. This is only possible if there is an upward trend in the number of years of education demanded by the employers.

As expected, the results are lower when the three-year degree university graduates are included in the sample (Appendix 2), since those individuals holding lower years of education are included in the sample. In this latter case, the likelihood of being over-educated reduces around 4 percentage points compared with the sample excluding the three-year degree university graduates.

Table 3. 13. Predicted probabilities for a 25-64-year-old university-graduate being over-educated by year. Three-year degree university graduates excluded. Spain, 1995-2018

	Men			Women		
	Pre. Pr.	Std. Err.		Pre. Pr.	Std. Err.	
1995	0.4436	*0.0105*	***	0.4915	*0.0108*	***
1996	0.4806	*0.0142*	***	0.5274	*0.0142*	***
1997	0.4746	*0.0140*	***	0.5214	*0.0141*	***
1998	0.4793	*0.0137*	***	0.5260	*0.0137*	***
1999	0.5024	*0.0134*	***	0.5489	*0.0133*	***
2000	0.5214	*0.0132*	***	0.5675	*0.0130*	***
2001	0.5169	*0.0131*	***	0.5630	*0.0129*	***
2002	0.5195	*0.0130*	***	0.5655	*0.0128*	***
2003	0.5217	*0.0127*	***	0.5676	*0.0125*	***
2004	0.5181	*0.0127*	***	0.5641	*0.0124*	***
2005	0.5700	*0.0127*	***	0.6150	*0.0122*	***
2006	0.5646	*0.0123*	***	0.6098	*0.0118*	***
2007	0.5751	*0.0120*	***	0.6198	*0.0115*	***
2008	0.5783	*0.0121*	***	0.6229	*0.0115*	***
2009	0.5756	*0.0120*	***	0.6204	*0.0115*	***
2010	0.5804	*0.0120*	***	0.6253	*0.0114*	***
2011	0.4105	*0.0073*	***	0.4403	*0.0073*	***
2012	0.4669	*0.0096*	***	0.4980	*0.0094*	***
2013	0.4736	*0.0098*	***	0.5053	*0.0096*	***
2014	0.4827	*0.0097*	***	0.5146	*0.0095*	***
2015	0.4956	*0.0098*	***	0.5270	*0.0095*	***
2016	0.5078	*0.0092*	***	0.5389	*0.0090*	***
2017	0.5160	*0.0090*	***	0.5467	*0.0087*	***
2018	0.5243	*0.0089*	***	0.5546	*0.0086*	***

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: SLFS (2nd quarters), INE

Furthermore, the positive relationship between the economic crisis and the probability of being over-educated, anticipated by the estimated coefficients, is also confirmed by these predicted probabilities, for both periods and samples.

Overall, the predicted probabilities over time point out the difficulties for the Spanish labour market to accommodate university graduates in proper jobs. The following figures depicts the differences in these predicted probabilities among the categories of the variables included in the model.

Figure 3.3 and 3.4 show the predicted probabilities for a 25-64-year-old university graduate to be over-educated conditional on the fact that s/he is already working as a employee by age group, nationality and gender. Figures 3.5 and 3.6 present the predicted probabilities by activity sector and gender. And Figures 3.7 and 3.8 depict the predicted probabilities by region (NUTS-1) and gender. Figures representing the predicted probabilities when the sample including all university graduates is used can be found in Appendix 2.

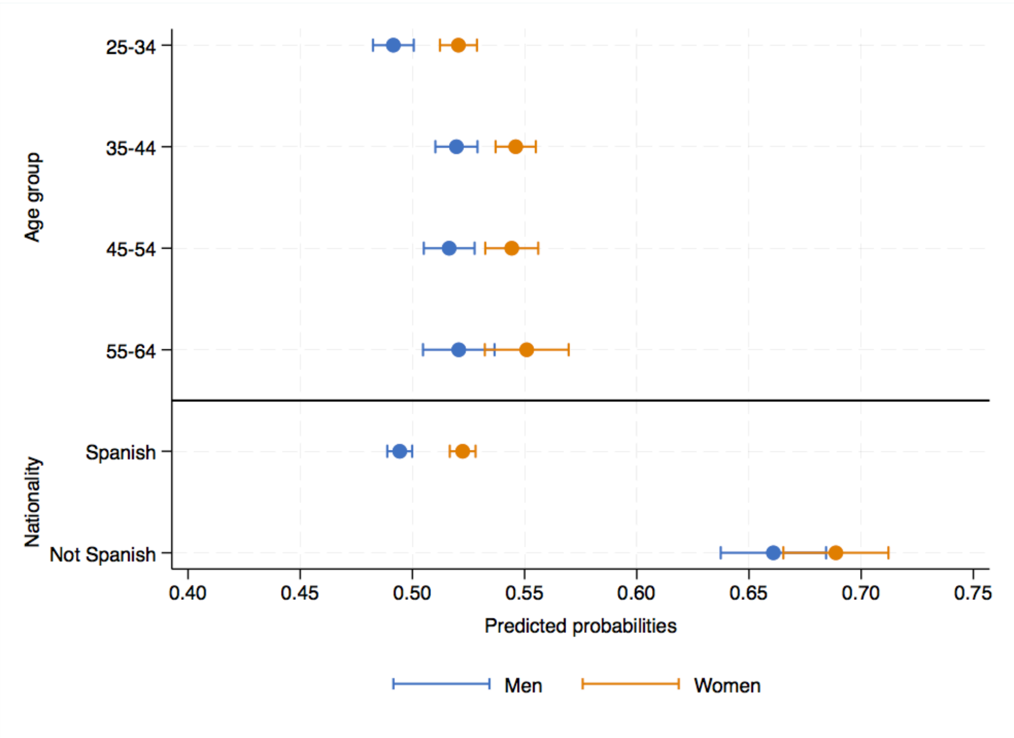
For all age-groups and nationalities, the probability of being over-educated is higher for women than for men when the sample excluding the three-year degree university graduates is used (Fig. 3.3 and 3.4), whereas the other way round is true when all university graduates are considered (Appendix 2). This result just confirms what has already been stated in the analysis of the previous sections of this chapter. Including or not including the more professional degrees which disappeared by the end of the 90s does make a difference in the results and in the likelihood of being over-educated. The reason: the fact that over-education has been measured using the years of education needed for the job and the fact that these more professional degrees required three years of education instead of five.

Furthermore, more differences are found between periods and between categories. Focusing on the sample excluding the three-year university graduates (Fig. 3.3 and 3.4), three main features can be extracted from the figures. First, in the first period (Fig. 3.3) younger generations seem to be less likely to be over-educated than the oldest ones. This result may be showing the increase in the supply of university graduates in the 90s as a consequence of the increase in the supply of and demand for university studies showed in the previous chapter as well as in the earlier sections of this chapter.

Second, this initial trend changes in the second period (Fig. 3.4) where the probability of being over-educated seems to decrease when the age of the individuals increase. These latter results might be either supporting the Career Mobility Theory which states that over-education is a temporary phenomenon that mainly affect younger generations when entering the labour market, or suggesting the hypothesis of this chapter that the Spanish labour market has not been able to accommodate the increasing number of university graduates coming out from the Spanish universities in the last decades. However, given the positive relationship between being over-educated and time, the latter hypothesis seems to be more likely.

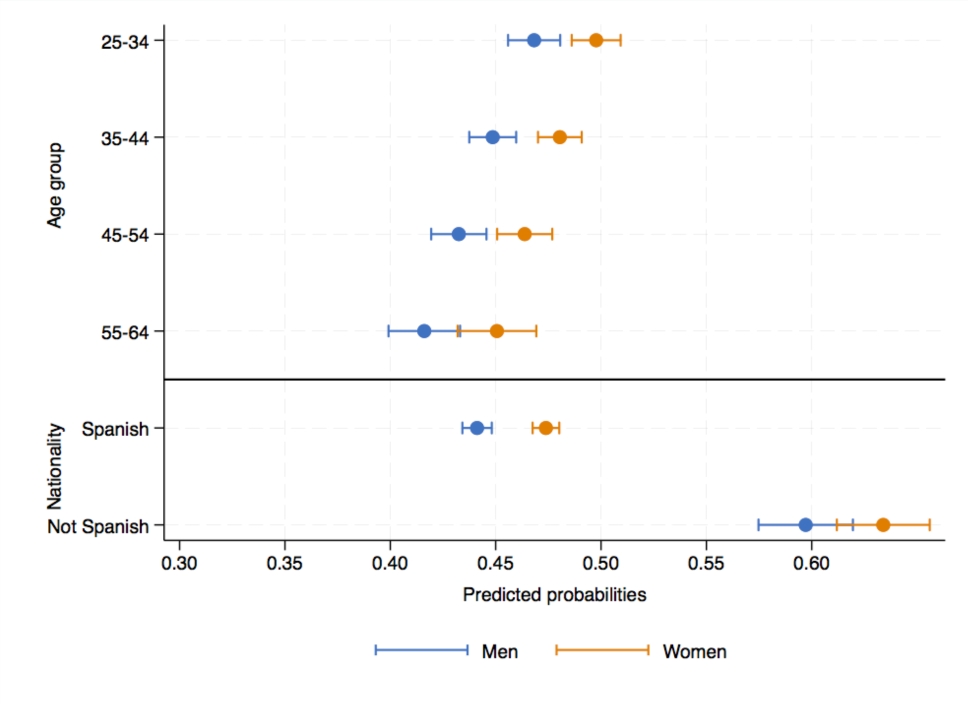
Third, the variability in the predicted probabilities increases with the age of the university graduates, being therefore higher for the oldest generations.

By nationality, all models and periods depict lower probabilities of being over-educated for those university graduates holding a Spanish nationality. Moreover the variability in the predicted probabilities is also higher among those holding a nationality different from the Spanish one. The negative effect that holding a degree from a foreign university can have on the likelihood of being employed in a highly skilled job has already been treated in this chapter. It is especially significant of those degrees related to medicine, law or other fields of study which professional development imply different legislations or ways of working across countries.



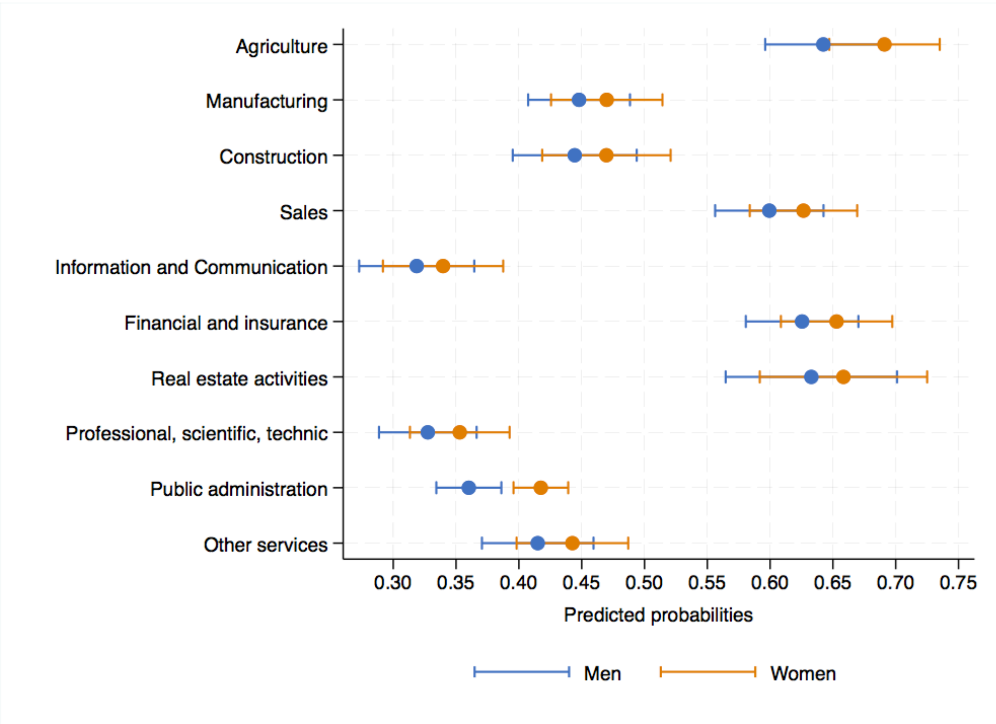
Source: SLFS (2nd quarters), INE

Figure 3. 3. Predicted probabilities for a university graduate to be employed and over-educated by age-group, nationality, and gender. Three-year degree university graduates excluded. Spain, 1995-2010



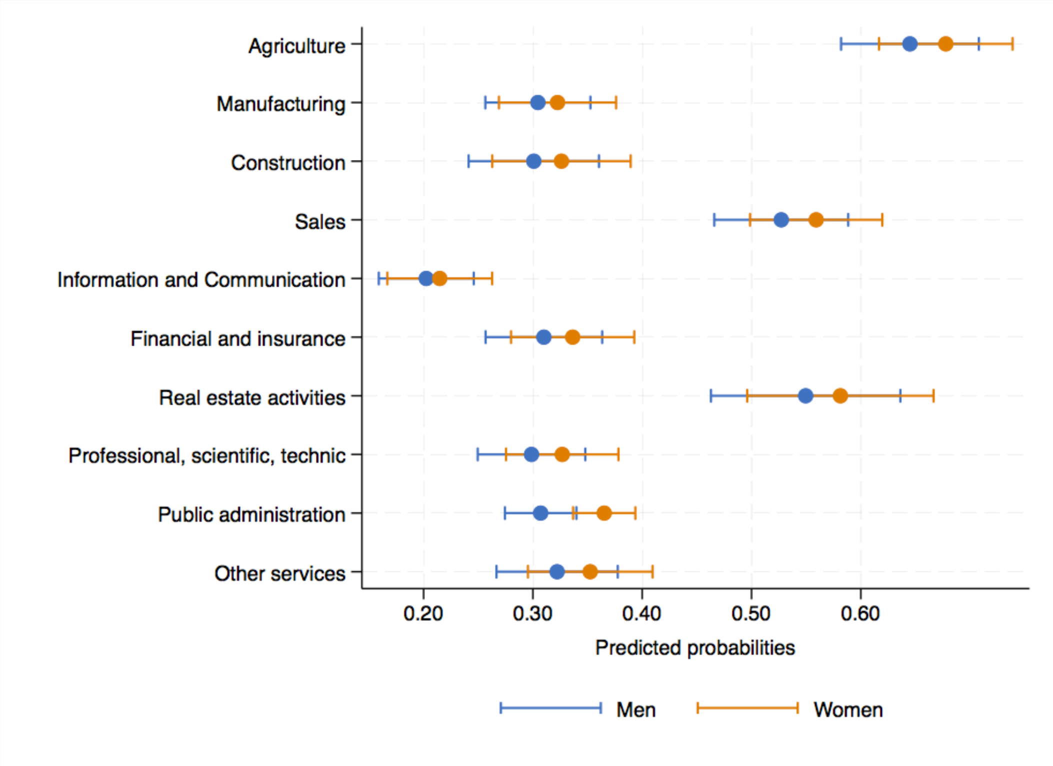
Source: SLFS (2nd quarters), INE

Figure 3. 4. Predicted probabilities for a university graduate to be employed and over-educated by age-group, nationality, and gender. Three-year degree university graduates excluded. Spain, 2011-2018



Source: SLFS (2nd quarters), INE

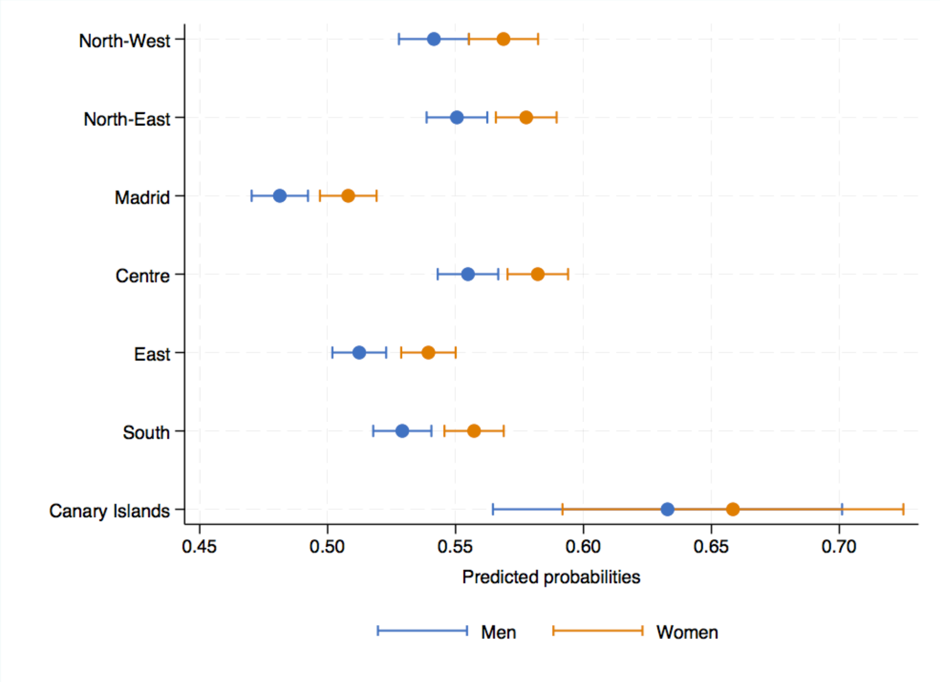
Figure 3. 5. Predicted probabilities for a 25-64-year-old university graduate to be employed and over-educated by activity sector and gender. Three-year degree university graduates excluded. Spain, 1995-2010



Source: SLFS (2nd quarters), INE

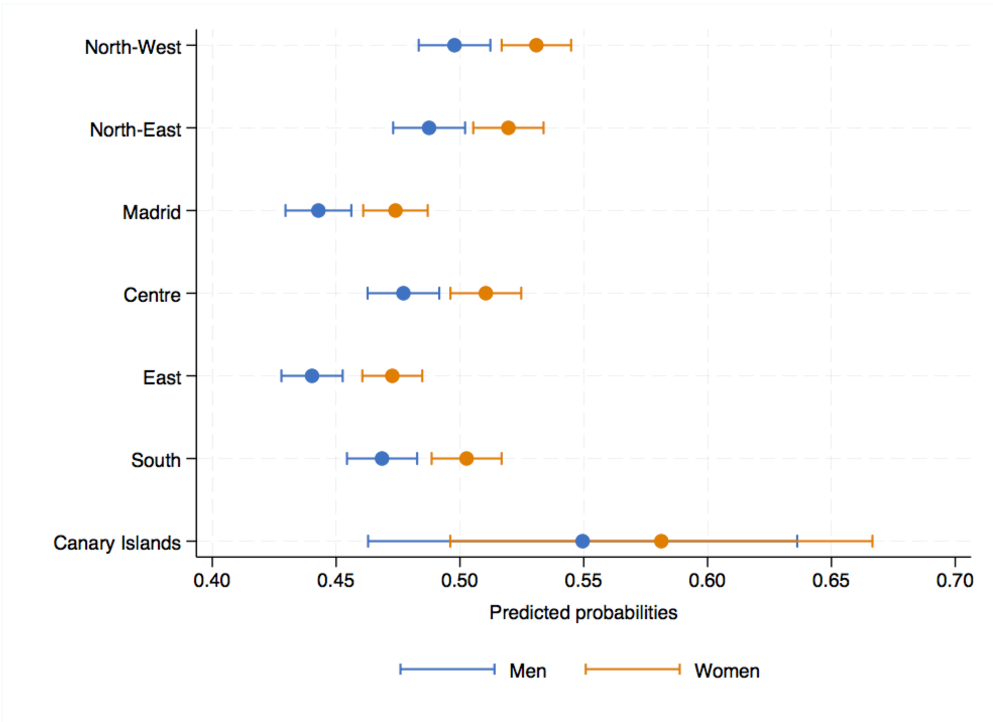
Figure 3. 6. Predicted probabilities for a 25-64-year-old university graduate to be employed and over-educated by activity sector and gender. Three-year degree university graduates excluded. Spain, 2011-2018

By activity sector, the predicted probabilities depict higher variation within categories for each one, being even higher when the predicted probabilities have higher values, as for instance, those employees working in the 'Agriculture, forestry, and fishing', 'Sales activities', or 'Real estate activities'. This is especially true for the sample excluding the three-year degree university graduates (Fig. 3.5 and 3.6). This variability is a consequence of the different demand for high skills from the occupations included in each of the activity sectors. Furthermore, those university graduates working on the most technological sector 'Information and communication' are the ones with lower probabilities of being over-educated, this is true for both periods and both samples.



Source: SLFS (2nd quarters), INE

Figure 3. 7. Predicted probabilities for a 25-64-year-old university graduate to be employed and over-educated by region and gender. Three-year degree university graduates excluded. Spain, 1995-2010.



Source: SLFS (2nd quarters), INE

Figure 3. 8. Predicted probabilities for a 25-64-year-old university graduate to be employed and over-educated by region and gender. Three-year degree university graduates excluded. Spain, 2011-2018.

By regions, slight differences appear among periods when the three-year degree are excluded from the sample (Fig. 3.7 and 3.8). Higher differences appear in the second period when the sample includes all the university graduates (Appendix 2).

University graduates living in the area of Madrid are less likely to be over-educated especially during the first period (Fig. 3.7). In the second period, the differences in the predicted probabilities among regions reduces, especially between the East region and the area of Madrid (Fig. 3.8). Predicted probabilities for university graduates to be over-educated are highly variable in the Canary Islands, they can vary more than 15 percentage points in both periods.

In short, the econometric analysis of the data has demonstrated the decreasing ability of the Spanish labour market to accommodate the university graduates from 1995 to 2018. Besides, the partition of the data series in both periods because of the occupational classification changes has made the identification of the upgrading on the requirements of the Spanish labour market easier. Finally, the estimated coefficients and the predicted probabilities have demonstrated that over-educated university-graduates share a set of characteristics (being young or foreign and not working in 'Information and communication activities'), and these characteristics are shared by both genders.

3.7. Has the Spanish labour market been balanced?

Focusing on university graduates, this chapter accounts for the probability of university graduate employees not being accommodated in a 'highly skilled job' by the Spanish labour market. Assuming a certain homogeneity among university graduates and occupations, I used statistical methods to identify the over-educated university graduates, assessing as over-educated those individuals who have been in school for more years than those required by the job.

Given the data, the second quarters of the Spanish Labour Force Survey, statistical methods have been considered to be the best option to classify the individuals into under-educated, adequately educated, and over-educated. The data do not contain subjective information and the use of objective methods needs a detailed description of the requirements of the occupations that does not exist for the Spanish labour market. Therefore, the mean of the years of education for each occupational category plus/minus one standard deviation is the upper- and lower-limit to determine the level of qualification mismatch. Using this educational mismatch classification, the results show lower differences in the qualification mismatch between genders.

Furthermore, an econometric analysis is run to assess the changes on the probability of being over-educated in the Spanish labour market over the period 1995-2018. Different variables take into account personal and labour market characteristics in addition to region (NUTS-1) and time period. The characteristics of the data and the dependent variable (dichotomous) suggest using a variation of the Heckman two-step model. The censored probit model accounts for a possible sample selection, which has been confirmed by the analysis of the statistical significance of the estimated coefficients of the select equation in the first-step and the significant level for the value of *athrho* and *rho*.

In order to be able to compare the results with a baseline model, a probit model has also been estimated using the same sample and the same covariates as the Heckman two-step model. The affinity of the results between the models run confirms their robustness. Most of the estimated coefficients are highly statistically significant and coincide in terms of sign, although they vary slightly in terms of magnitude.

All in all, the results are in accordance with the previous literature analysing all the educational levels and previous periods. They confirm that between 1995 and 2018 the Spanish university graduates had more difficulties to find a job commensurate with their educational level. This indicates a saturation of the Spanish labour market. Furthermore, despite this general upward trend, a positive relationship between economic crisis and increasing probability of being over-educated has also been detected.

The predicted probabilities for the rest of the covariates are also in broad agreement with the earlier literature. These results confirm the initial hypothesis that those over-educated university graduates share a set of characteristics. In particular, they are more likely to be over-educated when they are young and foreign, whereas those living in the area of Madrid and working in 'Information and communication activities' are less likely to be over-educated.

In contrast to other studies, the analysis of these data suggests that when the three-year degree university graduates are included, university graduate women are less likely to be over-educated than men. This can be explained by a higher representation of women among the three-year university degree, whereas men are more represented among postgraduate studies and doctorates. Three-year degrees require fewer years of formal education and postgraduates and doctors need longer years to get their degree.

All in all, these results should be treated with some caution since some homogeneity among individuals and occupations has been assumed. As a consequence, some of these results can reflect unobserved heterogeneity among university graduates or among occupations considered in the same group of occupations in the two-digit classification, as for example skill heterogeneity.

"Differences in the skills young persons obtain from their home and social environment also produce differences in workplace competencies. Young persons with nominally the same schooling score differently on standardized international test of attainment."

Freeman and Schettkat (2001, p. 585)

Chapter 4. To what extent does skill heterogeneity explain education mismatch for university graduates in the Spanish labour market?

This chapter analyses the effects that accounting for cognitive skill heterogeneity has on the probabilities of being over-educated in the Spanish labour market. Using OECD data from the Survey of Adults Skills (PIAAC), I estimate a pairwise comparison of the variation in individuals' skills across countries. The Heckman two-step model is then used to measure to what extent these variations in skills contribute to changing the probability for a university graduate to be over-educated in the Spanish labour market.

4.1. Introduction

Individuals are not homogeneous and neither are their skills and abilities. Their capabilities to learn and to develop new skills are, of course, related to the potential of their innate abilities and how well they can incorporate the new learning within their existing knowledge. Besides, individuals' skills and abilities also depend on the quality and quantity of training or experiences they have received. Hence, training and learning new concepts and ways of working are the best inputs to improve individual's skills.

Education and training are the main instruments to develop an individual's skills. In fact, and as described in the literature review chapter and in previous chapters, education and training, and especially the level of formal education, have been traditionally used as indicators to measure an individual's skill level in studies about education mismatch in the labour market (Aguilar Ramos and García Crespo, 2008; Alba-Ramírez and Blázquez, 2003; Battu, Belfield and Sloane, 1999; Chevalier, 2003; Chevalier and Lindley, 2007; Dolton and Vignoles, 2000; Murillo, Rahona and Salinas, 2010). This is the case, indeed, of the previous chapter (Chapter Three) where the statistical mean approach was used to measure education mismatch.

However, as is well known from the literature, an individual's educational level does not always define an individual's skill level (Alba-Ramírez and Blázquez, 2003; Badillo-Amador, García-Sánchez and Vila, 2005; Green, 2013; Green, McIntosh and Vignoles, 1999; Halaby, 1994; Hanushek and Woessmann, 2015). Apart from the differences among the individuals' potential abilities, different factors may influence the learning process or the type of skills acquired. For instance, the quality of teaching, often related to the quality of the teaching institution, especially in those countries with higher gaps between public and private schools, high schools, colleges or universities might influence the abilities acquired by the students. This is especially important in university studies and in those areas of study where higher economic resources are needed to develop the teaching and learning process, for instance, science or medicine. As regards tertiary education and the type of skills acquired in this education level, the field of study may be considered as a signal of the main type of skills taught in the degree (Montt, 2015).

The *objective* of this fourth chapter is twofold. It first aims to measure *the variation in skills among five European countries: England, France, Italy, Sweden, and Spain*. Second, and focusing on the Spanish labour market, the chapter *measures to what extent accounting for skill heterogeneity affects the likelihood of being over-educated*, and whether this probability varies by age, gender or region. In particular, I use the literacy and numeracy PIAAC scores as a measure of the individuals' level of the cognitive skills.

Using data from the Survey of Adults Skills (PIAAC) carried out by the OECD in 2011-2012, I first do a pairwise comparison to evaluate the variation in numeracy and literacy skills in five European countries to identify whether there is skill variation across countries.

This chapter aims at evaluating the effect that accounting for observable cognitive skills, literacy or numeracy, has on the probability of being over-educated. This previous analysis is useful to detect whether there is a variation in cognitive skills among individuals. In the case that the results indicate that there is no variation in literacy and numeracy skills, it will not make sense to use these observable cognitive skills to evaluate whether accounting for them affects the probability of being over-educated, because they will not be good measurements for skills heterogeneity among individuals.

Besides, the pairwise comparison of other European countries is useful to locate Spain in the European context and improve the understanding of the results. According to the classification of Esping-Andersen and Regini (1999), it is expected that Southern countries share some characteristics, since their labour market structures are similar. Indeed, the Italian and Spanish labour markets are defined as more regulated than the ones of France or England which are considered to be more flexible. In between, comes Sweden, whose labour market is considered to give security but also the necessary flexibility.

The results of the pairwise comparison will not be useful to conclude what part of the variation in skills among countries is due to their different labour market policies, but taking into account the structure and functioning of the labour market can help to interpret the results of the pairwise comparisons.

Using also PIAAC data but focusing on the Spanish labour market, I applied the Heckman two-step model to estimate the probability for an individual to be employed and over-educated in the Spanish labour market accounting for their skills level. As in the previous chapter, the idea behind this is that over-educated individuals are not randomly distributed and they share a set of characteristics. Therefore, in order to get unbiased estimated coefficients, the Heckman two-step model is applied here to account for this sample selection.

Furthermore, as in Chapter Three, a probit model using the same variables as the Heckman two-step model has also been used as a baseline model to see whether when the probability of being employed is not considered (probit model), the likelihood of being over-educated is higher than when a possible bias selection is taken into account (the Heckman two-step model). As is well established in the literature, and has been seen in Chapter Three, the probit model tends to overestimate the estimated coefficients when there is a sample selection. Indeed, in this case, results would only be equal if there were not unemployment in the labour market, that is, if all the individuals were working, and therefore the probability of being employed was equal to 1. Unfortunately, this is never the case, and consequently, it is expected to find a lower likelihood of being over-educated when the probability of being employed is considered (the Heckman two-step model).

The analysis aims at measuring whether accounting for cognitive skills reduces the likelihood for an individual to be over-educated, since it is expected that skills level has a negative effect on the likelihood of being over-educated. According to the Job Competition Model (Thurow, 1970), and the Assignment Theory (Tinbergen, 1956; Sattinger, 1980; Hartog, 1981), it is expected that the highest skilled university graduates are more likely to secure highly skilled jobs, and therefore, to be matched in the labour market (Pryor and Schaffer, 1999). As explained in the literature review chapter the Job Competition Model defends the idea that the most valuable workers, meaning the more skilled workers, will secure the higher-skilled jobs, whereas the less-skilled workers will be employed in highly skilled jobs only if there is enough supply of these sorts of jobs. Otherwise, they will be employed in less-skilled jobs. The analysis of the previous chapters suggests that the Spanish labour market has not been able to offer enough highly skilled jobs.

Similarly, Assignment Theory states that the best match is to find an equilibrium between the supply of and demand for skilled workers. Only if there is a skill balance will the individual, the firm, and therefore, the economy and the society benefit from it. In this sense, in order to find this equilibrium, the more skilled workers have to be less over-educated, and therefore, a negative relationship between higher skills and the probability of being over-educated should be found.

The rest of the chapter is organised as follows. The source and the sample used are detailed in the next section (4.2). The empirical framework and the methodology are explained in Section 4.3, and the results of the calculations are depicted in Section 4.4. The main contributions of the chapter are presented in the last section (4.5).

4.2. The source: The Survey of Adult Skills (PIAAC)

The Survey of Adult Skills has been boosted inside the Programme of the International Assessment of Adults Competences (PIAAC) to assess adults' proficiency in some key skills. It focuses on skills highly demanded by the current economies and societies: literacy, numeracy, problem solving in technology-rich environments and reading components.

The PIAAC was carried out in 24 countries: 22 OECD member countries and two partner countries, Cyprus and the Russian Federation (Jimeno, Lacuesta and Villanueva, 2013; OECD, 2013c). It includes information related to: (a) basic demographic characteristics and background of respondents, (b) educational attainment and participation in learning activities, (c) labour-force status, work history and job characteristics, (d) social participation and health, and (e) literacy and numeracy practices and the use of skills (Hernández Lahiguera and Serrano Martínez, 2013; OECD, 2013c).

The population of the survey are individuals from 16 to 65 years old, and the survey was carried out in two main stages. The individuals answered first the background questionnaire, and later made the cognitive assessment. The background questionnaire includes information related to basic demographic characteristics, background, educational attainment, labour force status and social participation (from (a) to (d)). The cognitive assessment evaluates the use of skills in literacy, numeracy, problem solving in technology-rich environments and reading components. However, the problem solving in technology-rich environments questionnaire was not evaluated in Cyprus, France, Italy and Spain (OECD, 2013c, p. 51).

The Survey of Adults Skills did not use the same questionnaire to assess all the individuals. Literacy, numeracy and problem solving in technology-rich environments questionnaires were randomly distributed by respondents who answered the test in two different stages. From these responses and from the background questionnaire more plausible values using Item Response Theory (IRT) were derived to have a greater database. [88]

Table 4.1 summarises the main characteristics of the PIAAC data and of the sample used in this chapter for Spain, and Table 4.2 depicts the same data for each of the other four countries considered in this chapter (England, France, Italy, and Sweden). The countries chosen here are the same as those used in Chapters Two and Three to contextualise the Spanish labour market into the European context. Following the classification of Esping-Andersen and Regini (1999); Vaughan-Whitehead (2011) regarding the characteristics and structure of their labour markets apart from Spain, I have included a country to represent each of their categories: England for the Anglo-Saxon countries, France for Continental Europe, Italy for Southern Europe, and Sweden for the Scandinavian countries.

---

[88] A more detailed description of the design and the evaluation of the PIAAC assessment can be found in the reader's companion edited by OECD (OECD, 2013c).

Table 4. 1. Main characteristics of the Spanish sample

PIAAC Sample	**Sample frame**	The Population registry, 2011
	People included	aged 16-65 years, residing in the country at the time of data collection
	People excluded	None
	Data collection	1st Sept 2011 – 1st May 2012
	Language of assessment	Spanish, Basque, Catalan, Galician, Valencian
	Literacy and numeracy	Yes
	Problem solving	No
	Reading components	Yes
	Total observations	6,055
	Men	49 %
	Women	51 %
Sample Chap. 4	**Selected population**	25-64-year-old population
	Observations	3,835
	Men	47 %
	Women	53 %

Sources: PIAAC http://www.oecd.org/skills/piaac/Technical%20Report_Part%204.pdf and http://www.mecd.gob.es/dctm/inee/internacional/piaac/piaac2013vol1.pdf?documentId=0901e72b81741bbc

For instance, in the Spanish case (Table 4.1), the sample frame was the population registry of 2011, and the target population, people *"aged 16-65 years, residing in the country at the time of data collection"* (Hernández Lahiguera and Serrano Martínez, 2013, p. 51). The total Spanish PIAAC sample contains 6,055 observations (49 percent men and 51 percent women). In this chapter, a sample including the 25-64-year-old population has been used. Therefore, the sample used contained 3,835 observations (47 percent men and 53 percent women). Unlike the previous chapters, I use here the entire 25-64-year-old population because when the sample is restricted to only 25-64-year-old university graduates the model does not converge due to the fact that there are not enough data in each category of the independent variables. However, a variable accounting for education level has been included to see the difference among groups and to be able to calculate the predicted probabilities for university graduates to be over-educated.

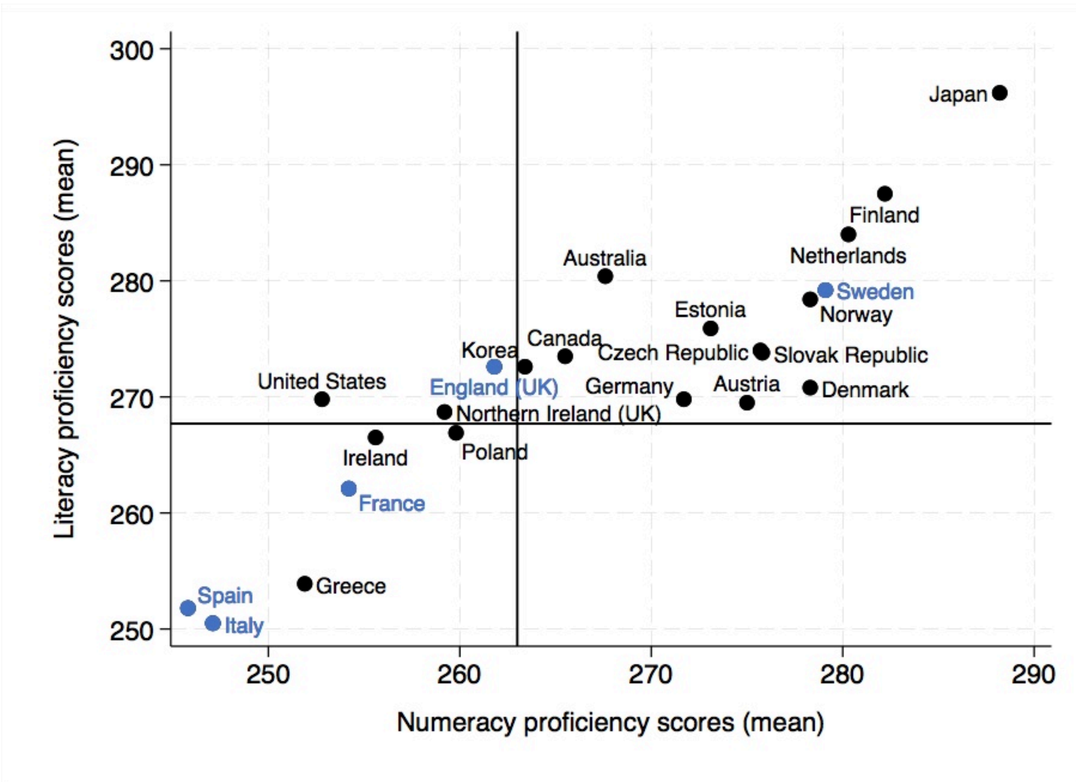
Table 4. 2. Main characteristics of the sample by country

Country	**France**	**Italy**	**Sweden**	**England**
Sample frame	Master sample from census data file, 2010	Household registries held by municipalities, 2011	Population registry 2011	The filtered version of the 2011 Postcode Address File
People included	aged 16-65 years, residing in the country at the time of data collection			aged 16-65 years, residing in private households in the country at the time of data collection
People excluded	- Young adults who have never claimed any income and are not attached to their parents' households - Some illegal migrants	- Adults in non-institutional group quarters - Illegal migrants	Undocumented immigrants	- Individuals living in private residences that are not listed on the address database.
Data collection	Sept – Nov 2012	End 2011 –2012	Aug. 2011 – June 2012	Aug 2011 – April 2012
Language of assessment	French	Italian	Swedish	English
Literacy and numeracy	Yes	Yes	Yes	Yes
Problem solving	No	No	Yes	Yes
Reading components	No	Yes	Yes	Yes
Total observations	6,993	4,621	4,469	5,131
Men	49 %	48.4 %	50.4 %	42.8 %
Women	51 %	51.6 %	49.6 %	57.2 %

Sources: PIAAC (http://www.oecd.org/skills/piaac/Technical%20Report_Part%204.pdf); ***France (****http://www.oecd.org/skills/piaac/Country%20 note%20-%20France%20(FR).pdf);* ***Italy*** *(http://bw5.cineca.it/bw5ne2/ShowFile.aspx?FileName=UDcknu5M1mzeShSPemUn1M7WwVEmY1 ujsBVc+X6DIi0cRJArVkUArVIOaRhA+OuZ&SID=INAPDEFAULTe3ctfkhpes3hasgq4eopdmmh);* ***Sweden*** *(https://www.oecd.org/skills/piaac/ Country%20note%20-%20Sweden.pdf);* ***England*** *(http://www.oecd.org/skills/piaac/Country%20note%20-%20United%20Kingdom.pdf)*

In Table 4.2 the main characteristics of the samples for the rest of the countries (France, Italy, Sweden, and England) are presented. The table shows how, for instance, Italy used as a sample frame for the PIAAC survey the household registries held by the municipalities of 2011, excluding adults in non-institutional group quarters and illegal migrants. All in all, the Italian PIAAC sample is the smallest sample and it contains 4,621 observations (48.4 percent men and 51.6 percent women). Table 4.2 shows how the distribution of the data is similar for the other countries considered.

As mentioned before, PIAAC measures individuals' cognitive skills. In particular, numeracy and literacy skills are assessed. The score results for literacy and numeracy skills are based on a scale of 500 points. Including all the PIAAC data, Figure 4.1 depicts the distribution of the mean of the numeracy and literacy PIAAC scores by country.



Source: PIAAC and OECD (2016)

Figure 4. 1. Mean literacy and numeracy PIAAC scores by country.

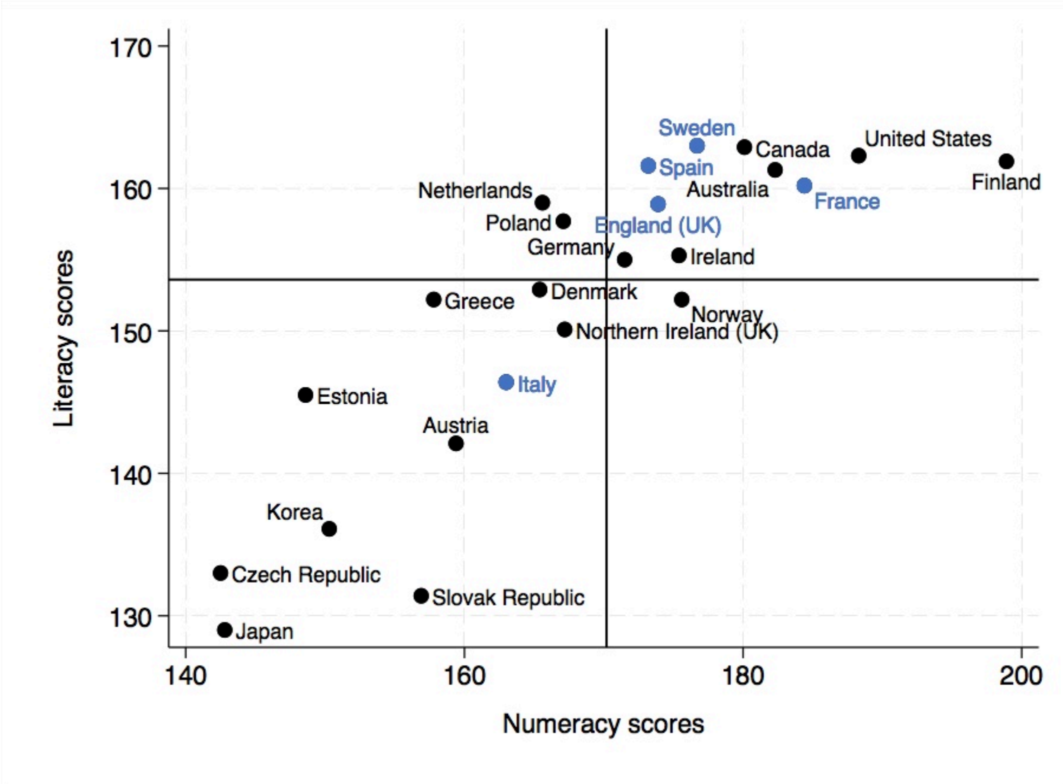
In the Spanish case, the mean of the distribution of the proficiency scores are around 250 score points in both skills: 252 score points in literacy (the second lowest average score, only exceded by Italy) and 246 in numeracy (the lowest average score). The average level within the participating OECD countries is 273 score points for literacy and 269 score points for numeracy (2nd proficiency level out of 5);[89] whereas the highest scores levels are for Japan (296 points in literacy skills and 288 in numeracy skills, this is the 3rd proficiency level in both skills). In fact, the numeracy scores for the four countries considered here are lower than the OECD average, and only people from England obtained on average higher literacy (but not numeracy) scores than the OECD average (OECD, 2016) (Fig. 4.1).

According to Figure 4.1, three of the countries considered (France, Italy and Spain) are countries with lower skilled population. The French scored around 10 PIAAC points less than the OECD average (both in literacy and in numeracy), and the Spanish and the Italians scored around 15 PIAAC proficiency points lower than the OECD average. English workers scored higher than the PIAAC proficiency points, although this was only for literacy, and Sweden is the only country of the ones included in this chapter whose workers scored higher than the PIAAC proficiency points in both cognitive skills, literacy and numeracy.

The dispersion between the highest and the lowest score value is even more significant than the average score (Fig. 4.2). According to OECD data, *"on average, 152 score points separate the highest and lowest 5 percent performers in literacy"* (OECD, 2013a, p. 73) (184.7 points for those in the 5th percentile, the 1st proficiency level; and 338.3 points for those in the 95th percentile, the 4th proficiency level) and Spain is one of the *"countries with comparatively large variations in scores"* (OECD, 2013a, p. 73) with a range of 162 points (163.5 points for performers in the 5th percentile, below the 1st proficiency level; and

[89] PIAAC literacy and numeracy scores are divided into six levels. The score-point ranges and the main tasks required for each level can be found in OECD (2013c, pp. 69-70).

325.1 points for performers in the 95th percentile, the 3rd proficiency level) (OECD, 2013a; OECD, 2016).

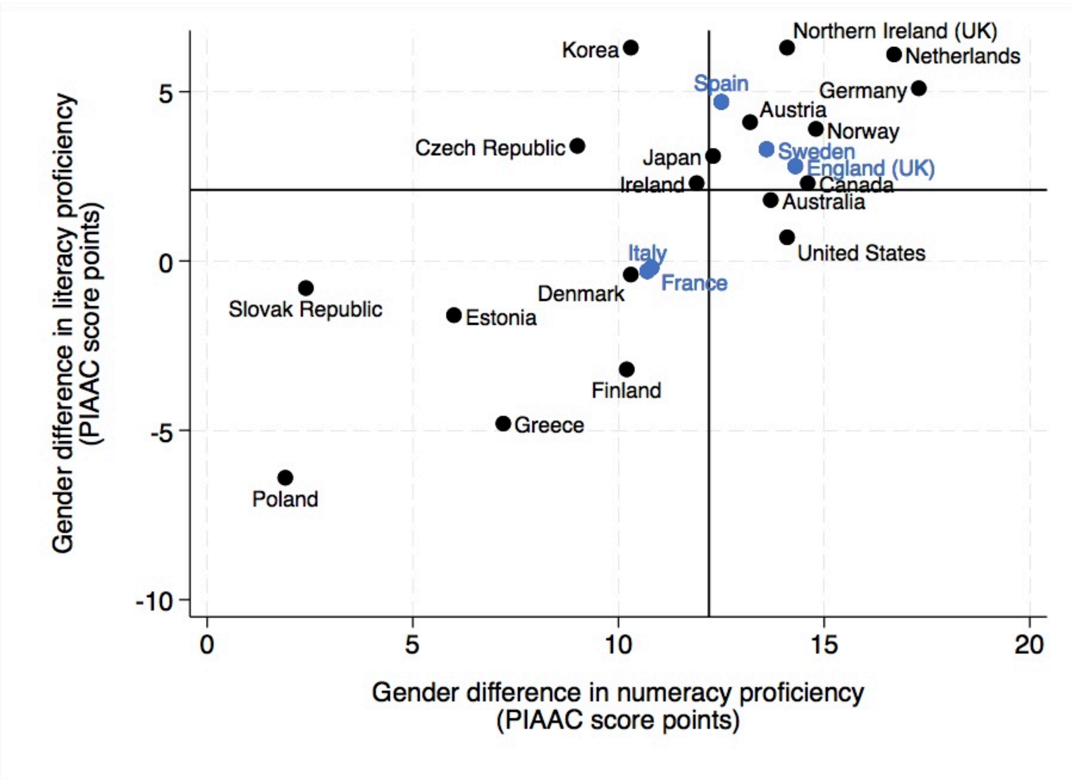


Source: PIAAC and OECD (2016)

Figure 4. 2. Dispersion between the highest and the lowest 5 percent PIAAC literacy and numeracy scores values by country.

The average range between the highest and lowest 5 percent performers in numeracy is 167 score points for the average of the OECD countries (171.1 points for those in the 5th percentile, below the 1st proficiency level; and 341.9 points for those in the 95th percentile, the 4th proficiency level) and 173 in the Spanish case (149.1 points for those in the 5th percentile, below the 1st proficiency level; and 322.3 points for those in the 95th percentile, the 3rd proficiency level) (OECD, 2013a; OECD, 2013b; OECD, 2016) (Fig. 4.2).

The degree of variation between the highest and the lowest score points indicates that, in Spain, the individual who scored the highest score value performs twice as well in literacy or numeracy as the individual with the lowest score value. This is an important variation, and it indicates considerable differences within the country. Among the five countries analysed in this chapter, England, France, Sweden and Spain are the ones which have higher differences between the skills levels within the country. In the figure, they are on the upper-right quadrant, just above the OECD average either for numeracy or literacy scores. In contrast, Italy, with the lowest within country differences among the countries considered here, is situated on the lower-left quadrant of the figure (Fig. 4.2).



Source: PIAAC and OECD (2016)

Figure 4. 3. Gender difference in PIAAC literacy and numeracy proficiency scores by country.

Figure 4.3 accounts for the gender difference in the average literacy and numeracy PIAAC score points across countries. The OECD average (continuous lines) indicates a higher gender difference, with higher average scores for men than for women, and in numeracy proficiency than in literacy ones (2.1 for literacy scores and 12.2 for numeracy scores). A similar distribution has the gender difference of the PIAAC scores in the five countries considered here. In Italy and France, there are smaller gender differences than the average of the OECD countries. In particular, Italy and France barely have gender differences in relation to the literacy proficiency scores (-0.3 and -0.2, respectively). In contrast, in England and Sweden, gender differences appear to be higher in numeracy scores than in literacy ones. In Spain, the data are the other way round: gender differences in numeracy proficiency are nearer to the average of the OECD countries than literacy ones.

Therefore, one of the characteristics of the countries selected here is their relatively low PIAAC scores. Besides, four of them (England, France, Spain, and Sweden) also depict a high dispersion between the highest and the lowest 5 percent PIAAC literacy and numeracy scores values (Fig. 4.2). Of those, England, Sweden, and Spain also register a gender difference, both in numeracy and literacy scores (Fig. 4.3). In contrast, Italy seems to be the most homogeneous country of the five considered here. It has the lowest dispersion in both indicators, the dispersion between the highest and the lowest 5 percent PIAAC proficiency scores, and the gender difference in both literacy and numeracy proficiency scores. Overall, the data shown so far do not depict a clear difference across countries. Pairwise comparisons will confirm whether or not these differences exist.

In this chapter, the 25-64-year-old individuals who are not in the army have been considered. Self-employed, employers, and unpaid family workers have not been included in the sample because of the difficulty to determine the real level of education they need to carry out their job (García-Mainar and Montuenga-Gómez, 2005; Verdugo and Verdugo, 1989).

Regarding age, and as seen in previous chapters, the youngest age, 25 year old, is the age at which most of the university students enrolled at the age of 18-19 year old have graduated, at least in the Spanish educational system.[90] Even if graduates of all education levels are considered here, the low threshold aims at

[90] In Spain, in the last academic year, around 70 percent of the students enrolled at university for the first time were under 25 years old (57 percent of them under 22 years old). They will probably finish the degree when they are older than 25, and therefore, it seems sensible to analyse only the university graduates from the age of 25.

accounting for the age at which most of the individuals, including university students, have finished their formal education. The upper threshold, 64 year old, is marked by the legal age of retirement at 65 years old.

All five countries considered in this chapter assessed the literacy and numeracy skills in PIAAC, and England was the only one where the problem solving in technology-rich environments was assessed. France is the only one where the reading components were not evaluated. Therefore, in this chapter only literacy and numeracy skills are considered.

4.3. Methodology

4.3.1. Variation in individual's cognitive skills across countries

Using data for the five countries considered in the chapter (England, France, Italy, Sweden, and Spain), I focus on the 25-64-year-old population to measure the degree of variation in skills across countries. One-way analysis of variances (ANOVA) is applied to disentangle the variation in the mean of the literacy and numeracy PIAAC scores across countries. The purpose of these calculations is to check if the variation in the literacy and numeracy PIAAC scores among individuals in Spain is significant, and therefore, that it is sensible to use them as a measurement for cognitive skill heterogeneity among individuals. Subsequently, the Heckman two-step model is used in this chapter to measure the effect of skills heterogeneity on the probability of being over-educated in the Spanish labour market. The results of these calculations are presented in the following section (4.4).

Throughout the thesis, I have mentioned the variation in abilities (innate and acquired) among individuals. Data used in Chapter Three do not allow us to account for this heterogeneity and, therefore, some homogeneity among them was assumed. As a consequence, the proportion of over-educated university graduates found in Chapter Three may be overestimated, since these results may include some university graduates with low skills. Even so, the previous chapter is useful to account for the general trend in the proportion of university graduates who are over-educated, since supposedly it shows some consistency in the distribution of the skills among individuals among years.

PIAAC data do account for some of this possible skill heterogeneity and, therefore, it will be considered in this fourth chapter. In particular, PIAAC enables one to account for literacy and numeracy proficiency, given the different values for the two observable cognitive skills.

Other skills would also be important or affect the possibility of securing a highly skilled job, but most of them are not observable or are difficult to measure. There is, for instance, the ability of the worker to respond to changes in the work organisation or to unforeseen situations, or their ability to work in different teams or being multitasking, all of them being useful attributes in nowadays' economies. The continuous and quick technological changes and innovations have made that these sorts of attributes are every day more and more required in the jobs offered in the labour market.

PIAAC gives ten plausible values for each cognitive skill (literacy and numeracy) for each individual. These plausible values have been calculated using the Item Response Theory (IRT). In order to have one unique value for each individual and cognitive skill, for the analysis of the variation in skills I use the mean of these ten plausible values for each of the cognitive skills considered, literacy and numeracy.

The dependent variable (literacy or numeracy scores) is continuous and the independent variable (country) is categorical, then a one-way ANOVA analysis seems to be the adequate test to determine whether the mean of the literacy and numeracy scores of each country are different. However, before underpinning the one-way ANOVA some previous analysis of the data has to be done.

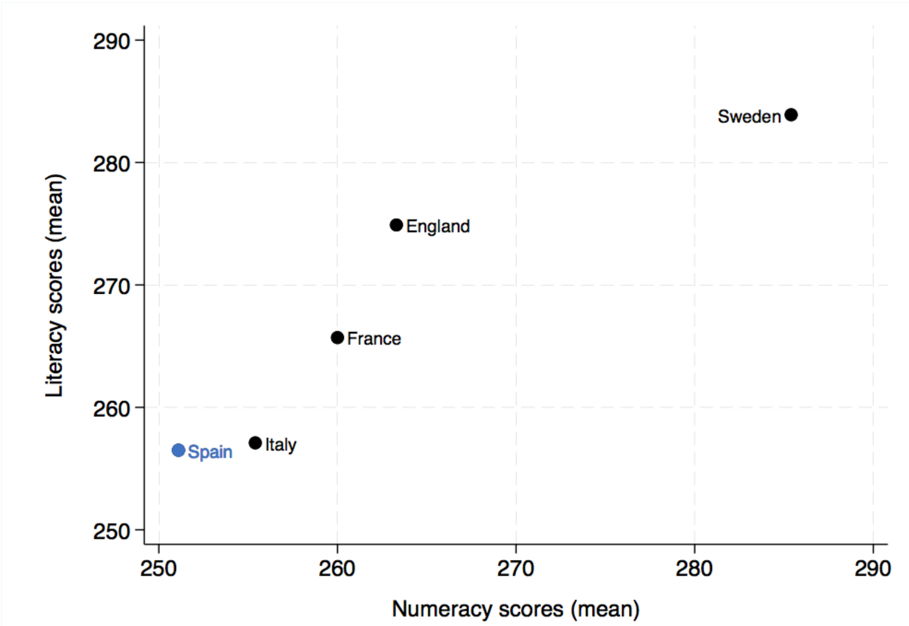
First, I examined a boxplot to identify possible outliers, and to eliminate them and their biasing effect on the results, I kept the data ranged between the 1st and the 99th percentile. The final sample used to measure the variation on individuals' skills contains 21,088 observations (45.8 percent men and 54.2 percent women) (Table 4.3).

Table 4. 3. Observations' distribution by country

	Total (N)	Country's sample (%)
France	5,510	26.1
Italy	3,528	16.7
Spain	4,620	21.9
Sweden	3,385	16.1
England	4,045	19.2
Total	21,088	100.0

Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Figure 4.4 depicts the distribution of the mean of the literacy and numeracy PIAAC scores for 25-64-year-old individuals across countries. Data show lower mean cognitive skills in the Southern countries (Spain and Italy) than in the other countries. The mean literacy and numeracy scores for 25-64-year-old individuals in France and England are higher than in the Southern countries, but lower than in Sweden. Pairwise comparison will measure the degree of difference between the literacy and numeracy mean scores by country to evaluate whether these differences are statistically significant. Besides, a Levene's test for equal variances will help to evaluate the hypothesis of equal variances among groups.



Source: PIAAC http://www.oecd.org/skills/piaac/publicdataandanalysis/

Figure 4. 4. Mean of PIAAC literacy and numeracy proficiency scores for 25-64-year-old individuals by country

As will be shown in the next section, where the results of the calculations are presented, the null hypothesis for the Levene's test is rejected for all the subsamples. Therefore, the equal variances among groups cannot be confirmed and it is plausible to expect differences among the distribution of the data across countries and between genders.

The one-way ANOVA confirms there are mean differences among at least some groups, although these differences are not identified yet. Finally, I use a pairwise comparison of means adjusted with a Tukey test to disentangle whether the variances are equal across groups. As the number of means to compare is large, it makes more sense to use the Tukey test. It is widely accepted in the statistical literature that the Tukey test gives powerful results when the number of means to test is higher, whereas, Bonferroni, for instance, is more adequate when the number of possible comparisons is small (Black, 2009; Keller, 2015).

The pairwise comparison of means identifies which of the mean differences among groups are statistically significant. The results of all these calculations can be found in Section 4.4.

4.3.2. Measuring the cognitive skills effect on the probability of being over-educated

After ensuring that there does exist some variation in observable cognitive skills among individuals, I measure the degree of over-education to then calculate the likelihood of being over-educated. The goal is to conclude to what extent accounting for individuals' observable cognitive skill heterogeneity influences the probability of being over-educated.

According to the Job Competition Model (Thurow, 1970), Assignment Theory (Tinbergen, 1956; Sattinger, 1980; Hartog, 1981), or even Job-Matching Theory (Jovanovic, 1979a; Jovanovic, 1979b; Sicherman, 1987; Sicherman and Galor, 1990), a negative relationship between the skills level and the likelihood of being over-educated is expected. Thus, those individuals who are more skilled are less likely to be over-educated because the labour market should have matched them with higher skilled jobs, since they hold higher skills.

Regarding education mismatch, and as has been described and explained in the literature review and in the previous chapter, it can be mainly measured in three ways: objective, subjective, and statistical. However, as Green, McIntosh and Vignoles (1999) have already stated, the main problem is that there is still lacking a unique and efficient measurement for education mismatch, and applying one or another measurement would change the results, due to their sensitivity.

In Chapter Three, I used the statistical mean approach since the SLFS data do not allow one to apply another type of measurement. However, PIAAC data do allow to calculate the three approaches. Some authors have already used these data to calculate these mismatches using alternative measurements. For instance, Martínez García (2013) used objective, subjective and statistical methods and found that the degree of over-qualification ranges from 13.8 percent (statistical method) to 20.7 percent (subjective method) for the working population between 25 and 65 years old. Flisi *et al.* (2014), on the other hand, found that the level of over-education level in Spain ranged from 12 percent when the education mismatch is measured using years of education (YEAR1, statistical mean approach) to 43 percent when they use subjective methods (SUB_EDU1, using the level of education and self-reported opinion on the real level of education needed for the current job). According to previous studies, the proportion of mismatched workers varies in relation to the approach used to measure them, being higher when subjective methods are used.

In this chapter, I measure the education mismatch using two different approaches: statistical mean and subjective. For the statistical mean approach and following Verdugo and Verdugo (1989), I apply the same rule as in the previous chapter. Thus, I calculate the mean of years of education for the population employed in each two-digit break-up occupation group. The mean of years of education for each occupation group plus/minus a standard deviation is used as a threshold to divide the workers as over-educated, under-educated or adequately matched. Therefore, individuals are classified as adequately matched when their years of education are between the upper and the lower thresholds. They are classified as over-educated if the years of education they have done are higher than the ones of the mean plus a standard deviation of the rest of the counterparts working in the same occupational group. And they are considered as under-educated if the years of education they hold are lower than the low threshold. Finally, those individuals with a number of years of education between the upper and the lower thresholds are classified as education adequately matched.

In contrast, the subjective method calculates the difference between the years of attained education and the years of education required to get the job to determine the number of years of education mismatch (over-/required/under- education) for each individual. In PIAAC data, both variables *yrsqual* (years of attained education) and *yrsget* (years of education required to get the job) are derived from individuals' answers. Individuals answer in ISCED education levels, and those are transformed to the minimum years of education to obtain that grade.[91] The concrete questions are: "Which of the qualifications on this card is the highest you have obtained?" for *yrsqual (*question *b_q01a*) and "Still talking about your current job: If applying today, what would be the usual qualifications, if any, that someone would need to GET this type of job?" for *yrsget* (question *d_q12a*). Therefore, according to the answers to these questions, individuals are classified as over-educated (under-educated) if the years of attained education (*yrsqual*) are higher (lower) than the years of education required to get the job (*yrsget*), and they hold the required education level (matched or adequately educated) if the years of attained education are the same as the ones required by the job.

The distribution of education mismatch in the Spanish labour market can be found in the next section. Moreover, a Pearson Chi-squared test for non-correlation is calculated to test the correlation between both of the education measurements.

[91] ISCED is the International Standard Education Classification (https://ec.europa.eu/eurostat/statistics-explained/index.php/International_Standard_Classification_of_Education_(ISCED))

Once the population has been classified according to the education mismatch, the probability of being over-educated is calculated. In order to account for a possible sample selection, the Heckman two-step model has been applied. To reiterate, the objective of using this model is to correct the possible bias from non-randomly selected samples. In this sense, and according to Trevis Certo *et al.* (2016), the significance of the estimated coefficients of the explanatory variables of the selection equation should be tested to corroborate whether there is a potential problem of sample selection. These calculations can therefore be used to validate whether there is a sample selection problem or not. Moreover, a probit model has also been estimated to be used as a baseline for comparison with the Heckman two-step estimations.

Regarding the Heckman two-step model, and as has been shown in the previous chapter, the assumptions required by this model imply the need to use it cautiously. First, there can be some problems when the sample is small or it contains few observed data points. In order to have an adequate sample, in this chapter I used all the 25-64-year-old population working as employees instead of using only data for 25-64-year-old university graduates. However, a variable accounting for qualification level facilitates the division of the individuals between those with a university degree and those with lower educational levels. Overall, the total sample includes 4,442 observations 22 percent of which are individuals holding at least one university degree. Regarding their labour situation, 58 percent of these people are working as employees and 42 percent are unemployed. From those 2,586 individuals working as employees, 26.8 percent are over-educated when the statistical mean approach is used, and 36.7 percent when the subjective approach is used.

Table 4. 4. Main statistics of the sample

	Total	Men (%)	Women (%)
Observations	4,442	46.7	53.3
University graduates	993	40.9	59.1
Employed	2,586	51.5	48.5
Over-educated (Statistical mean approach)	694	42.8	57.2
Over-educated (Subjective method)	948	47.3	52.7

Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Equally as in the previous chapter, I follow Trevis Certo *et al.* (2016) to confirm that the Heckman two-step model is adequate to analyse these data. First, as was demonstrated in the previous paragraph, the dependent variable of the outcome equation (being over-educated) is only observed in a subsample. Second, the effect of other possible covariates on the correlation between the error terms should also be considered. In this sense, the *size of the company* or the *industrial sector* are also two covariates that may be considered. A negative relationship between the size of the company and the likelihood of being over-educated is expected. But, the influence of the industrial sector is not clear. It depends on the category of reference and the way the sectors are classified. Therefore, it is not possible to conclude which will be the direction of the correlation between the error terms when the other potential omitted variables are considered. Nevertheless, the important issue to consider here is that if these variables were included in the model, they would reinforce (positively or negatively) the correlation between the error term of the two equations (outcome and selection).

Third, in the following section we will see how the significance of the Wald test of independent equations (*athrho*) rejects the null hypothesis of no correlation between the error terms, at least when the statistical mean approach is used to measure the degree of education mismatch. This result validates the appropriateness of using the Heckman two-step model.

Formally, the Heckman two-step model can be defined as follows:

$$\Pr(\text{Over-educated}) = \beta_0 + \beta_j X_{ij} + \varepsilon_i \qquad \text{Outcome equation} \qquad [4.1]$$

$$Z^* = \alpha_j W_{ij} + e_i \qquad Z = \begin{cases} 1 \text{ if } Z^*>0; \\ 0 \text{ otherwise} \end{cases}$$

$$\Pr(Z=1) = \phi(\alpha_j W_{ij}) \qquad \text{Selection equation} \qquad [4.2]$$

In this case, the variables included in the outcome and selected equations will be defined as follows:

BASELINE SPECIFICATION

Outcome equation [4.1]

- **Dependent variable**: whether the individual is over-educated (dummy variable: 1 if s/he is over-educated. 0 if s/he is not over-educated (s/he is adequately educated or under-educated)).
- **Independent variables:** *personal characteristics* (woman, age group, nationality, and qualification), *a spatial variable* (NUTS-1 region of residence), *and a variable to define the job characteristics* (public/private sector).

  Where,
  - *Woman* (dummy variable: 1 if woman, 0 if man).
  - *Age group* (dummy variable (4 groups): 25-34, 35-44, 45-54, and 55-64; reference category: 25-34).
  - Not Spanish (dummy variable: 1 if not Spanish; 0 if Spanish).
  - *University graduates* (dummy variable: 1 if s/he holds a university degree or a higher educational level, 0 if s/he holds an educational level lower than university degree).
  - Private sector (dummy variable: 1 if public sector; 0 if private sector).
  - *Region* (dummy variable using the NUTS-1 classification; reference category: area of Madrid).

Selection equation [4.2]

- **Dependent variable:** *whether the individual is employed* (dummy variable: 1 if s/he is an employed employee, 0 if s/he is an unemployed employee).
- **Independent variables:** *personal characteristics* (woman, age group, nationality, qualification, and health), *a spatial variable* (NUTS-1 region of residence).

Where,

- *Woman* (dummy variable: 1 if woman, 0 if man).
- *Age group* (dummy variable (4 groups): 25-34, 35-44, 45-54, and 55-64; reference category: 25-34).
- Not Spanish (dummy variable: 1 if not Spanish; 0 if Spanish).
- *University graduates* (dummy variable: 1 if s/he holds a university degree or a higher educational level, 0 if otherwise).
- *Excellent-good health* (dummy variable: 1 if the subject has declared to have a good, very good or excellent health status; 0 if the subject has declared to have a poor or very poor health status).
- *Region* (dummy variable using the NUTS-1 classification; reference category: area of Madrid).

Using the previous classifications of education mismatch (statistical mean approach and subjective approach), the dependent variable of the outcome equation [4.1] is a dummy variable that has value 1 when the individual is classified as over-educated, and value 0 when the individual is not over-educated, thus s/he is under-educated or adequately matched.

The independent variables of the outcome equation can be classified into three main groups: those accounting for the personal characteristics of the individual, those considering the characteristics of the region where the individual is living, and a third group related to the job characteristics.

Regarding the first group, the personal characteristics, all of them are dummy variables. W*oman*, *Not Spanish* and *University graduates* are single dummy variables since the variable only has two categories. They only have value '0' or '1'. The value of the '0' and '1' for each of the variables has been clearly explained in the previous description of the model. For instance, *university graduates* is a variable accounting for the educational level of the individual and, as has already been explained, it has value '1' when the individual has an educational level equal or higher to ISCED 5B (first stage of tertiary education), and value '0' when the individual has a lower educational level.

The variable *age group* requires a more precise description. *Age group* includes four categories: ‘25-34’, ‘35-44’, ’45-54’, and ‘55-64’. From them, four dummy variables has been created to include in the model. The youngest (‘25-34’) is the reference category and it aims at including all the individuals who are finishing their formal studies and entering the labour market for the first time. The group ‘35-44’ includes, in the Spanish case, those individuals who are starting or have just started to live independently from the parents, either living alone or as a couple with a partner. Most of them will also probably have children. The second middle group (’45-54’) mainly includes families with or without children. The difference from the previous group is in the fact that in this latter group individuals have been in the labour market for longer. According to the Job Matching Theory it would indicate they are less likely to be over-educated, since they have had time to adjust their educational level to their job. Finally, the oldest group, ‘55-64’, will include the last years of the career, meaning, the consolidation of the career, and even, in some cases, pre-retirement, although the latter will be less likely to try to extend the years of active work and to delay retirement thanks to the new policies.

About the *region* of residence, I have used here the NUTS-1. The same region classification is used in Chapters Three and Five. NUTS-1 is a group of large administrative areas and they are considered to be major socio-economic regions by the European Commission. Therefore, it makes sense to consider them from the point of view of the Spanish labour market. For the Spanish case, NUTS-1 includes seven regions.

As regards the job characteristics, only whether the workplace is in the public or private sector has been considered. In fact, this variable can be considered as a proxy of the size of the firm, since smaller companies are usually in the private sector, especially in the Spanish case where more than 85 percent of the private companies are small and midsize enterprises (SMEs).[92]

---

[92] https://industria.gob.es/es-es/Servicios/MarcoEstrategicoPYME/Marco%20Es tratégico%20PYME.pdf

I have considered to also including other variables as for instance, the industry sector, the field of study or whether the individual has participated in training activities recently.

A covariate accounting for the industrial sector was included in some models, however its inclusion reduced the statistical significance of the model. Moreover, the post-estimation Wald test suggests not including it. Therefore, in the end it was not included.

Regarding field of study, it has been demonstrated that it is likely that some fields of study teach more specific skills than others and that it influences the probability of being employed and being over-educated (Dolton and Vignoles, 2000; Green, McIntosh and Vignoles, 1999; Montt, 2015). In the previous chapters, little has been said regarding the possible difference in the probability of being over-educated that can emerge among university graduates due to the field of study of their degree. PIAAC data do include field of study of the degree, and therefore, the possible situations in the labour market or the difference level of skills they have acquired at university can be considered.

Despite the fact that field of study is an important variable to consider, the few observations of each of the categories when this variable is included made the estimation of the model impossible. Besides, the post estimation F test suggests not including them.

I have also tried to include a covariate accounting for whether the individual has participated in training activities since when accounting for abilities not only is it important to measure them, but also to take into account when they have been acquired and, furthermore, if they have been used since they were acquired. It is well known and frequently discussed in the literature that skills are not permanent, one can improve them, but they can also be reduced or worsened when they are not used. Time might have a negative effect on the skills acquired and part of this knowledge might be obsolete, especially when the individual is not using it (Brown, 2001; Desjardins and Warnke, 2012; Weiss, 1995). Consequently, the continuous training of the workers, either in or out of the workplace as well as the continuous use of these skills, becomes extremely important to maintain workers' skills. Continuous technical and productive innovation makes it necessary to update the workers' knowledge in general, but especially for those who are underusing their skills. This is especially true in those economies where the level of educational mismatch is not negligible and permanent, since those mismatched individuals are underusing their skills at work. However, the post estimation adjusted Wald test for 'training' was not statistically significant and therefore the null hypothesis could not be rejected indicating that the coefficients for this variable may be 0. Besides, the inclusion of this variable reduces its statistical significance. Therefore, I have not include it in the end.

The Heckman two-step model has a second equation, the selection equation [4.2]. This equation accounts for a possible sample selection of the data. In this case, it accounts for the probability for an individual to have already been chosen to be employed. The main problem is that, the dependent variable of the selection and the outcome equation are usually related, and therefore it is difficult to find a variable that only affects the dependent term of the equation [4.2]. In this case, different variables have been considered to be candidates for exclusion variables: whether the individual has children, whether their partner is working, or the general health status reported by the individual.

Regarding the partner's labour situation, it is likely that the labour situation of the partner affects the probability for the individual to be employed or not. When in the household there are two or more people and one of them is working, it may influence the likelihood that the other person accepts a job or not, depending on their characteristics. The sign of this impact is uncertain since it depends first on, whether the partner is working, and second, on which sort of job s/he has, or to be more precise, in the amount of the wage s/he receives. The post estimation F test confirms the statistical significance of the variable rejecting the null hypothesis and suggesting including the variable in the model. However, the higher standard errors for the estimated coefficients of the variables of the outcome equation suggest it is not a good exclusion restriction.

A similar impact is produced whether there are children at home or not, since the need to receive a regular wage is essential when there are children in the household. Again, the impact of this variable is not obvious because on the one hand, it is expected that those families with young children are more prone to stay at home with them. But, on the other hand, they will also need higher incomes to maintain them. In this case, the post estimation F test does not reject the null hypothesis of coefficient being equal to zero. These results suggest that the variable should not be included since it was not contributing to the model, and therefore, it has after consideration not been included.

Finally, a variable accounting for the general state of health of the individuals included in the background questionnaire has been included as a restriction variable. The individual answered the following question: *“In general, would you say your health is excellent, very good, good, fair, or poor?”*. The variable recoded here as a dichotomous variable with value ‘1’ if the individual replies that they have good, very good or excellent health status, and ‘0’ if otherwise. A formal Hausman-type test was run to confirm that after the recodification the regression coefficients for the variable health are indeed the same. Besides, the post-estimation F test confirms the significance of including it. Besides, the estimated coefficients for the outcome equations are either statistically significant or with lower standard errors.

The rest of the independent variables included in the selection equation [4.2] are also included in the outcome equation [4.1] and they have already been described and explained in the previous paragraphs.

The objective of the chapter is to allow for the effect of skill heterogeneity on the probability for university graduates to be over-educated. Therefore, in order to see how the inclusion of a variable accounting for skill heterogeneity affects the probability of being over-educated, I run three different specifications of the probit and the Heckman two-step models. The baseline specification (BS) is the one described above. As was detailed in the formal definition of the model, the other two specifications derived from BS include a new covariant in the outcome equation [4.1]. This new independent variable is a numeric one and it accounts for the literacy PIAAC scores in the literacy specification (LS) and for the numeracy PIAAC scores in the numeracy specification (NS). I run two different models including the cognitive skills separately because both scores cannot be included jointly since they are highly correlated (Flisi *et al.*, 2014; Levels, Van der Velden and Allen, 2013; Nieto and Ramos, 2011). LS and NS are also specified below.

Previous literature has pointed out the possible problems of endogeneity in the use of educational level and cognitive skills to analyse their effect on over-education, since it is expected that those individuals with higher educational level also have higher cognitive skills. However, the importance of accounting for skills heterogeneity recommend including this variable.

LITERACY SPECIFICATION (LS)

Outcome equation [4.1]

- **Dependent variable**: as in the Baseline Specification (BS), the dependent variable is whether the individual is over-educated (dummy variable: 1 if s/he is over-educated. 0 if s/he is not over-educated (s/he is adequately educated or under-educated)).
- **Independent variables.** This specification includes the same as in the BS: *personal characteristics* (woman, age group, and qualification), *a spatial variable* (NUTS-1 region of residence), *and a variable to define the job characteristics* (public/private sector). But it also includes an additional variable accounting for the literacy proficiency skills.
    - *Literacy* (continuous variable accounting for PIAAC literacy proficiency scores)

Selection equation [4.2]

- **Dependent variable**: as in the BS, the dependent variable is *whether the individual is employed* (dummy variable: 1 if s/he is an employed employee, 0 if s/he is an unemployed employee).
- **Independent variables.** This specification includes the same as in the BS: *personal characteristics* (woman, age group, nationality, qualification, and health), *a spatial variable* (NUTS-1 region of residence). But it also includes an additional variable accounting for the literacy proficiency skills.

- *Literacy* (continuous variable accounting for PIAAC literacy proficiency scores)

NUMERACY SPECIFICATION (NS)

Outcome equation [4.1]

- **Dependent variable**: as in the BS, the dependent variable is whether the individual is over-educated (dummy variable: 1 if s/he is over-educated. 0 if s/he is not over-educated (s/he is adequately educated or under-educated)).
- **Independent variables.** This specification includes the same as in the BS: *personal characteristics* (woman, age group, and qualification), *a spatial variable* (NUTS-1 region of residence), *and a variable to define the job characteristics* (public/private sector). But it also includes an additional variable accounting for the literacy proficiency skills.
    - *Numeracy* (continuous variable accounting for PIAAC numeracy proficiency scores)

Selection equation [4.2]

- **Dependent variable**: as in the BS, the dependent variable is *whether the individual is employed* (dummy variable: 1 if s/he is an employed employee, 0 if s/he is an unemployed employee).
- **Independent variables.** This specification includes the same as in the BS: *personal characteristics* (woman, age group, nationality, qualification, and health), *a spatial variable* (NUTS-1 region of residence). But it also includes an additional variable accounting for the literacy proficiency skills.
    - *Numeracy* (continuous variable accounting for PIAAC numeracy proficiency scores)

All the models have been run taking into account the design of the sample. That is, the 80 replications of the jackknife method have been applied to correct the estimation bias and to obtain more robust standard errors. Besides, when the literacy and numeracy scores are included as covariates (LS and NS, respectively), all ten plausible values have been imputed using multiple imputation methods to reduce the test error caused by the IRT applied by PIAAC to get the ten plausible values for these two variables.

4.4. Results

This section is divided into three sub-sections in order to facilitate the analyses of the results. The following subsection depicts the results for the pairwise comparison of the PIAAC proficiency scores by country. Subsection 4.4.2 depicts the estimated coefficients for both models (probit and the Heckman two-step) when the education mismatch has been calculated using the statistical mean approach. Subsection 4.4.3 analyses the predicted probability for a 25-64-year-old university graduate to be employed and over-educated conditional on the probability of being already employed.

4.4.1. Variation in individuals' cognitive skills across countries

Using PIAAC data for the five countries considered in the previous parts of this chapter (England, France, Italy, Spain, and Sweden), I analyse here to what extent the difference between the means of the literacy and numeracy PIAAC scores across countries are statistically significant (Table 4.5).

The test of means difference between groups and the Levene's test for equal variances confirms in both cases, for literacy and numeracy scores, that neither the means nor the variances are equal among groups. Furthermore, the contrast between the mean by country confirms the country's dispersion shown in Figure 4.4.

Table 4.5 shows, for instance, that the French scored on average 10.2 points more than the Spanish in the literacy proficiency test, and 9.6 points more in the numeracy proficiency test. Likewise, the Swedish scored 29.2 literacy proficiency points more than the Spanish, and 36.7 points more in numeracy, on average. The lowest differences were between the Italian and Spanish cognitive skills.

Indeed, the Tukey test does not identify a statistically significant difference at the 5 percent level for the mean of the literacy skills for the Italian and the Spanish.

Table 4. 5. Pairwise comparison of the means of the literacy and numeracy scores by country.

		Literacy		Numeracy	
Spain	**France**	10.1776 *0.8510*	***	9.6600 *0.9593*	***
	Italy	0.9144 *0.9462*	n.s.	4.5306 *1.0667*	***
	Sweden	29.2362 *0.9737*	***	36.7170 *1.0977*	***
	England	19.9312 *0.9176*	***	14.1500 *1.0343*	***
Test of means' differences between groups (F)		323.08	***	316.49	***
Levene's test for equal variances (F)		17.4172	***	35.2610	***

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Note: Standard errors are in italics.
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Therefore, from the five countries considered here, Spain is the country where their workers have scored lower literacy and numeracy PIAAC proficiency scores. The larger differences are in the numeracy scores between the Swedish and the Spanish. Besides, the standard errors for both cognitive skills and for all countries are relatively small, indicating that the difference of the means is rather precise. These results contrasted with the fact that the proportion of over-educated university graduates in the Spanish labour market is one of the highest in European countries.

Differences in the educational system, but also in the labour market structures may be the main causes of this disparity. However, a pairwise comparison of the means does not allow one to conclude what is causing this effect. Whether it is due, for instance, to the excess in the use of fixed-term contracts which reduces the probability for the worker to follow on-the-job training; or if it is due to the high rates of unemployment and especially long-term unemployment which can lead to a deterioration of the individuals' skills cannot be extracted from this analysis.

Nevertheless, it is important to bear in mind that lifelong learning is more and more essential in both cases, in life and at work. The continuous improvement of the technologies, and especially, all the innovations related to the Information and Communication Technologies (ICT) make continuous learning necessary to update and maintain individuals' skills and knowledge.

4.4.2. Measuring the cognitive skills effect on the probability of being over-educated

Focusing now on the Spanish labour market, a variation of the Heckman two-step model is used here to estimate the probabilities for an individual to be over-educated and what effect accounting for individuals' skills heterogeneity has on this probability. Besides, a probit model is also run to use as a baseline. In both cases, the dependent variable is whether the individual is over-educated. Therefore, individuals should be first classified by their education mismatch in the labour market.

Table 4. 6. Education mismatch in the Spanish labour market by gender and educational level. Statistical mean approach (in percentages)

Education mismatch	Lower than university education		University education or higher	
	Men	Women	Men	Women
Under-educated	32.5	31.3	0.0	0.0
Matched	60.3	61.7	57.6	53.6
Over-educated	7.2	7.0	42.4	46.4
Total	**100.0**	**100.0**	**100.0**	**100.0**

Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Tables 4.6 and 4.7 show the proportion of education mismatch in the labour market by gender and educational level. Both over-education measurements are categorical variables and all the cells of the two-way table contain more than five observations. Therefore, a Pearson Chi-squared test is adequate to check the independence between both measurements. The high statistical significance of this test (p=0.000) confirms the observed differences between both measurements are statistically significant.[93]

Table 4. 7. Education mismatch in the Spanish labour market by gender and educational level. Subjective approach (in percentages)

Education mismatch	Lower than university education		University education or higher	
	Men	Women	Men	Women
Under-educated	26.0	16.8	6.0	3.7
Matched	42.0	38.8	50.2	58.7
Over-educated	32.0	44.4	44.0	37.5
Total	**100.0**	**100.0**	**100.0**	**100.0**

Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

According to the results from Tables 4.6 and 4.7, the difference in the proportion of university-graduate women who are over-educated is 8.9 percentage points (pp), 37.5 percent using the subjective method and 46.4 percent using the statistical mean approach. For men, this difference is much lower, 1.6 pp (44.0 when the subjective approach is used and 42.4 percent using the statistical mean approach). Moreover, for men the subjective approach gives higher proportions of over-educated university graduates, whereas for women the higher percentage of over-educated university graduates is when the statistical mean approach is calculated.

These differences among the proportions of over-education by measurement may be explained by the differences in the nature of the measurements themselves, and especially by the nature of the subjective measurement. In this sense, and regarding men, the subjective approach gives a higher proportion of over-educated workers. This result is in concordance with the studies which conclude that the subjective measurements tend to overestimate the mismatches (Hartog, 2000). In contrast, the proportion of over-educated women is higher when the statistical mean approach is applied. These latter results may be more related with women’s lower expectations in their jobs.

---

[93] Pearson chi2(4) = 289.1027   Pr = 0.000

According to PIAAC data, the subjective method uses the worker's point of view to determine which qualification s/he would need to get the job if they were applying for it today. Concretely, the question that workers are answering is: "Still talking about your current job: If applying today, what would be the usual qualification, if any, that someone would need to GET this type of job?". Therefore, it seems sensible to suppose there is a gender difference in the relationship among the proportions of over-educated and their jobs' expectations.

Higher differences between both over-education measurements appear when individuals with lower educational levels are considered (lower than university education). In this latter case, the proportion of over-educated men or women is lower than 10 percent when the statistical mean approach is calculated, and it rises up to 32 percent for men and 44.4 for women when the subjective measurement is used. Notice here that the proportion of under-educated university graduates is zero when the subjective education mismatch measurement is applied.

The following tables show the estimated coefficients for the probit model when the statistical mean approach is used to calculate the proportion of education mismatch (Tables 4.8 to 4.10). The same results using subjective education mismatch measurement can be found in Appendix 3.

Estimated coefficients from a probit model cannot be interpreted in a straightforward way since the accumulative function of the probit model is based on a standard normal distribution. Therefore, the estimated coefficients account for the effect that a change in the independent variable has on the dependent variable in z-score units. However, here we are keen on anticipating the positive or negative effect that a specific variable has on the dependent variable maintaining constant the rest of the variables. Moreover, predicted probabilities for an individual to be over-educated and employed conditional on whether s/he is employed are also depicted in the next subsection.

For the three specifications (Baseline, Literacy, and Numeracy) and for both over-education measurements (statistical mean approach and subjective method), the statistic *F* is highly statistically significant when the probit model is applied (Tables

4.8 to 4.10 for the statistical mean approach and Appendix 3 for subjective methods).

Table 4. 8. Estimated coefficients of the probit model. BS and statistical mean approach

(Y=over-educated)

Independent variables (reference category)		Coef.	Jackknife Std. Err.	Sig. Level
Woman (Man)		-0.0852	0.0744	n.s.
Age-group (25-34)	**35-44**	-0.0484	0.0704	n.s.
	45-54	-0.2568	0.0989	*
	55-64	-0.2909	0.1198	*
Not Spanish (Spanish)		0.2911	0.1242	*
University graduates (Lower educational levels)		1.1647	0.0886	***
Public (Private)		-0.2226	0.0871	*
Region (Madrid)	**North-West**	0.3742	0.1520	*
	North-East	0.3385	0.1053	**
	Central Area	0.1335	0.1195	n.s.
	East	0.0630	0.1030	n.s.
	South	0.1524	0.1142	n.s.
	Canary Islands	-0.1067	0.2389	n.s.
Constant		-1.5767	0.1195	***
Number of observations		**2,459**		
F(13,67)		**17.26**	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

In contrast to what might be expected, the estimated coefficients for the probit model suggest a positive relationship between higher cognitive proficiency scores (literacy or numeracy) and the probability of being over-educated. Indeed, these estimated coefficients are only statistically significant for the LS when the subjective method is used to calculate the education mismatch (Appendix 3). For the NS when the subjective education mismatched measurement is used and for the literacy and numeracy specifications when the statistical mean measurement is used, the estimated coefficients are statistically not significant (Tables 4.8 to 4.10). However, these results may hide the positive relationship between being highly skilled and being employed since the probit model does not account for the sample selection effect. Therefore, it is possible that the positive effect that literacy or numeracy skills have on the probability of being employed is higher than the negative effect that they have on the probability of being over-educated. As a result, the total effect that the estimated coefficients of the probit model are showing in the probability of being over-education are, in fact, the latent effect of being employed.

Moreover, data suggest a higher effect for literacy than for numeracy scores. The inclusion of the literacy proficiency scores has a higher effect on the estimated coefficients of the other covariates. See for instance, the reduction in the estimated coefficient for the education level variable when the literacy and the numeracy scores are included in the estimation model. The higher effect of the literacy skills on the estimated coefficients of the probit model may be caused by the general higher literacy proficiency PIAAC scores than the numeracy ones for the Spanish population (Fig. 4.1).

Table 4. 9. Estimated coefficients of the probit model. LS and statistical mean approach

(Y=over-educated)

Independent variables (reference category)		Coef.	Jackknife Std. Err.	Sig. Level
Woman (Man)		-0.0675	0.0753	n.s.
Age-group (25-34)	**35-44**	-0.0464	0.0706	n.s.
	45-54	-0.2431	0.0979	*
	55-64	-0.2542	0.1241	*
Not Spanish (Spanish)		0.3242	0.1299	*
University graduates (Lower educational levels)		1.1134	0.0985	***
Literacy		0.0014	0.0011	n.s.
Public (Private)		-0.2302	0.0867	*
Region (Madrid)	**North-West**	0.3638	0.1503	*
	North-East	0.3448	0.1037	*
	Central Area	0.1295	0.1189	n.s.
	East	0.0580	0.1027	n.s.
	South	0.1469	0.1142	n.s.
	Canary Islands	-0.0973	0.2414	n.s.
Constant		-1.9371	0.3153	***
Number of observations		**2,459**		
F(14,76.9)		**19.01**	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

In the BS (Table 4.8), the estimated coefficient for the education variable was 1.1647, whereas it is 1.1134 in the LS (Table 4.9) and 1.1155 in the NS (Table 4.10). The positive sign of this estimated coefficient and its highly statistical significance confirm the higher probability for university graduates to be over-educated compared to those individuals holding lower educational levels. This is also confirmed by the literature.

Similar results are found when the subjective over-education measurement is used, even if the estimated coefficients for the education covariant are statistically not significant. The estimated coefficient for the education level in the BS is 0.0977, whereas it drops to 0.0910 for the NS and to 0.0624 for the LS (Appendix 3).

Table 4. 10. Estimated coefficients of the probit model. NS and statistical mean approach

(Y=over-educated)

Independent variables (reference category)		Coef.	Jackknife Std. Err.	Sig. Level
Woman (Man)		-0.0576	0.0761	n.s.
Age-group (25-34)	**35-44**	-0.0471	0.0705	n.s.
	45-54	-0.2434	0.0987	*
	55-64	-0.2573	0.1247	*
Not Spanish (Spanish)		0.3229	0.1283	*
University graduates (Lower educational levels)		1.1155	0.0967	***
Numeracy		0.0013	0.0010	n.s.
Public (Private)		-0.2284	0.0869	*
Region (Madrid)	**North-West**	0.3712	0.1508	*
	North-East	0.3466	0.1036	**
	Central Area	0.1342	0.1196	n.s.
	East	0.0616	0.1025	n.s.
	South	0.1569	0.1152	n.s.
	Canary Islands	-0.0885	0.2427	n.s.
Constant		-1.9308	0.2926	***
Number of observations		**2,549**		
F(14,77)		**19.5**	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Including cognitive skills as covariates also reduced the estimated coefficient for the variable accounting for gender (*Woman*), whereas it increases the differences between the probability of those workers being over-educated and employed in the public sector and those employed in the private sector. The estimated coefficients for the variable *Public* is -0.2226 in the baseline specification and -0.2302 and -0.2284 in the literacy and numeracy specification, respectively (Tables from 4.8 to 4.10). Estimated coefficients for this independent variable are only statistically significant when the statistical mean approach is used to measure the degree of education mismatch (Tables 4.8 to 4.10). In this case, the estimated coefficients are negative indicating that those working in the public sector are less likely to be over-educated. As has already been stated in the methodology section, this public/private sector difference in the probability of being over-educated may be hiding the size company effect. Private companies are more likely to be smaller than the public ones. Especially in the Spanish labour market where the SMEs represent more than 85 percent of the business network.

Regarding gender difference, the results are not conclusive since the estimated coefficients for the gender variable are statistically not significant in all of the specifications for the statistical mean approach. However, in all three specifications the estimated coefficients are negative indicating the lower probability for women to be over-educated than men. These results are not in concordance with the main literature, indicating that some latent variables, as for instance the lower probability for women being employed, have an influence on the gender differences in the probability of being over-educated.

Besides, and as expected, the results suggest a reduction in the gender difference in the probability of being-over-educated when the cognitive skills are included as independent variables in the model, indicating that the Spanish labour market is taking into account individuals' skill heterogeneity. In this sense, it appears that accounting for numeracy skills further reduces the value of the estimated coefficient. See, for instance, the estimated coefficient in Tables from 4.8 to 4.10, it is -0.0852 when the cognitive scores are not considered (Table 4.8), -0.0675 when the literacy skills are introduced as a covariant (Table 4.9), and -0.0576 when the numeracy skills are included (Table 4.10).

As regards *age-group*, estimated coefficients for the probit model for the 35-44 age-group are statistically not significant in all the specifications when the statistical mean approach is used to measure the degree of over-education (Tables 4.8 to 4.10). However, for all the specifications and measurements, the estimated coefficients are negative and, as in the previous variables, accounting for cognitive skills reduces the differences in the effect that being in one group or another has on the probability of being over-educated. The negative sign of the estimated coefficient for the oldest generations indicates that they are less likely to be over-educated than the youngest ones (the reference category 25-34). Indeed, for the three specifications when the statistical mean approach is used, the coefficient is higher in the oldest generation (-0.2434 for the 45-54 year-old and -0.2573 for the 55-64 year-old in the NS).

The estimated coefficients for the 35-44 age-group are statistically not significant, whereas estimated coefficients for the 45-54 age-group are only statistically significant when the statistical mean approach measurement is applied. For these latter two groups, the sign of the estimated coefficients is positive when the subjective method is used and negative when the statistical mean approach is applied. Therefore, results are not conclusive for these age-groups.

By *region*, few estimated coefficients are statistically significant for the three specifications and both education mismatch measurements. Moreover, the estimated coefficients do not follow a same pattern and the effect of including the cognitive skills is different in each of the categories. Nevertheless, two main tendencies can be extracted from these data. First, for all specifications the sign of the estimated coefficients is the same for each of the education mismatch measurements, indicating the robustness of the results. Second, including a variable accounting for cognitive skills does vary the estimated coefficient, although it is not in the same direction in all regions. These latter results suggest that accounting for them is relevant when the likelihood of being over-educated is calculated.

Overall, the estimated coefficients for the probit model does not confirm the negative effect between higher cognitive skills and the probability of being over-educated, although the sample selection might hide this effect. Nevertheless, including cognitive skills as covariates does affect the estimated coefficients of the rest of the covariates. In most of the cases, the inclusion of literacy and numeracy plausible values reduces the effect that other independent variables have on the probability of being over-educated. This result suggests the important effect that cognitive skills have on individuals' labour market opportunities, and therefore, the importance of accounting for cognitive skills when the likelihood of being over-educated is calculated.

The previous probit models have been estimated ignoring the possible sample selection. Estimated coefficients of Tables 4.8 to 4.10 and the fact that 40 percent of the population are not affected by the dependent variable of the selection equation (employed) suggests considering the fact that data are not randomly distributed. The estimated coefficients for the Heckman two-step model presented in the previous section are depicted in the following tables. The dependent variable of the outcome equation [4.1] is the probability of being over-educated, whereas the dependent variable of the selection equation [4.2] is the probability of being employed.

Using these PIAAC data, and to confirm that there is sample selection, estimated coefficients of the probit model of the selection equation are presented in Table 4.11 for the BS, and in Table 4.12 and 4.13 for the LS and NS, respectively.

These coefficients are statistically significant for most of the variables, although they are statistically not significant for some of the categories of age-group and region. These results suggest a potential sample selection problem and imply the need to use the Heckman two-step model to obtain more robust estimates (Trevis Certo *et al.*, 2016). In this case, the results indicate that to estimate the probability of being over-educated is relevant to condition it on the probability that the individual is already employed. This is true for both over-education measurements, since the selection equation does not include it as a variable. The potential sample selection of the data is also confirmed by the fact that only part of the population can be over-educated, those who are employed (60 percent).

Table 4. 11. Estimated coefficients of the probit model of the selection equation [4.2]. BS

Y= Employed

Independent variables (reference category)		Coeff.	Jackknife Std. Err.	Sig. Level
Woman (Man)		-0.2901	0.0381	***
Age-group (25-34)	35-44	0.0189	0.0599	n.s.
	45-54	-0.0651	0.0627	n.s.
	55-64	-0.7560	0.0637	***
Not Spanish (Spanish)		-0.2671	0.0640	***
University graduates (Lower educational levels)		0.6584	0.0547	***
Excellent-good health (Poor health)		0.1784	0.0508	**
Region (Madrid)	North-West	-0.1275	0.0834	n.s.
	North-East	-0.0244	0.0854	n.s.
	Central Area	-0.1991	0.0727	**
	East	-0.1901	0.0746	*
	South	-0.3981	0.0684	***
	Canary Islands	-0.7305	0.1293	***
Constant		0.6133	0.0872	***
Number of observations		4,213		
F(13,67)		50.27	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Table 4. 12. Estimated coefficients of the probit model of the selection equation [4.2]. LS

Y= Employed

Independent variables (Reference category)		Coeff.	Jakknife Std. Err.	Sig. Level
Woman (Man)		-0.2772	0.0395	***
Age-group (25-34)	35-44	0.0315	0.0603	n.s.
	45-54	-0.0181	0.0649	n.s.
	55-64	-0.6555	0.0680	***
Not Spanish (Spanish)		-0.1687	0.0647	*
University graduates (Lower educational levels)		0.5149	0.0565	***
Literacy		0.0034	0.0005	***
Excellent-good health (Poor health)		0.1522	0.0520	**
Region (Madrid)	North-West	-0.1500	0.0834	n.s.
	North-East	-0.0123	0.0865	n.s.
	Central Area	-0.2338	0.0739	**
	East	-0.2135	0.0744	**
	South	-0.4003	0.0684	***
	Canary Islands	-0.7091	0.1298	***
Constant		-0.2282	0.1602	n.s.
Number of observations		4,213		
F(14,76.8)		64.14	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Table 4. 13. Estimated coefficients of the probit model of the selection equation [4.2]. NS

Y= Employed

Independent variables (reference variables)		Coeff.	Jackknife Std.Err.	Sig. Level
Woman (Man)		-0.2418	0.0404	***
Age-group (25-34)	**35-44**	0.0306	0.0602	n.s.
	45-54	-0.0019	0.0652	n.s.
	55-64	-0.6308	0.0691	***
Not Spanish (Spanish)		-0.1461	0.0648	*
University graduates (Lower educational levels)		0.4715	0.0559	***
Numeracy		0.0043	0.0005	***
Excellent-good health (Poor health)		0.1503	0.0519	**
Region (Madrid)	**North-West**	-0.1299	0.0830	n.s.
	North-East	-0.0076	0.0856	n.s.
	Central Area	-0.2216	0.0740	**
	East	-0.2084	0.0731	**
	South	-0.3737	0.0687	***
	Canary Islands	-0.6675	0.1316	***
Constant		-0.4567	0.1521	**
Number of observations		**4,213**		
F(14,76.9)		**69.9**	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

A significance test for the exclusion variable included in the selection equation [4.2] has been run to check its validity. The highly statistically significant results for the post estimation test validate the use of the variable '*health*' as the exclusion variable.[94] Even so, and as expected, the estimated coefficients for this variable show higher probabilities of being employed for those reporting good or excellent health (Lawrence, 2017).

The statistical significance of the *athrho* and *rho* for the Heckman two-step model when the model includes the over-education measured by statistical approach confirms the correlation between the error terms of the outcome and selection

[94] For the sample where the education mismatch was calculated using the statistical mean approach, the adjusted Walt test for the exclusion variable was F(1,79)= 12.33 p=0.0007 for the BS; F(1,77)=8.56 p=0.0045 for the LS; and F(1,76.7)=8.38 p=0.0049 for the NS.

equation. This result suggests the use of the selection model to get unbiased estimated coefficients. The correlation between the error terms of both equations of the model cannot be confirmed when the subjective method is used to measure the degree of over-education since the *athrho* and *rho* are statistically not significant. The estimated coefficients for the dependent variables are not statistically significant either.

The estimated coefficients for the Heckman two-step model when the statistical mean approach is used to measure the degree of education mismatch are depicted in the following tables (Tables 4.14 to 4.16), whereas the estimated coefficients for the three specifications when the subjective education mismatch measurement is used can be found in Appendix 3.

Important differences in the estimated coefficients appear when the possible sample selection is considered compared to the results for the probit model. As has already been explained, an over-estimation of the coefficients for the probit model is expected, since it has not accounted for the fact that only individuals who are already employed are likely to be over-educated. The results depicted from Tables 4.14 to 4.16 not only show more accurate estimations but also sign changes. It is worth highlighting here, for instance, the change in the sign of the gender and cognitive skills variables.

The estimated coefficients of the selection equation show a positive relationship between higher cognitive skills and the probability of being employed indicating that those abler workers are the ones selected by the employers to work. Moreover, the negative sign of the variables accounting for cognitive skills for both measurements and specifications suggest the validity of the positive relationship between having higher skills and being employed in highly skilled jobs. It has already been stated by Assignment Theory (Tinbergen, 1956; Sattinger, 1980; Hartog, 1981), Job Competition Model (Thurow, 1970), and Job-Matching Theory (Jovanovic, 1979a; Jovanovic, 1979b; Sicherman, 1987; Sicherman and Galor, 1990).

Table 4. 14. Estimated coefficients of the Heckman two-step model. BS and statistical mean approach

Y=(over-educated |employed)

	Independent variables (Reference category)		Coef.	Jackknife Std. Err.	Sig. Level
Outcome equation	**Woman (Man)**		0.1162	*0.0681*	n.s.
	Age-group (25-34)	**35-44**	-0.0416	*0.0589*	n.s.
		45-54	-0.1224	*0.0851*	n.s.
		55-64	0.2971	*0.1675*	n.s.
	Not Spanish (Spanish)		0.3496	*0.0944*	***
	University graduates (Less than university education)		0.4303	*0.2175*	n.s.
	Public (private)		-0.1607	*0.0654*	*
	Region (Madrid)	**North-West**	0.3569	*0.1197*	**
		North-East	0.2669	*0.0948*	**
		Centre	0.2307	*0.0995*	*
		East	0.1617	*0.0795*	*
		South	0.3268	*0.0920*	**
		Canary Islands	0.3706	*0.2028*	n.s.
	Constant		-0.8037	*0.1927*	***
Selection equation	**Woman (Man)**		-0.3056	*0.0384*	***
	Age-group (25-34)	**35-44**	0.0145	*0.0613*	n.s.
		45-54	-0.0857	*0.0630*	n.s.
		55-64	-0.7597	*0.0666*	***
	Not Spanish (Spanish)		-0.2674	*0.0654*	***
	University graduates (Less than university education)		0.6593	*0.0556*	***
	Excellent-good health (Poor health)		0.1697	*0.0544*	**
	Region (Madrid)	**North-West**	-0.1092	*0.0799*	n.s.
		North-East	-0.0012	*0.0840*	n.s.
		Centre	-0.1651	*0.0699*	*
		East	-0.1739	*0.0704*	*
		South	-0.3730	*0.0678*	***
		Canary Islands	-0.7055	*0.1273*	***
	Constant		0.5995	*0.0841*	***
Number of observations			**4,158**		
F (13, 67)			**5.49**	***	
athrho			**-1.3304**	**	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Table 4. 15. Estimated coefficients of the Heckman two-step model. LS and statistical mean approach

Y=(over-educated |employed)

	Independent variables (Reference category)		Coef.	Jackknife Std Err.	Sig. Level
Outcome equation	Woman (Man)		0.1395	*0.0562*	*
	Age-group (25-34)	35-44	-0.0469	*0.0569*	n.s.
		45-54	-0.1233	*0.0763*	n.s.
		55-64	0.3105	*0.1263*	*
	Not Spanish (Spanish)		0.3046	*0.0924*	**
	University graduates (Less than university education)		0.3783	*0.1516*	*
	Literacy		-0.0013	*0.0008*	n.s.
	Private (Public)		-0.1485	*0.0615*	*
	Region (Madrid)	North-West	0.3470	*0.1079*	**
		North-East	0.2462	*0.0868*	**
		Centre	0.2536	*0.0953*	**
		East	0.1838	*0.0754*	*
		South	0.3407	*0.0870*	***
		Canary Islands	0.4171	*0.2003*	*
	Constant		-0.3944	*0.2934*	n.s.
Selection equation	Woman (Man)		-0.2895	*0.0398*	***
	Age-group (25-34)	35-44	0.0230	*0.0611*	n.s.
		45-54	-0.0401	*0.0641*	n.s.
		55-64	-0.6618	*0.0688*	***
	Not Spanish (Spanish)		-0.1651	*0.0655*	*
	University graduates (Less than university education)		0.5132	*0.0572*	***
	Literacy		0.0034	*0.0005*	***
	Excellent-good health (Poor health)		0.1360	*0.0532*	*
	Region (Madrid)	North-West	-0.1283	*0.0791*	n.s.
		North-East	0.0136	*0.0842*	n.s.
		Centre	-0.1936	*0.0697*	**
		East	-0.1919	*0.0702*	**
		South	-0.3699	*0.0674*	***
		Canary Islands	-0.6754	*0.1271*	***
	Constant		-0.2602	*0.1551*	n.s.
Number of observations			**4,158**		
F (13, .)			**.**		
athrho			**-1.5959**	***	
rho			**-0.9210**	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Table 4. 16. Estimated coefficients of the Heckman two-step model. NS and statistical mean approach

Y=(over-educated |employed)

Independent variables (Reference category)			Coef.	Jackknife Std. Err.	Sign. Level
Outcome equation	Woman (Man)		0.1303	0.0538	*
	Age-group (25-34)	35-44	-0.0482	0.0561	n.s.
		45-54	-0.1262	0.0744	n.s.
		55-64	0.3124	0.1156	**
	Not Spanish (Spanish)		0.2870	0.0890	**
	University graduates (Less than university education)		0.3711	0.1277	**
	Numeracy		-0.0020	0.0008	*
	Private (Public)		-0.1411	0.0589	*
	Region (Madrid)	North-West	0.3333	0.1059	**
		North-East	0.2362	0.0845	**
		Centre	0.2516	0.0936	**
		East	0.1847	0.0732	*
		South	0.3350	0.0843	***
		Canary Islands	0.4117	0.1982	*
	Constant		-0.1786	0.2727	n.s.
Selection equation	Woman (Man)		-0.2518	0.0407	***
	Age-group (25-34)	35-44	0.0173	0.0605	n.s.
		45-54	-0.0273	0.0640	n.s.
		55-64	-0.6387	0.0694	***
	Not Spanish (Spanish)		-0.1415	0.0658	*
	University graduates (Less than university education)		0.4676	0.0565	***
	Numeracy		0.0044	0.0005	***
	Excellent-good health (Poor health)		0.1285	0.0522	*
	Region	North-West	-0.1059	0.0790	n.s.
		North-East	0.0194	0.0833	n.s.
		Centre	-0.1786	0.0698	*
		East	-0.1866	0.0688	**
		South	-0.3395	0.0677	***
		Canary Islands	-0.6307	0.1288	***
	Constant		-0.5033	0.1474	**
Number of observations			**4,158**		
F (14, .)			**.**		
athrho			**-1.7187**	***	
rho			**-0.9377**	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Furthermore, accounting for cognitive skills affects the estimated coefficients of the rest of the covariates of the selection model, the probability of being employed. However, the same effect has not been found for the covariates of the outcome equation (probability of being over-educated).

Regarding gender, the positive sign of the estimated coefficients of the outcome equation and the negative ones for the selection equation are consistent with the main literature. Women are less likely to be employed and more likely to be over-educated when they are employed than men. The value of the estimated coefficients is higher for the literacy and numeracy specifications than for the baseline specification, indicating that including a variable accounting for skills heterogeneity increases the gender differences in the probability of being over-educated and reduces the gender differences in the probability of being employed (selection equation). These results may hide some latent effect from unobservable variables as for instance the gender differences in literacy and numeracy PIAAC proficiency scores (Fig. 4.3).

Regarding age-groups, the results from the Heckman two-step model differ depending on the education mismatch measurements. The estimated coefficients when the statistical mean approach is used depicts a higher probability of being over-educated for the oldest generations than the youngest, whereas the results are reversed when the subjective education mismatch measurement is used. Moreover, for all specifications and education mismatch measurement the estimated coefficients are statistically not significant.

By educational level, and as expected and already stated by the probit model, the estimated coefficients of the Heckman two-step model also confirm the higher probability of being employed and being over-educated for those workers holding a university degree. These results suggest that holding a university degree is still advantageous in the Spanish labour market, even if it is not in a ‘perfect job’, since being a university graduate increases the likelihood of being employed in the labour market although the tasks of the job do not require a university graduate to do them.

Spaniards still have lower probabilities of being over-educated than those holding a foreign nationality. As has already been explained above, the fact of holding a foreign degree may reduce the likelihood for foreign university graduates to be working in a highly skilled job in the Spanish labour market.

In general terms, accounting for skills' heterogeneity reduces the effect that the independent variables have in each of the categories in both dependent variables: the probability of being employed and the probability of being over-educated. Therefore, when the skills' heterogeneity is not considered the effect of the covariates on the dependent variable (the probability of being over-educated) are over-estimated, and in fact, they are hiding, in part, the effect of this skills' heterogeneity. This is the case, for instance, of the predicted probabilities for a university graduate to be over-educated seen in Chapter Three. As seen before, the values might be over-estimated, even if their general trends are still being valid.

4.4.3. Conditional predicted probabilities for a 25-64-year-old university graduate to be over-educated

Predicted probabilities for being over-educated have been calculated by gender, educational level, and age-group for each of the specifications. The objective is to test whether the probability of being over-educated varies, when the skill heterogeneity measured by observable cognitive skills, is included in the model. Tables from 4.17 to 4.19 show the predicted probabilities for a 25-64-year-old individual to be over-educated conditional on being employed (predicted probabilities for the Heckman two-step model when the dependent variable of the outcome equation is '1=over-educated' conditional on that the dependent variable of the selection equation is also '1=employed'). In all the cases, the proportion of over-educated has been measured using the statistical mean approach. Similar tables showing the predicted probabilities when the subjective educational mismatch measurement is used can be found in Appendix 3.

The predicted probabilities of success conditional on the success of the selection equation, that is, the predicted probability for a 25-64-year-old individual to be employed and over-educated conditional on being already employed reveal some differences among groups. Different values are found in the probability of being over-educated depending on the education mismatch measurement used.

Table 4. 17. Predicted probabilities for a university graduate being over-educated in the Spanish labour market by gender, age-group, and educational level. BS

Pr(Over-educated=1, Employed=1)/Pr(Employed=1)

		Lower than university graduates			University graduates		
		Pred. Prob.	Std. Err.	Sig. Level	Pred. Prob.	Std. Err.	Sig. Level
Men	25-34	0.0799	0.0139	***	0.3523	0.0278	***
	35-44	0.0610	0.0097	***	0.3167	0.0289	***
	45-54	0.0416	0.0099	***	0.2767	0.0295	***
	55-64	0.0320	0.0124	*	0.3063	0.0471	***
Women	25-34	0.0619	0.0119	***	0.3525	0.0300	***
	35-44	0.0460	0.0093	***	0.3109	0.0338	***
	45-54	0.0297	0.0089	**	0.2661	0.0322	***
	55-64	0.0203	0.0092	*	0.2744	0.0458	***

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Table 4. 18. Predicted probabilities for a university graduate being over-educated in the Spanish labour market by gender, age-group, and educational level. LS

Pr(Over-educated=1, Employed=1)/Pr(Employed=1)

		Lower than university graduates			University graduates		
		Pred. Prob.	Std. Err.	Sig. Level	Pred. Prob.	Std. Err.	Sig. Level
Men	25-34	0.0821	0.0145	***	0.3436	0.0268	***
	35-44	0.0629	0.0099	***	0.3072	0.0295	***
	45-54	0.0439	0.0103	***	0.2702	0.0301	***
	55-64	0.0356	0.0141	*	0.3072	0.0452	***
Women	25-34	0.0653	0.0131	***	0.3491	0.0312	***
	35-44	0.0490	0.0098	***	0.3071	0.0340	***
	45-54	0.0324	0.0092	**	0.2657	0.0322	***
	55-64	0.0234	0.0106	*	0.2827	0.0449	***

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Table 4. 19. Predicted probabilities for a university graduate being over-educated in the Spanish labour market by gender, age-group, educational level. NS

Pr(Over-educated=1, Employed=1)/Pr(Employed=1)

		Lower than university graduates			University graduates		
		Pred. Prob.	Std. Err.	Sig. Level	Pred. Prob.	Std. Err.	Sig. Level
Men	25-34	0.0822	0.0145	***	0.3455	0.0269	***
	35-44	0.0613	0.0096	***	0.3065	0.0289	***
	45-54	0.0430	0.0102	***	0.2705	0.0296	***
	55-64	0.0345	0.0139	*	0.3084	0.0445	***
Women	25-34	0.0678	0.0136	***	0.3530	0.0307	***
	35-44	0.0495	0.0098	***	0.3085	0.0330	***
	45-54	0.0331	0.0092	**	0.2689	0.0313	***
	55-64	0.0238	0.0108	*	0.2892	0.0454	***

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Despite the different value on the predicted probabilities regarding the education mismatch measurement used, some general features for the results for both measurements can be summarised as follows. First, there is a higher probability for men to be over-educated in each age-group or educational level, although the difference here appears to be rather small. For instance, the predicted probabilities of being over-educated when the statistical mean approach is used for 25-34-year-old university graduate men is 33 percent whereas it is 32 percent for their female counterparts.

Second, there is a higher likelihood for university graduates to be over-educated than those with lower educational levels for each age-group and gender. The probability of being over-educated is lower than 10 percent for those holding lower educational levels than a university degree, whereas it is always above 20 percent and sometimes above 30 percent for those individuals with university education when the statistical mean approach is used. For the subjective measurement (Appendix 3), the probabilities are higher than 25 percent for both educational levels, but they are even higher for those holding a university degree.

Third, there is little increase in the probability of being over-educated for those with lower educational levels than a university degree when the literacy and numeracy skills are considered, whereas it decreases for those holding a university degree. These results indicate the higher effect that accounting for cognitive skills has on the education mismatch for those individuals with higher educational levels. A higher demand for those skills among highly skilled jobs may be the main cause of these results.

In brief, according to the results, in the Spanish labour market, the likelihood for a university graduate to be over-educated is not much influenced by their level of cognitive skills, and in particular, by their level of literacy and numeracy skills. The results are consistent, regardless of the model (probit or the Heckman two-step model), the measurement of over-education (statistical mean approach or subjective methods), or the type of cognitive skill (literacy or numeracy). The magnitude of result is also, quite similar, typically a change of around 2 percentage points: therefore, the probability of being over-educated changes only to a small extent when skill heterogeneity is taken into account. Therefore, the results tend to confirm the lower probability of being over-educated for those university graduates with higher observable cognitive skills. These results are in concordance with the Job Competition Model (Thurow, 1970) and the Assignment Theory (Tinbergen, 1956; Sattinger, 1980; Hartog, 1981).

This does not mean that in the Spanish labour market the level of skills is not taken into account. Maybe, literacy and numeracy are not the best indicators to assess the level of skills of the university graduates in the Spanish labour market. There are other skills that could also affect the probability of being over-educated that might be more important in the Spanish labour market itself, as for instance, social skills or problem solving abilities. Unfortunately, these abilities were not measured in the Spanish PIAAC survey and therefore they could not be included in this analysis.

4.5. Has skills heterogeneity conditioned the likelihood of a university graduate being over-educated in the Spanish labour market?

The main objective of this chapter is to measure to what extent accounting for the individual's skills heterogeneity affects the probability of being over-educated. It was expected to find a negative relationship between the cognitive skills and the probability of being over-educated. Assignment theory and Job-Matching Theory suggest that those individuals with higher cognitive skills will be the ones who will be more likely to be hired in the best (skilled) jobs, and therefore who will be less likely to be over-educated. In contrast, those university graduates with lower cognitive skills will be more likely to work in jobs which require lower skills, but as they hold higher formal education degrees, they will be classified over-educated, even if they are not really over-skilled. The idea behind this is that those individuals who are considered as over-educated, because they have been at school for more years than the ones required or needed for the job they are securing, were, in reality, low skilled workers. That is, they have the degree, the formal qualification, but they hold lower abilities than their counterparts. Or in other words, they will be considered as over-educated but not as over-skilled.

In this sense, it may be useful to recall Chevalier's distinction between the 'apparent' and the 'genuine' over-educated (Chevalier, 2003). The 'apparent' is the individual who does not really have the skills despite holding the educational level, whereas the 'genuine' will be the one that holds the educational and the skills level. Nevertheless, in both cases, they are working in a job where the educational level required or needed is lower than the one they formally hold.

Using the Spanish PIAAC data, this chapter has analysed *to what extent the individuals' observable skills heterogeneity affects the likelihood for a university graduate to be over-educated in the Spanish labour market.* In this sense, the chapter has two main contributions. First, and using the PIAAC data of five European countries (England, France, Italy, Sweden, and Spain), the variation in skills across countries is calculated. Using the classification of Esping-Andersen and Regini (1999) regarding the European labour markets, the pairwise comparisons across countries show big differences between Scandinavian and South European countries for both cognitive skills, whereas Anglo-Saxon and Continental Europe tend to be less substantial.

Second, and focusing on the Spanish labour market, a probit model and a variation of the Heckman two-step model are run in three different variations, including and not including the literacy and numeracy proficiency. Therefore, each one of the three specifications has been run four times: using probit model and using the Heckman two-step model, and each of the models using first data where the proportion of over-educated university graduates has been measured using the statistical mean approach and then using the subjective approach.

The comparison of the results of the three specification is used then to analyse whether the inclusion of cognitive skills reduces the likelihood of being over-educated, at least as far as university graduates are concerned. Nevertheless, and according to the results from the predicted probabilities of success, the impact that this effect has on the probability for a university graduate to be over-educated is very limited, independent from the gender and age-group.

Two key implications can be derived from this result. First, including only observable cognitive skills is not enough to consider skill heterogeneity among university graduates, at least, as regards the Spanish labour market. The results indicate that other unobservable factors should be affecting the likelihood for university graduates to be over-educated in Spain. This may be related to a higher demand for other sort of skills for university graduates such as team work competences or ability to adapt to new situations, which are usually more characteristic of highly skilled jobs.

Literacy and numeracy skills are basic cognitive skills required in people's everyday life. However, highly skilled jobs usually require more specific cognitive skills, for instance, the ability to do research (social, scientific, or even historical research) or to develop independent work (as an independent professional or as an employee) or to lead team work. Some of them are innate abilities, and other are the sort of abilities that one might learn in on-the-job trainings or with experience.

Second, a negative relationship between the literacy and numeracy skills and the probability of being over-educated has been found, even if it is very low. This latter result is therefore consistent with the Theory of Job Competition and the Assignment Theory. This may indicate that a precise measurement of the cognitive skills will increase its effect on the probability of being over-educated. In this sense, it would be useful, for instance, to have the proficiency scores for problem solving in technology-rich environments and reading components available for the Spanish case. It would likely help to address some of the weaknesses of accounting only for literacy and numeracy proficiency scores.

> *"Simulation is a particular type of modelling. Building a model well-recognized way of understanding the world: something we do all the time, but which science and social science has refined and formalized."*
>
> Gilbert and Troitzsch (2005, p. 2)

Chapter 5. The future employability of university graduates in the Spanish labour market

Knowing the tendency for more skilled workers to be adequately employed in highly skilled jobs, and, above all, knowing the tendency for younger generations to enter at university and for them to be employed in an adequate job could be used to design more adequate educational and active labour market policies by policy makers. Using data from the Spanish Labour Force Survey and population and demographic figures from the National Statistical Institute (INE), this chapter forecasts the number of over-educated university graduates by single ages, genders, and regions in the Spanish labour market for the next ten years, from 2019 to 2030.

5.1. Introduction

The previous chapters have shown how in Spain, as in most developed countries, the expansion of the supply of and demand for university studies were in parallel with the growth of the education mismatches in the labour market, and especially, over-education. Using data from the Spanish Labour Force Survey (SLFS) for the whole population, Chapter Two displays the main changes in the educational level of the supply of and demand for workers in the Spanish labour market in the last decades of the 20^{th} century and beginning of the 21^{st}. It depicted the increase in the educational level of the population and the decrease in the unskilled jobs supplied by the labour market, and especially for women.

Using also data from the SLFS, but focusing on the 25-64-year-old university graduates, Chapter Three analysed the probability for them to be over-educated in the Spanish labour market from 1995 to 2018. An increase in the probability of being over-educated has been found in the Spanish labour market during the last decades of the 20^{th} century and beginning of the 21^{st}. This result reflects the mismatch in the intensity of the growth of the supply of and demand for university graduates, suggesting the possibility that the Spanish labour market has not been able to accommodate the increasing supply of university graduates.

Finally, Chapter Four used data from the PIAAC to measure the impact that accounting for skill heterogeneity has on the probability for a 25-64-year-old university graduate to be over-educated. The results of the fourth chapter demonstrated that accounting for either literacy or numeracy skills heterogeneity is not enough to justify the difference in the likelihood for university-graduates to be over-educated either by gender or age-group.

Keeping in mind all these results, this fifth chapter aims at using all this information as a starting point to forecast the probability for university graduates to be over-educated in the Spanish labour market the next years, from 2019 to 2030. The objective is to forecast future trends for the supply of and demand for university graduates in order to achieve a better educational balance for them in the Spanish labour market.

Therefore, in order to offer a starting point for policy makers to design future educational or active labour market policies, this fifth chapter uses different scenarios to forecast the trend of the education mismatch for university graduates in the Spanish labour market.

Using different sources, I create here a sample with actual data for the Spanish population by gender, single ages and region. It also includes the proportion of individuals employed and not employed by educational level and occupation. This is the actual data used as a baseline to analyse the trend in the proportion of university graduates and in their tendency to be employed by sort of occupation (skilled or not skilled). Then, three scenarios are described to project the future trend of the parameters of the model in order to foresee the number of over-educated university graduate in the Spanish labour market from 2019 to 2030.

Traditionally mathematics has been used in social sciences, and especially in economics to formalise theories. It has already been used for the Classic Economic Theory. However, simulations and forecasts have spread in parallel with the use of computers. It started in the 1960s when the first computers where used at research institutes (Gilbert and Troitzsch, 2005). Improvements in the technology and the emergence of personal computers which are cheaper and increasingly powerful and easy to use propelled the use of computer simulations among professionals, researchers and students.

A simulation in social science aims at understanding a current real situation in order to be able to model the relationship among the variables implied and, therefore, to be able to apply mathematical methods to forecast the near future. However, this is not a simple job. According to Gilbert and Troitzsch (2005), simulation-based research should include the following six stages: first, it is important to delimitate the research variable, which variable it is and which future trend we would like to know. Second, some real data regarding the behaviour of this variable are needed to fix its past and actual trend. Third, and perhaps the most important part of the simulation, or at least the part which will make the difference in the outcome, is the model design. In this stage, the assumptions regarding the future trend of the model should be delimited. It is important here to be able to rationally justify the predicted behaviour of the independent variables. For instance, if I assume an increase in the number of students entering at university it should be, for example, due to (a) an increase in the number of young people because of, for instance, an increase in the birth rate in the previous years, or (b) a reduction in the university fees, or (c) more opportunities for them to access to university studies, e.g. a change in admissions policy or an increase in the supply of universities or university studies.

Following Gilbert and Troitzsch (2005), the last three stages of the simulation are directly related to check the robustness of the model. The fourth stage is about verifying that the model works. Thus, there are enough data for all variables and the model is consistent or it runs without any hindrance. Fifth, the results of the model should be validated, that is, the output of the model is consistent with the previous behaviour of the variable and the assumptions made. And sixth, there is a sensitivity analysis to understand the behaviour of the model when limited changes are included in the variables behaviour.

Therefore, this chapter attempts to offer a starting point for policy makers to identify the most critical points and to design the best educational and active labour market policies to combat them. The aim has already been delimited. It is to foresee the trend for the proportion of 25-64-year-old over-educated university graduates, by single ages, gender, and regions. As a result, the research variable is the number of university graduates working in a non-highly-skilled job.

According to Gilbert and Troizsch's stages of a simulated-based research, the first stage has, then, already been defined. The remaining stages will be developed in this chapter which is structured as follows: Section 5.2 describes the sources used to analyse the actual behaviour of the parameters used in the forecast. The methodology used to forecast the degree of over-education in the next years in the Spanish labour market is explained in Section 5.3. Section 5.4 presents the results of the forecast for the three scenarios, and Section 5.5 includes a discussion accounting for the effects that the unforeseen external shock of COVID19 is causing in the Spanish economy, and therefore, their effects on the Spanish labour market. Section 5.6 concludes the chapter.

5.2. Sources, data, and variables

The data analysis from the previous chapters has demonstrated the education mismatch problems of the Spanish labour market, and especially for 25-64-year-old university graduates. It was already evident in the 1990s, it has increased over time and especially during the economic crisis, and it is still noticeable in 2018. These previous results demonstrate that accounting for gender as well as regional differences is important when the educational mismatch is measured. Therefore, the forecast of this chapter will also account for these differences and the calculations and analysis will therefore be done using data for single ages, genders, and regions (NUTS-1).

The second stage of a simulated-based research is, according to Gilbert and Troitzsch (2005), to know about the past and current trend of the research variable. Given that this chapter aims at forecasting the proportion of over-educated university graduates in the Spanish labour market, the data used in this second stage should describe the actual trend for this variable in order to fix their recent behaviour. To do so, this chapter uses different data from the National Statistics Institute (INE).

Single individual data accounting for the characteristics of over-educated university graduates in the Spanish labour market has not been found. Therefore, individual data from the SLFS from 2014-2018 is used here to create a dataset in which each observation refers to a single age, gender, and region (NUTS-1). For instance, the identification number (Id) 180 is for the number of men who were 25 years old in 2016, that is, those males who were born in 1991 and living in the economic region of Madrid. Therefore, the same Id (180) is for the number of men who were 26 years old in 2017 and living in the economic region of Madrid, and for those men living in the economic region of Madrid who in 2018 were 27 years old. As a result, for each observation (Id), I have categorised the number of individuals by educational level, and the number of university graduates employed, and whether they are working in a highly skilled job. Table 5.1 summarises the actual data for Spain included in the sample.

Table 5. 1. Population and university graduates by age and year. Spain, 2014-2018

Year	Men		Women	
Population				
Year	Number	% out of total population	Number	% out of total population
2014	13,116,312	0.50	13,086,934	0.50
2015	13,044,699	0.50	13,042,651	0.50
2016	12,968,842	0.50	13,011,651	0.50
2017	12,932,310	0.50	13,000,419	0.50
2018	12,949,414	0.50	13,042,298	0.50
University graduates				
Year	Number	% out of total men	Number	% out of total women
2014	1,800,112	13.72	2,080,880	15.90
2015	1,855,286	14.22	2,147,550	16.47
2016	1,932,608	14.90	2,256,235	17.34
2017	1,995,469	15.43	2,357,600	18.13
2018	2,049,187	15.82	2,474,631	18.97
University graduates employed				
Year	Number	% out of university graduate men	Number	% out of university graduate women
2014	1,497,962	83.21	1,590,237	76.42
2015	1,569,849	84.61	1,670,085	77.77
2016	1,598,156	82.69	1,781,061	78.94
2017	1,729,843	86.69	1,890,705	80.20
2018	1,781,103	86.92	2,003,924	80.98
University graduates employed in highly skilled jobs				
Year	Number	% out of university graduate men employed	Number	% out of university graduate women employed
2014	1,096,432	73.19	1,067,657	67.14
2015	1,125,101	71.67	1,110,808	66.51
2016	1,171,262	73.29	1,194,928	67.09
2017	1,231,310	71.18	1,265,982	66.96
2018	1,241,905	69.73	1,294,163	64.58

Source: SLFS (2nd quarters) (INE).

Actual data from the previous table show the gender equal distribution of the population and its low variability. Regarding the number of university graduates, two main features can be depicted: first, the constant year-on-year increase, and second, the higher proportion of university graduates among women than men although the variation is low (between 2 and 3 percentage points). For instance, according to the data, in 2018, almost 19 percent of 25-64-year-old women were university graduates, whereas for men this percentage was 15.8. In contrast, a higher proportion of university graduate men than women were employed in the Spanish labour market. In 2018, for example, these percentages were 86.9 for men and 80.9 for women. Moreover, the proportion of university graduates matched in highly skilled jobs were also higher for men than for women. Data from the previous table show that in 2018 69.7 percent of university graduate men employed were employed in a highly skilled job, that is, they were adequately matched, whereas for women this percentage was 64.6. These results are in concordance with the results of the previous chapters which indicate the higher probability for women to be over-educated than men.

On the website of INE, population and demographic statistics and datasets are available. In particular, I have used INE's population projection for the years from 2019 to 2030. All these data can be found by single ages, gender, and regions. Therefore, I do not hypothesise regarding birth rates, mortality rates and possible internal or external migration, and take the official projections data as reasonable and applicable. Hence, using as a baseline the distribution and tendency of the official demographic projections, I can focus on the supply of and demand for university graduates in the Spanish labour market in the near future, that is in ten years' time.

Table 5. 2. Population by gender and year. Spain, 2014-2030

Year	Men		Women	
	N	%	N	%
2014	13,116,312	0.50	13,086,934	0.50
2015	13,044,699	0.50	13,042,651	0.50
2016	12,968,842	0.50	13,011,651	0.50
2017	12,932,310	0.50	13,000,419	0.50
2018	12,949,414	0.50	13,042,298	0.50
2019	13,023,462	0.50	13,114,170	0.50
2020	13,008,442	0.50	13,127,234	0.50
2021	12,979,117	0.50	13,140,779	0.50
2022	12,940,803	0.50	13,137,127	0.50
2023	12,897,000	0.50	13,124,674	0.50
2024	12,852,956	0.49	13,113,089	0.51
2025	12,808,833	0.49	13,096,547	0.51
2026	12,767,522	0.49	13,084,074	0.51
2027	12,723,320	0.49	13,065,293	0.51
2028	12,678,129	0.49	13,044,869	0.51
2029	12,624,136	0.49	13,009,239	0.51
2030	12,576,914	0.49	12,979,244	0.51

Year	Variation rate	
	Men	Women
2014-15	-0.55	-0.34
2015-16	-0.58	-0.24
2016-17	-0.28	-0.09
2017-18	0.13	0.32
2018-19	0.57	0.55
2019-20	-0.12	0.10
2020-21	-0.23	0.10
2021-22	-0.30	-0.03
2022-23	-0.34	-0.09
2023-24	-0.34	-0.09
2024-25	-0.34	-0.13
2025-26	-0.32	-0.10
2026-27	-0.35	-0.14
2027-28	-0.36	-0.16
2028-29	-0.43	-0.27
2029-30	-0.37	-0.23

Source: From 2014 to 2018 SLFS (2nd quarters) (INE), and from 2019 to 2030 the population projection from INE (http://www.ine.es)

Table 5.2 shows the population by gender and year, and their year-on-year variation for each gender from 2014 to 2030. Actual data have been depicted in blue and data from INE's projection in black. The year-on-year variation is rather low each year, and showing a downward trend in most of the years. Moreover, the highest decrease is in the number of men, provoking an increase in the proportion of 25-64-year-old women living in Spain. Nevertheless, the data used in this chapter is disaggregated data. As has already been explained I have created a dataset where the proportion of individuals by educational level, etc. are divided by gender, single ages, and regions (NUTS-1).

Table 5. 3. NUTS-1 and NUTS-2 correspondence classification

NUTS-1	NUTS-2
North-West	Asturias
	Cantabria
	Galicia
North-East	Aragon
	Navarra
	Basque Country
	La Rioja
Madrid	Madrid

NUTS-1	NUTS-2
Central area	Castile and Leon
	Castile-La Mancha
	Extremadura
East	Catalonia
	Valencia
	Balearic Islands
South	Andalusia
	Murcia
Canary Islands	Canary Islands

Source: https://ec.europa.eu/eurostat/web/nuts/national-structures-eu

Regarding regions, INE uses the more disaggregated geographical classification of NUTS-2 (17 regions), but then it can be grouped in NUTS-1 (7 regions) using the correspondence table found on the Eurostat webpage (Table 5.3). The main reason for using NUTS-1 as a geographical classification is to have significant data for each of the observations. Even if there are no real 'individual' data, it is classified by regions, gender, and single ages, therefore in some ages the number of observations can be low. This geographical classification is the same as the one used in the previous chapters, and using a more aggregated geographical classification also makes the presentation of the results easier to understand.

The total sample includes, therefore, 20,483 observations with 1,078 different 'individuals'. The data including all information for 'individuals' (number of university graduates, number of university graduates employed, and number of university graduates employed in highly skilled jobs) are from 2014 to 2018; whereas the number of total population by single ages, gender and regions is available from the SLFS from 2014 to 2030. Actual data from the SLFS from 2014 to 2018, and forecast data from the projection of population from INE from 2019 to 2030.

In the following sections the figures showing the projected data will also depict the real data. In these, the upward tendency that each one of the variables (number of university graduates, number of university graduates employed, and number of university graduates employed in highly skilled jobs) has recently followed can be clearly identified. Moreover, the forecast is also run backwards to 2007 to check the consistency of the projection's results and actual data.

5.3. Scenarios and methodology

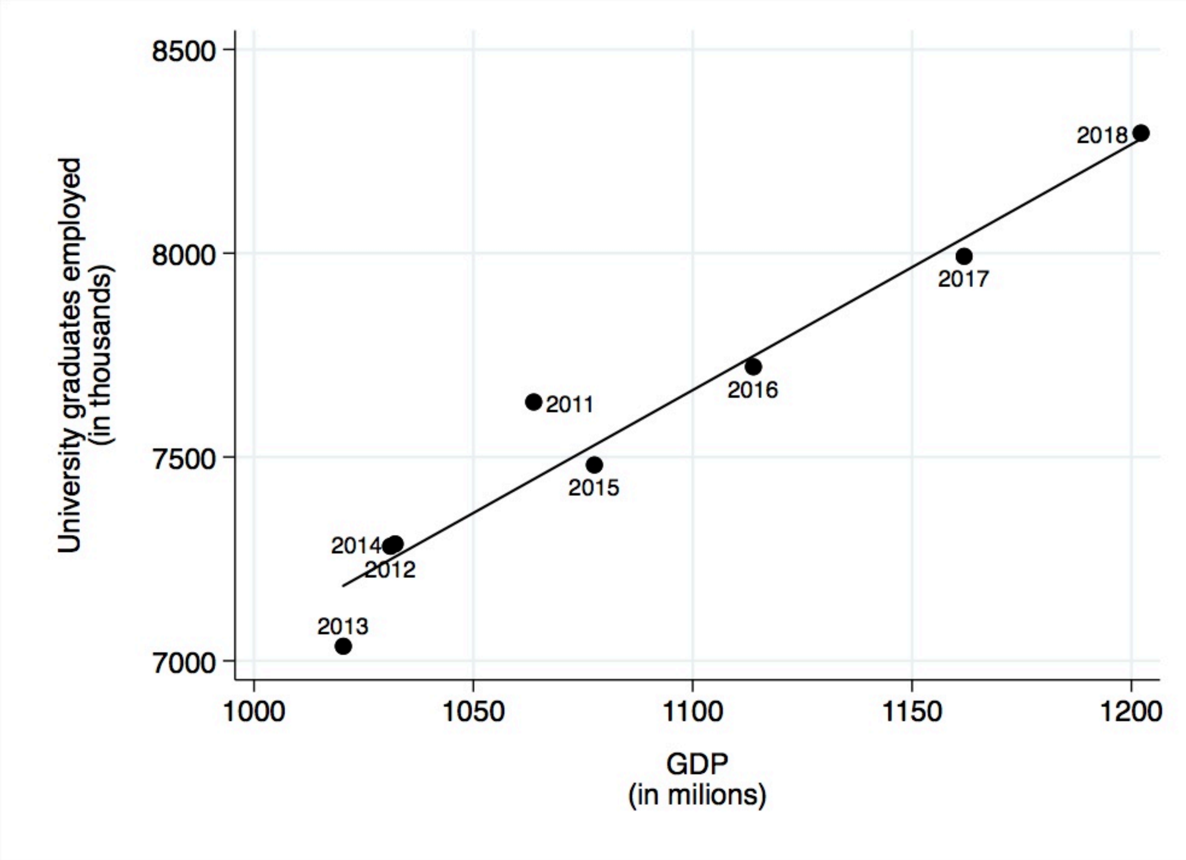
According to Gilbert and Troitzsch (2005), and as was stated in the Introduction of this chapter, this section is the cornerstone of the simulation. It does not mean that the remaining parts of the process are not important, however, the decisions taken here are the ones that will determine the outcome and therefore the validity and the suitability of the results; although foreseeing the near future for demographic, economic or social variables is always a good practice.

Generally speaking, economic, social, and population simulations or projections tend to be cautious and a little hazardous. The main reason is that it is difficult to foresee both the direction and the real effect of an external shock on the population or on the economy. Therefore, a forecast is usually derived from the real data and the current or the planned economic and social policies. Besides, the outcome of the economic and social policies needs some time to be reflected in the real world. Therefore, it seems sensible to opt for more conservative suppositions, especially when a short period of time is foreseen.

Nevertheless, sometimes reality strikes with totally unforeseen events which could alter the real data completely, and therefore the results of the simulation. However, even when this occurs, the model may still be valid if a new scenario is included. For instance, in the Spanish case, in population projections of the 90s, nobody was expecting that the number of people migrating to the country would increase as it did at the end of the 90s, and even less that the upward trend would last for such a long time. The main reason was that Spain had traditionally been a country of emigration and not the other way round. Nevertheless, it happened and the number of people living in Spain and participating in the formal and informal labour market increased, since most of the migrations were driven by economic causes and therefore they were people of working age.

Overall, using the projection of the population of the National Statitstical Institute (INE) as baseline data, this projection considers three scenarios. All three scenarios use the estimated coefficients of a regression model including the real data from 2014 to 2018 estimated by gender and region, as initial data. Then, Scenario 1 assumes a quite steady tendency and supposes that the last two years' average tendency for each of the variables and for each of the ages, genders, and regions remains for the remaining years. Therefore, it adds to the initial data the last two years' average tendency (2017-2018) of the number of university graduates (U1), university graduates employed (W1), and university graduates employed in highly skilled jobs (HSJ1) by gender and region to foresee the following years, from 2019 to 2030.

Scenario 2 opts for the same hypothesis as Scenario 1 as regards to the number of university graduates (U1) and the number of university graduates employed in highly skilled jobs (HSJ1). Therefore, the difference between Scenarios 1 and 2 is on the number of university graduates employed (W). For them, this second scenario uses the last year-on-year average of the Gross Domestic Product (GDP) of each region to increase the number of university graduates employed. The hypothesis is that the variation in the number of university graduates employed will be half of the year-on-year average variation of the GDP of each region (W2). Therefore, in this case, regional differences are considered, but not gender, since there are no gender differences in the GDP data. Even so, in this case, slight regional differences appear, since year-on-year variation on the GDP in 2017-2018 was positive and around 3 percent in most of the Spanish regions. Only the Madrid area and the East region had higher variations (4.2 and 6.6 percent, respectively, see Table 5.5).



Y (University graduates employed)= 1031.51 + 6.0297 GDP

R^2=0.9451 Adj. R^2=0.9359

Source: Spanish National Accounts. http://www.ine.es

Figure 5. 1. Gross Domestic Product and university graduates employed

Figure 5.1 shows the positive relationship between GDP and the number of university graduates employed in Spain during the last decade. Even if the data does not confirm a causal relationship, it shows that there is some relationship between economic expansion valued as the variation in the GDP growth and the number of university graduates employed in the labour market. Indeed, for most of the years, the dots are scattered around the line. Only data from 2011 fall above the line, showing that for this year lower levels of GDP are related to a higher number of university graduates employed compared to the other years. The variation in the components of the GDP as well as the composition of this employment should be deeply analysed in order to be able to extract solid conclusions from this data. However, taking into account the general behaviour of the rest of the years, the confirmation of this positive relationship between GDP and the number of university graduates employed is enough for the aim of this research.

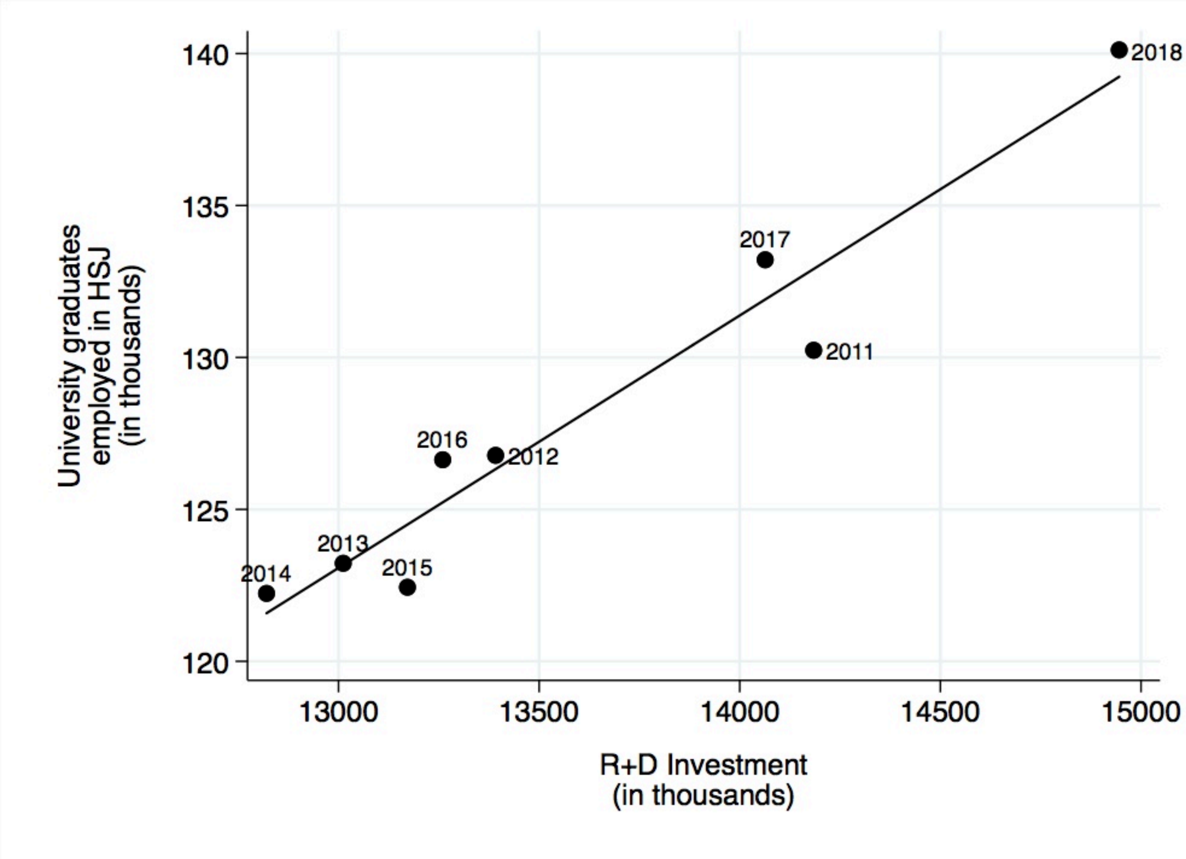
Scenario 3 is, indeed, a variation of Scenario 2, and uses the same hypothesis for the number of university graduates (U1) and the number of university graduates employed (W2), but it adds a variation in the number of university graduates employed in highly skilled jobs (HSJ2). In this third scenario, therefore, half of the proportion of Research and Development (R+D) investment out of the GDP has been considered as a percentage of variation for the number of university graduates employed in highly skilled jobs. The idea behind it is that R+D investment allows the firms to get a skill biased technological change which in parallel is expected to increase the demand for university graduate workers. This latter hypothesis seems sensible, at least in the Spanish case. Using data from 1980 to 2000 De Juan Asenjo and López Santiago (2004) analysed the effect that technological change had on the occupational distribution in Spain. They concluded that the technological change was skewed promoting the professional and technical jobs. More recently, Gonzalez Pandiella (2014) also conclude an increase in the creation of jobs in the Spanish labour market, and especially, jobs highly skill demanding. Besides, according to the following figure (Fig. 5.2) in more recent years this positive relationship between R+D investment and number of university graduates employed in highly skilled jobs is still in effect.

It has already been seen that the technological change may increase the demand for skilled or highly skilled workers, but it might also increase the demand for low skilled workers, all depending on whether it is a capital or a work intensive technology. However, R+D investment implies the work of highly skilled jobs in order to systematically develop experimental research.[95] Besides, the translation of this new knowledge to modify the production system will also require highly skilled workers. Therefore, it does not seem unreasonable to suppose a positive relationship between the variation in R+D investment and the demand for university graduate workers.

Recent actual data show a positive relationship between R+D investment and university graduates employed in highly skilled jobs (Fig. 5.2). As in the GDP case, the causal relationship cannot be demonstrated here. However, data show that in

[95] https://ec.europa.eu/eurostat/statistics-explained/index.php/R_%26_D_expenditure

recent years, higher R+D investment has been related to higher employment rates in highly skilled jobs for university graduates.



Y (*University graduates employed in HSJ*)= 15.066 + 0.0083 *RD Investment*
R^2=0.9387 Adj. R^2=0.9285

Source: Statistics on R+D Activities. http://www.ine.es

Figure 5. 2. Research and Development investment and university graduates employed in highly skilled jobs

Previous figures show the general behaviour of the variables for the whole of Spain and some regional differences appeared when data is split in regions (NUTS-1) (Table 5.5 and 5.7). As a result, the most favourable scenario regarding the situation for university graduates being adequately employed in the Spanish labour market varies depending on the region. However, in general terms, when the results for Spain are considered, the most favourable scenario is the third one. The second scenario simulates an unchanged tendency in the number of university graduates out of the population (U1) and an upward tendency in the number of university graduates employed (W2) and in the number of university graduates employed in highly skilled jobs (HSJ2). Results from Scenario 3 are better than the second one, because the average tendency in the last years for all the regions has been the reduction in the number of university graduates being employed in highly skilled jobs. Therefore, the third scenario offers a higher reduction in the number of over-educated university graduates, and in their proportion out of all university graduates, and out of university graduates employed.

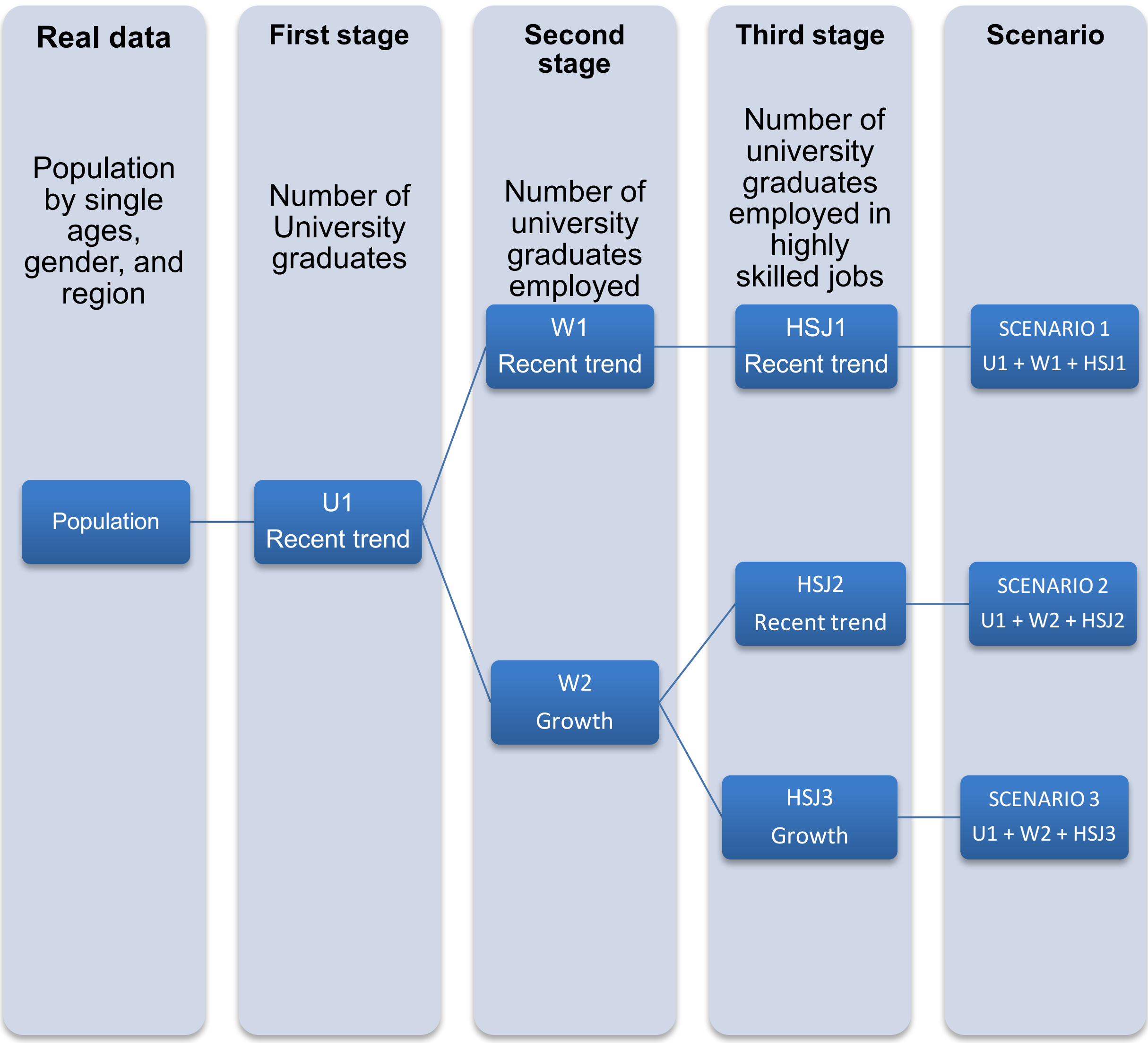


Figure 5. 3. Scenarios and hypothesis for the projection

The projection is built into three stages; each one of them has one or two hypotheses (Fig. 5.3). In the first stage, the number of university graduates by age, gender, and region is calculated for each year of the projection (2019-2030) (U). The results of this first stage are used in the second stage to calculate the number of university graduates employed in the Spanish labour market by age, gender, and region for each of the years considered (W). The third and final stage is the calculation for the number of university graduates working in a highly skilled job in the Spanish labour market (HSJ). Each hypothesis is calculated separately by single ages, genders and regions (NUTS-1), and the regional and gender outcome is finally grouped to obtain a national result: the proportion of university graduates adequately matched in the Spanish labour market from 2019 to 2030. The projection is also run backwards to 2017 in order to see whether the results are consistent with the real data. The connexion between the real data and the projected data can be seen in the figures presented in the next section, where the real data and the projected data are depicted in each of the figures.

Following therefore the structure of the previous figure (Fig. 5.3) where the scenarios and hypothesis of the projection are divided by stages, the formalisation of the calculations is presented in the next paragraphs.

Stage 1

Using the real data for the 25-64-year-old population by single ages, gender, and region (NUTS-1) and the number of university graduates by single ages, gender, and region, the following model has been regressed for each gender and region:

$$U_i = \beta_{Pop_i} * Pop_i + \beta_i * age_i + \beta_i * Year_i \quad [5.1]$$

Where,

U_i is the number of university graduates, Pop_i is the number of individuals, age_i is the age of the 'individual', and $Year_i$ are dummy variables, using the first year (2014) as a reference category. And,

i = {North-West; North-East; Region of Madrid; Central Area; East; South; Canary Islands}, accounts for region (NUTS-1)

For each hypothesis, the estimated coefficients of these linear regressions [5.1] for each region have been applied to the population data from the demographic projection of INE to get the number of university graduates from 2017 to 2030. In general terms, the estimated coefficients are positive for population and year variables, whereas they are negative for age. These results suggest, as expected, an increase in the number of university graduates by time (in recent years the proportion of university graduates out of all the population has increased every year) and fewer individuals holding university studies in older generations.

Table 5. 4. Variation in the number of university graduates by gender and region. Spain, 2017-2018 (in percentages)

Region	Men	Women
North-West	4.96	-0.22
North-East	-0.41	5.13
Madrid	1.36	-1.06
Central area	1.25	6.24
East	6.43	10.22
South	2.54	5.97
Canary Islands	-1.60	1.79
Spain	**2.08**	**4.01**

Source: SLFS (2n quarters; 2017-2018), INE

Then, the variation of the last two years in the number of university graduates has been added to the general tendency in order to adjust the model (Table 5.4). Important differences across variables and regions, but also between genders arise from the actual data. See, for instance, data from the East area, where the number of university graduate men has increased more than 6 percent and the number of university graduate women more than 10 percent (Table 5.4), but the number of university graduate men employed has slightly reduced (Table 5.5), and the number of university graduate men and women adequately matched educationally has also reduced (Table 5.7).

The result is the projected number of university graduates in the Spanish labour market by single ages, gender, region, and year (Section 5.4). As the three scenarios share the same hypothesis regarding the estimated trend for the number of university graduates (U1), these projected data from Stage 1 will be used as the baseline data for the estimation of Stages 2 of all three scenarios.

Stage 2

In the second stage, the number of university graduates estimated in Stage 1 (U1) will be used to project the number of university graduates employed by gender, age, region, and year.

As in the previous stage, the first step is to regress the real data, and it can be formally summarised as follows:

$$W_i = \beta_{U_i} * U_i + \beta_{Pop_i} * Pop_i + \beta_i * age_i + \beta_i * Year_i \quad [5.2]$$

Where,

W_i is the number of university graduates employed, and U_i is the number of university graduates, Pop_i is the number of individuals, age_i is the age of the 'individual', and $Year_i$ are dummy variables, using the first year (2014) as a reference category. And,

i = {North-West; North-East; Region of Madrid; Central Area; East; South; Canary Islands}, accounts for region (NUTS-1)

The estimated coefficients of the regression equation [5.2] are used in Scenario 1 to calculate the projected number of university graduates employed (W1). On the whole, the estimated coefficients show a positive relationship for the number of university graduates and the population, and a negative relationship between the variable age and the number of university graduates employed. Year dummy variables show a higher variability in the signs of the estimated coefficients. They vary depending on the year and region considered.

Therefore, the projected number of university graduates of Stage 1 (U1) is used here as the observable data to estimate the number of university graduates employed from 2017 to 2030. Furthermore, in Scenario 1, the Spanish economic structure is assumed not to suffer big changes and the number of university graduates employed in the labour market will be increasing or decreasing in the same proportion as in the last two years. As a result, it is expected that the tendency of the recent two years (2017-2018) will be maintained for the next ten years (Table 5.5).

Table 5. 5. Variation in the number of university graduates employed by gender and region. Spain, 2017-2018 (in percentages)

Region	Men	Women
North-West	1.00	0.33
North-East	0.77	3.76
Madrid	1.89	1.93
Central area	1.22	0.74
East	-0.04	0.70
South	1.99	-0.01
Canary Islands	0.89	3.70
Spain	1.10	1.59

Source: SLFS (2n quarters; 2017-2018), INE

Again, gender and region differences are displayed in Table 5.5. In this case, regional differences are more evident than in the previous tables. For instance, the number of female university graduates in the North-East area and in the Canary Islands has increased more than 3 percent, whereas the number of male university graduates in these regions has increased less than 1 percent.

In Scenarios 2 and 3, an increase in the number of university graduates demanded by the labour market is expected (W2). In particular, it is expected that in each region the increase in the number of university graduates will be the same for both genders, but different by region, since it is expected to follow half of the proportion of the year-on-year variation of the GDP. As has already been seen above, the year-on-year variation of the GDP has been positive and higher than 3 percent in the last years for all regions. The region of Madrid and the East area are the zones with the highest values being the year-on-year variation on the GDP: 4.2 percent in the region of Madrid and 6.6 percent in the East area (Table 5.6).

Table 5. 6. Variation in the GDP by region. Spain, 2017-2018 (in percentages)

Region	GDP
North-West	3.60
North-East	3.40
Madrid	4.20
Central area	3.20
East	6.60
South	3.00
Canary Islands	3.20
Spain	**3.89**

Source: Spanish National Accounts. http://www.ine.es

The outcome of both calculations from Stage 2 is used in Stage 3 to find the number of university graduates employed in highly skilled jobs in the Spanish labour market from 2017 to 2030.

Stage 3

The third and last stage is the projection of the number of university graduates being adequately educationally matched in the Spanish labour market that is working in a highly skilled job (HSJ). Again, using first the real data from 2014 to 2018, the estimated coefficients for the following regression equation are calculated and used then using the estimated values for W1 and W2 as observable data. Formally, the regression model used in this third stage can be summarised as:

$$HSJ_{ijk} = \beta_{UG_{ijk}} * W_{ijl} + \beta_{Pop_i} * Pop_i + \beta_{ijk} * age_{ijk} + \beta_{ijk} * Year_{ijk} \quad [5.3]$$

Where,

HSJ_i is the number of university graduates employed in highly skilled jobs, W_i is the number of university graduates employed, Pop_i is the number of individuals, age_i is the age of the 'individual', and $Year_i$ are dummy variables, using the first year (2014) as a reference category. And,

i = {North-West; North-East; Region of Madrid; Central Area; East; South; Canary Islands}, accounts for region (NUTS-1)

Assuming that in the next years the tendency in the variation of the number of university graduates employed in highly skilled jobs will follow the same tendency as the recent years (HSJ1) and using the number of university graduates employed from Scenario 1 stage 2 (W1), I project, in the first scenario, the outcome for the number of university graduates educationally matched in the Spanish labour market.

In this case, the estimated coefficients from the regression model [5.3] suggest a positive relationship between the independent variables and the number of university graduates employed in a highly skilled job, except for the dummy variables accounting for time (year).

Table 5. 7. Variation in the number of university graduates employed in HSJ (adequately matched) by gender and region. Spain, 2017-2018 (in percentages)

Region	Men	Women
North-West	-0.46	-0.53
North-East	0.56	-1.15
Madrid	-0.18	-1.88
Central area	-1.98	0.45
East	-1.83	-2.65
South	-0.37	-1.68
Canary Islands	-5.91	-6.14
Spain	-1.45	-1.94

Source: SLFS (2n quarters; 2017-2018)

In this case, during the last two years, the variation in the number of university graduates employed has been negative in most of the regions. This is one of the reasons why in those regions where this negative variation is higher the second scenario will report better results (lower levels of over-education among 25-64-year-old university graduates) than the third.

From here, I derive the number of over-educated university graduates and their proportion out of the total number of university graduates and total number of university graduates employed for Scenario 1. The total number of university graduates is divided by the university graduates who are employed and those who are unemployed. From those employed there is one part who are adequately educationally matched, those who are working in a highly skilled job and those who are over-educated, since they are working in a job which requires lower educational level or fewer years of schooling (Fig. 5.4). It has been seen in previous chapters that the probability for a university graduate to be under-educated is nearly null. Therefore, here I have divided the employed university graduates between those who are adequately educationally matched and those who are over-educated.

Alternatively, Scenarios 2 and 3 use the number of university graduates employed from Scenario 2 and Stage 2 (W2) as observable data for the period from 2017 to 2030. The difference between both scenarios is on the hypothesis regarding the number of university graduates being employed in a highly skilled job (HSJ) in the Spanish labour market (Stage 3). Scenario 2 uses the same hypothesis as Scenario 1 (HSJ1) (Table 5.7), whereas Scenario 3 accounts for a possible increase in the demand for highly skilled jobs (HSJ2) (Table 5.8). As a result, Scenario 3 becomes the most favourable scenario for reducing the proportion of mismatch in the Spanish labour market.

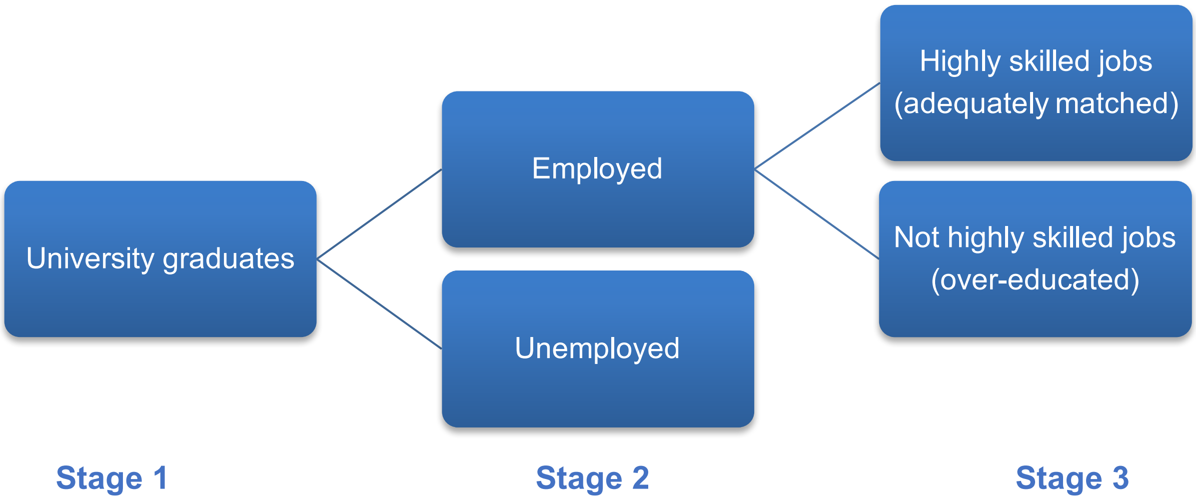


Figure 5. 4. Outline of the projection results

The hypothesis used in Scenario 3 is based on the positive relationship between R+D investment and university graduates employed in highly skilled jobs seen in Figure 5.2. As a consequence, the year-on-year variation of the R+D investment is used here to project the number of university graduates employed in highly skilled jobs. Table 5.8 summarised its year-on-year variation. It shows important differences by region.

Table 5. 8. Variation on the R+D investment. Spain, 2017-2018 (in percentages)

Region	I+D
North-West	0.82
North-East	1.31
Madrid	1.66
Central area	0.85
East	0.93
South	0.93
Canary Islands	0.47
Spain	**1.00**

Source: Statistics on R+D Activities. http://www.ine.es

Once the number of university graduates, the number of university graduates employed and the number of university graduates employed in highly skilled jobs have been projected, we can calculate the number of university graduates over-educated in the next years. Subtracting the number of university graduates employed in a highly skilled job from the total number of university graduates employed we obtain the number of university graduates who are over-educated (Fig. 5.4). We found these data by regions, genders, and single ages. And dividing the number of over-educated university graduates by the number of employed university graduates we got the proportion of university graduates who are over-educated.

Formally:

$$\text{Number of university graduates over-educated (NHSJ)} = W - HSJ$$

$$\text{Proportion of university graduates over-educated} = NHSJ / W$$

It is important to remember here, that all these calculations are derived from the population's projection of INE, which accounts for the possible internal and external migration. In this sense, I have not included extra possible population or labour force movement across regions, all of which are assumed to have been included in the population projection from INE.

5.4. Over-education in the Spanish labour market: Future trends

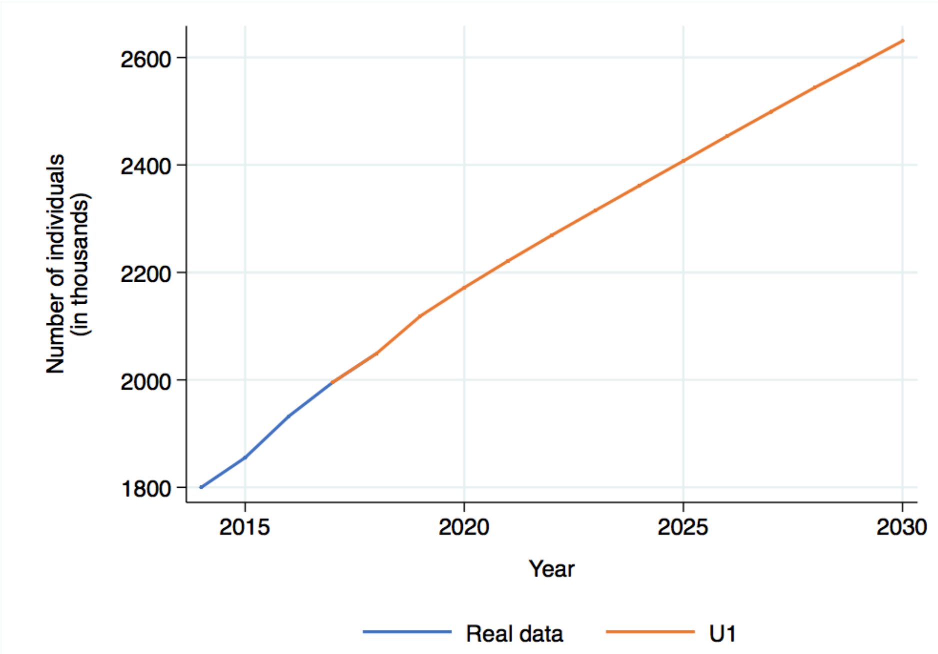
Once the research variable has been selected, its recent behaviour analysed, and the hypotheses and scenarios of the forecast designed, the results of the calculations should be analysed to validate the model. Besides, the results should also be validated to see the consistency of the model. In this sense, this section depicts the results of the forecast. The projection has also been run backwards two years (2017 and 2018) to check the consistency of the results.

In order to account for the differences across regions, which, as has already been seen, are quite significant, the forecast introduces regional variations in the hypothesis of the future tendencies used in each of the scenarios (Tables 5.4 to 5.8). However, in this section, the results are depicted for Spain as a whole and only gender differences have been included in order to make the discussion clearer. More detailed results by region can be found in Appendix 4 and some of these will also be introduced in the discussion to help understand the general tendency for Spain.

I will first discuss the outcome for each of the stages of the forecast to go through the combined results, and therefore the number and the proportion of university graduates who are over-educated in the Spanish labour market.

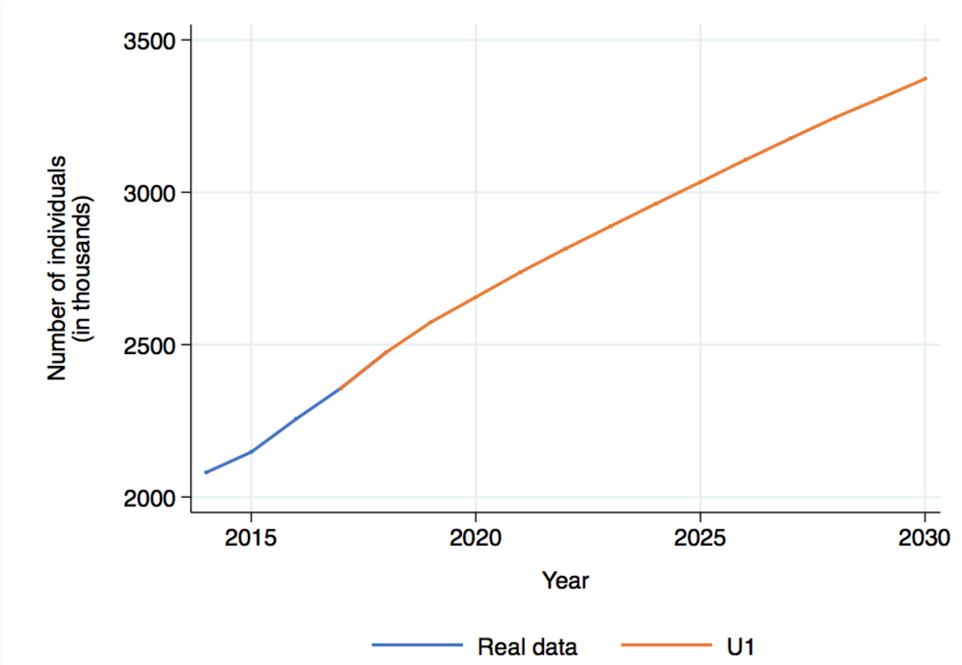
5.4.1. Stage 1: Results

Figures 5.5 and 5.6 show the upward tendency projected for the number of university graduates in Spain, for men and women, respectively. One unique hypothesis has been used in all three scenarios, therefore the figures show only one line (U1). Table 5.9 shows an increase in the number of university graduate men and women during all the period. However, the yearly variation rate shows a decrease in the intensity of this growth, being higher for women for both periods, actual and forecast data.



Note: U1: Number of 25-64-year-old university graduate men projected following Hypothesis 1 (Scenarios 1 to 3)
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Figure 5. 5. 25-64-year-old university graduate men in Spain (in thousands)



Note: U1 is the number of 25-64-year-old university graduate women projected following Hypothesis 1 (Scenarios 1 to 3).
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Figure 5. 6. 25-64-year-old university graduate women in Spain (in thousands)

The upward trend followed by the number of university graduates, both men and women, seen in the previous figures and tables is the combination of the expected growth of the population (number of individuals living in Spain) and the positive average trend for the last two years regarding the variation on the number of university graduates (Table 5.4). Furthermore, the data also show a higher increase in the number of university graduates for women than for men. Again, this is the consequence of a higher increase in the number of women than men, in the population, and in the number of university graduates. Regarding population, in 2014, 49 percent of the total population from 25 to 64 years old were women. The projection of the population from INE expects that in 2030, in Spain, there will be 51 percent of women in the 25-64-year-old population. Regarding the number of university graduates, in 2014, there were in Spain 3.9 million university graduates (46 percent of whom were men), there were 4.6 million in 2019 (45 percent were men), and, according to this forecast, it is expected that there will be 5.9 million in 2030 (44 percent men).

Table 5. 9. University graduates by gender and year. Spain, 2014-2030.

Year	Men		Women		Year	Yearly variation rate	
	N	%	N	%		Men	Women
2014	1,800,112	0.46	2,080,880	0.54	2014-15	3.07	3.20
2015	1,855,286	0.46	2,147,550	0.54	2015-16	4.17	5.06
2016	1,932,608	0.46	2,256,235	0.54	2016-17	3.25	4.49
2017	1,995,469	0.46	2,357,600	0.54	2017-18	2.69	4.96
2018	2,049,187	0.45	2,474,631	0.55	2018-19	3.40	4.01
2019	2,118,958	0.45	2,573,981	0.55	2019-20	2.49	3.19
2020	2,171,682	0.45	2,656,142	0.55	2020-21	2.29	3.10
2021	2,221,425	0.45	2,738,519	0.55	2021-22	2.16	2.80
2022	2,269,319	0.45	2,815,190	0.55	2022-23	2.04	2.61
2023	2,315,705	0.44	2,888,701	0.56	2023-24	1.99	2.54
2024	2,361,695	0.44	2,962,117	0.56	2024-25	1.94	2.41
2025	2,407,447	0.44	3,033,614	0.56	2025-26	1.92	2.40
2026	2,453,690	0.44	3,106,345	0.56	2026-27	1.85	2.26
2027	2,499,057	0.44	3,176,434	0.56	2027-28	1.80	2.18
2028	2,544,151	0.44	3,245,696	0.56	2028-29	1.68	1.92
2029	2,586,937	0.44	3,308,136	0.56	2029-30	1.69	1.94
2030	2,630,698	0.44	3,372,175	0.56			

Note: Actual data in blue and the projection's results in black
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030

The consistency of the projected tendency with the real data can be seen in the previous figures, only identified by the colour, since the tendency of the line is still the same. The yearly variation rate shown in Table 5.9 depicts the difference in the slope of both lines, for men and for women.

It is also important to bear in mind that the projected data also looks backwards for two years, therefore, data from 2017 to 2018 are represented twice in the figure, since only if the real and the projected data are not consistent will two different lines appear.

Table 5.10 shows the effect that the increase in the number of university graduates will have on the population. According to the forecast data, the proportion of university graduates out of the population will increase in the coming years in Spain. Concretely, it will increase more than 5 percentage points from 2019 to 2030 for both genders (from 16.3 in 2019 to 20.9 in 2030 for men, and from 19.6 to 26.0 for women).

Table 5. 10. University graduates out of the population by gender and year. Spain, 2014-2030 (in percentages)

Year	Men	Women
2014	13.7	15.9
2015	14.2	16.5
2016	14.9	17.3
2017	15.4	18.1
2018	15.8	19.0
2019	16.3	19.6
2020	16.7	20.2
2021	17.1	20.8
2022	17.5	21.4
2023	18.0	22.0
2024	18.4	22.6
2025	18.8	23.2
2026	19.2	23.7
2027	19.6	24.3
2028	20.1	24.9
2029	20.5	25.4
2030	20.9	26.0

Note: Actual data in blue and the projection's results in black
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

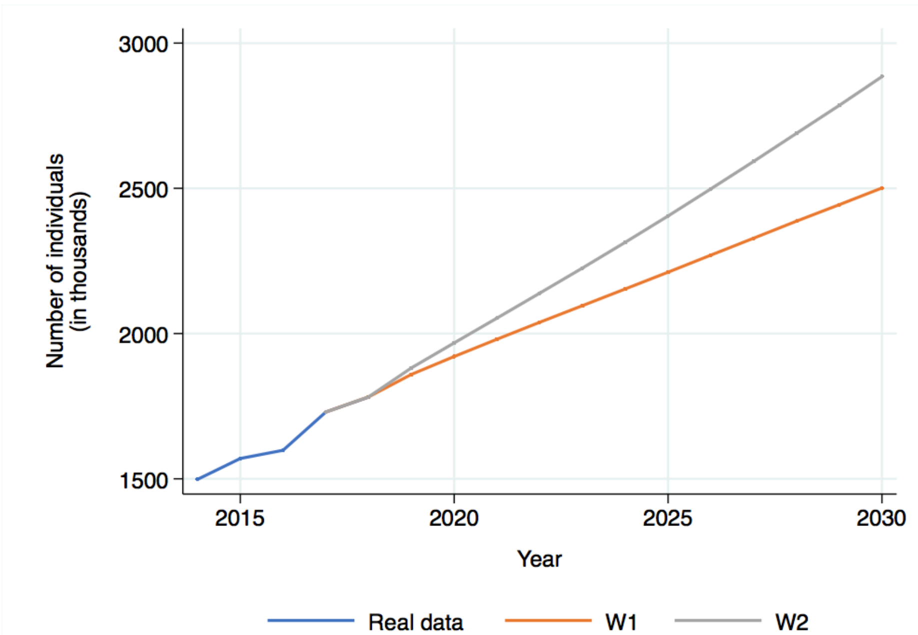
Given this constant increase expected in the number of university graduates in Spain, it seems highly likely that the proportion of over-educated university graduates in the Spanish labour market will continue being high, unless a significant growth in the number of university graduates employed, and especially employed in highly skilled jobs occur.

Data in Appendix 4 also show the upward trend in the number of university graduates in most of the regions. According to the data, the number of university graduate men will decrease in the North-East, the Central area and in the Canary Islands; whereas for women, the number of university graduates will follow a downward tendency in the North-West area and in the region of Madrid. In the other regions the number of university graduate women also follows an upward trend.

All in all, the overall tendency for Spain is an increase in the number of university graduates, and therefore, an increase in their supply in the Spanish labour market (Fig. 5.5 and 5.6).

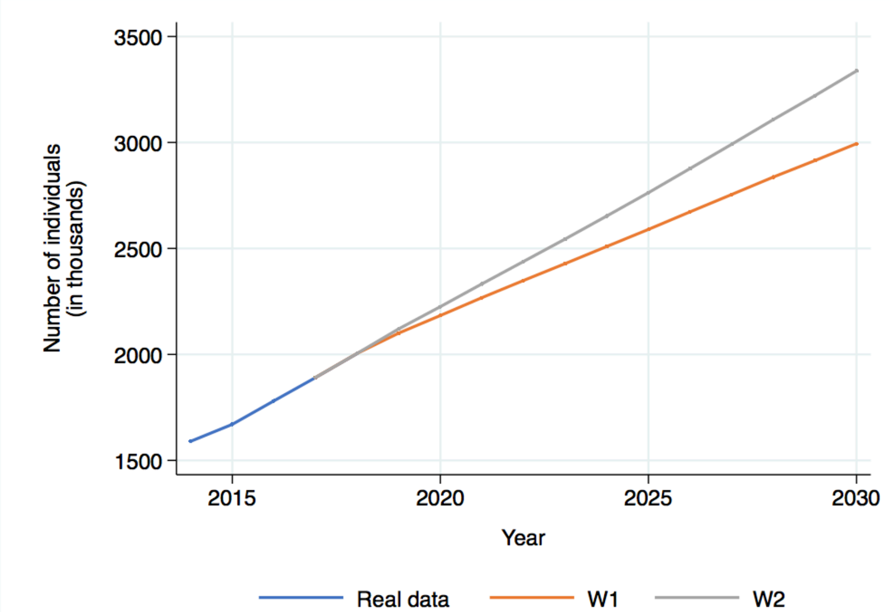
5.4.2. Stage 2: Results

Figures 5.7 and 5.8 focus on the outcome of the second stage of the forecast and depict the results for the calculations for the number of university graduates employed for men and women, respectively. W1 accounts for the number of university graduates employed when the hypothesis of continuity in the tendency of the last two years (2017-2018) is considered (Scenario 1), and W2 depicts the number of university graduates employed when an increase in the demand for workers in the Spanish labour market is supposed to be equal to half of the year-on-year variation of the GDP of each region (Scenarios 2 and 3).



Note: W1 is the number of 25-64-year-old university graduate men employed projected following Hypothesis 1 (Scenario 1), and W2 is the number of 25-64-year-old university graduate men employed projected following Hypothesis 2 (Scenarios 2 and 3).
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Figure 5. 7. 25-64-year-old university graduate men employed in the Spanish labour market (in thousands)



Note: W1 is the number of 25-64-year-old university graduate women employed projected following Hypothesis 1 (Scenario 1), and W2 is the number of 25-64-year-old university graduate women employed projected following Hypothesis 2 (Scenarios 2 and 3).
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Figure 5. 8. 25-64-year-old university graduate women employed in the Spanish labour market (in thousands)

As expected, W2 gives a more favourable scenario in the Spanish labour market than W1. This is true for Spain and for most of the regions, except for those 25-64-year-old university graduate women living in the North-East and in the Canary Islands, and for 25-64-year-old university graduate men living in the South area (Appendix 4). The reason is that, unlike the other regions, in these two regions the growth in the number of university graduates employed in the labour market during the last two years (W1) has been higher than half of the year-on-year variation of the GDP in these regions (Tables 5.5 and 5.6), a measurement that has been used to calculate W2.

Table 5.11 depicts the results for W1 and Table 5.12 for W2. At a national level, the number of university graduates employed projected by applying the second hypothesis (a variation associated with the GDP year-on-year variations) imply a higher increase in the employability of the university graduates. These results suggest that in the coming years the Spanish labour market will be more able to create new jobs. Whether these jobs are or are not highly skilled jobs will be seen in the following stage of the forecast. Meanwhile, the forecast data suggest a higher increase in the number of university graduate employed women than men, in both cases, W1 and W2. This is in concordance with the results from Stage 1 where a more intense growth in the number of university graduates was found for women.

Table 5. 11. University graduates employed by gender and year. Spain, 2014-2030 (Hypothesis 1 – W1)

Year	Men		Women	
	N	%	N	%
2014	1,497,962	0.49	1,590,237	0.51
2015	1,569,849	0.48	1,670,085	0.52
2016	1,598,156	0.47	1,781,061	0.53
2017	1,729,843	0.48	1,890,705	0.52
2018	1,781,103	0.47	2,003,924	0.53
2019	1,859,310	0.47	2,100,654	0.53
2020	1,920,981	0.47	2,183,567	0.53
2021	1,980,426	0.47	2,267,970	0.53
2022	2,038,892	0.46	2,348,954	0.54
2023	2,096,174	0.46	2,428,572	0.54
2024	2,153,601	0.46	2,509,387	0.54
2025	2,211,292	0.46	2,589,865	0.54
2026	2,269,981	0.46	2,672,679	0.54
2027	2,328,282	0.46	2,754,462	0.54
2028	2,386,772	0.46	2,836,857	0.54
2029	2,443,436	0.46	2,914,456	0.54
2030	2,501,103	0.46	2,994,342	0.54

Year	Yearly variation rate	
	Men	Women
2014-15	4.80	5.02
2015-16	1.80	6.64
2016-17	8.24	6.16
2017-18	2.96	5.99
2018-19	4.39	4.83
2019-20	3.32	3.95
2020-21	3.09	3.87
2021-22	2.95	3.57
2022-23	2.81	3.39
2023-24	2.74	3.33
2024-25	2.68	3.21
2025-26	2.65	3.20
2026-27	2.57	3.06
2027-28	2.51	2.99
2028-29	2.37	2.74
2029-30	2.36	2.74

Note: Actual data in blue and the projection's results in black
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Table 5. 12. University graduates employed by gender and year. Spain, 2014-2030 (Hypothesis 2 – W2)

Year	Men		Women	
	N	%	N	%
2014	1,497,962	0.49	1,590,237	0.51
2015	1,569,849	0.48	1,670,085	0.52
2016	1,598,156	0.47	1,781,061	0.53
2017	1,729,843	0.48	1,890,705	0.52
2018	1,781,103	0.47	2,003,924	0.53
2019	1,881,752	0.47	2,120,719	0.53
2020	1,967,715	0.47	2,225,298	0.53
2021	2,053,132	0.47	2,332,988	0.53
2022	2,139,052	0.47	2,438,740	0.53
2023	2,225,703	0.47	2,544,591	0.53
2024	2,314,185	0.47	2,653,240	0.53
2025	2,404,684	0.47	2,763,119	0.53
2026	2,497,973	0.46	2,877,030	0.54
2027	2,592,550	0.46	2,991,372	0.54
2028	2,689,067	0.46	3,107,888	0.54
2029	2,785,154	0.46	3,220,626	0.54
2030	2,884,454	0.46	3,337,438	0.54

Year	Yearly variation rate	
	Men	Women
2014-15	4.80	5.02
2015-16	1.80	6.64
2016-17	8.24	6.16
2017-18	2.96	5.99
2018-19	5.65	5.83
2019-20	4.57	4.93
2020-21	4.34	4.84
2021-22	4.18	4.53
2022-23	4.05	4.34
2023-24	3.98	4.27
2024-25	3.91	4.14
2025-26	3.88	4.12
2026-27	3.79	3.97
2027-28	3.72	3.90
2028-29	3.57	3.63
2029-30	3.57	3.63

Note: Actual data in blue and the projection's results in black
Source: SLFS (2^{nd} quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

The proportion of university graduates employed out of university graduates are presented in Table 5.13. This table is used to check whether the increase in the number of jobs in the Spanish labour market will be enough to employ all university graduates. In general terms, and according to the results of the following table, the creation of jobs in the Spanish labour market will be more intense than the increase in the number of university graduates. As a result, the data show an increase in the proportion of university graduates employed. There are even some years at the end of the period of the second hypothesis where the proportion is higher than 100. According to these data, it is expected that the demand for workers will exceed its supply. However, taking into account that I have only considered here the employment for university graduates, these data can be interpreted as there are enough jobs for all the population who holds a university graduate, but this does not mean they are adequately matched, as it will be shown in the following tables and figures.

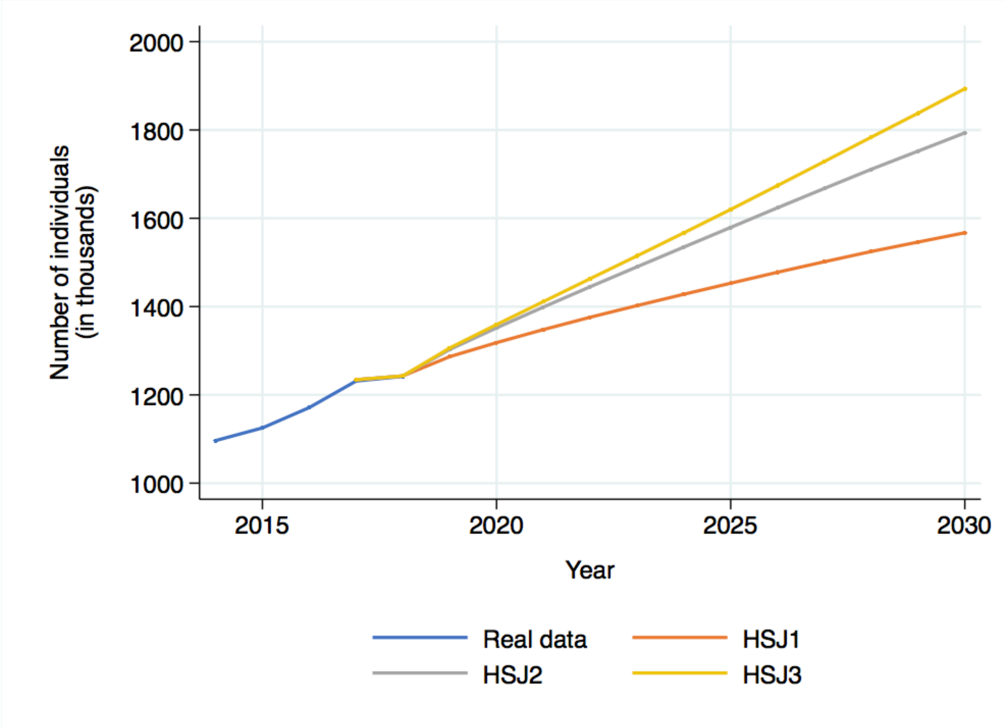
Table 5. 13. University graduates employed out of university graduates by gender, year, and hypothesis. Spain, 2014-2030 (in percentages)

Year	Men		Women	
	W1	W2	W1	W2
2014	83.2	83.2	76.4	76.4
2015	84.6	84.6	77.8	77.8
2016	82.7	82.7	78.9	78.9
2017	86.7	86.7	80.2	80.2
2018	86.9	86.9	81.0	81.0
2019	87.7	88.8	81.6	82.4
2020	88.5	90.6	82.2	83.8
2021	89.2	92.4	82.8	85.2
2022	89.8	94.3	83.4	86.6
2023	90.5	96.1	84.1	88.1
2024	91.2	98.0	84.7	89.6
2025	91.9	99.9	85.4	91.1
2026	92.5	101.8	86.0	92.6
2027	93.2	103.7	86.7	94.2
2028	93.8	105.7	87.4	95.8
2029	94.5	107.7	88.1	97.4
2030	95.1	109.6	88.8	99.0

Note: Actual data in blue and the projection's results in black
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

5.4.3. Stage 3: Results

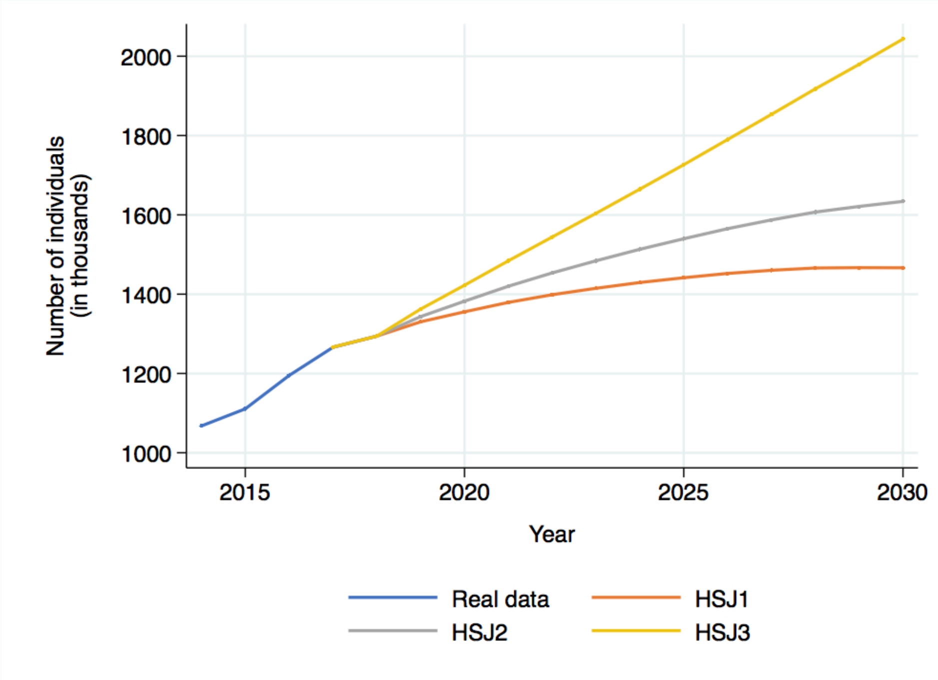
The next two figures (Fig. 5.9 and 5.10) depict the output of the third stage. The combination of the results of the two previous stages jointly with the two hypotheses for the number of university graduates employed in highly skilled jobs (HSJ) gives the three final outcomes for this third stage. As in the previous stages, Figure 5.9 shows the result for university graduate men, and Figure 5.10 for university graduate women. The results for each of the regions can be found in Appendix 4.



Note: HSJ1 is the number of 25-64-year-old university graduate men employed in highly skilled jobs projected following Hypothesis 1 (Scenario 1), and HSJ2 is the number of 25-64-year-old university graduate men employed in highly skilled jobs projected following Hypothesis 2 (Scenario 2), and HSJ3 is the number of 25-64-year-old university graduate men employed in highly skilled jobs projected following Hypothesis 3 (Scenario 3).
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Figure 5. 9. 25-64-year-old university graduate men employed in HSJ in the Spanish labour market (in thousands)

For Spain, the three scenarios show an upward tendency in the number of university graduates adequately matched in the labour market. Scenario 1 is the least favourable, since it is the one with the lower number of university graduates employed in highly skilled jobs, followed by the second scenario. Therefore, Scenario 3 is the most favourable to reduce the number of university graduates who are over-educated. This is the scenario where the variables calculated in Stage 1 (U) follow the tendency of the last two years (U1), whereas the number of university graduates employed (W), which is determined in Stage 2, is supposed to increase at the same rate as half of the year-on-year GDP variation (W2), and the number of university graduates educationally matched in the labour market (HSJ) is supposed to grow at the same rate as half of the investment in R+D (HSJ2). This is true for both genders, although the trend projected by Scenario 3 is higher for women (Fig. 5.10).



Note: HSJ1 is the number of 25-64-year-old university graduate women employed in highly skilled jobs projected following Hypothesis 1 (Scenario 1), and HSJ2 is the number of 25-64-year-old university graduate women employed in highly skilled jobs projected following Hypothesis 2 (Scenario 2), and HSJ3 is the number of 25-64-year-old university graduate women employed in highly skilled jobs projected following Hypothesis 3 (Scenario 3).
Source: SLFS (2^{nd} quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Figure 5. 10. 25-64-year-old university graduate women employed in HSJ in the Spanish labour market (in thousands)

Therefore, at this point of the analysis the result is uncertain, since the three variables are growing in the three scenarios. The result will, hence, depend on the intensity of the increase, the slope of the line of each of the variables. Table 5.14 anticipates the result. Showing the proportion of university graduates employed in a highly skilled job, it depicts the reduction in this proportion during the entire period, for all hypotheses and for both genders. As expected, Scenario 3 is the most favourable for both men and women; and higher proportions of university graduate men than women are employed in highly skilled jobs.

These results suggest that the tendency found in the previous chapters, where women are usually more likely to be more over-educated will probably continue in the next years, unless significant active labour market policies are introduced to switch the recent behaviour of the Spanish labour market.

Table 5. 14. University graduates employed in a HSJ (adequately educationally matched) by gender, year, and hypothesis. Spain, 2014-2030 (in percentages).

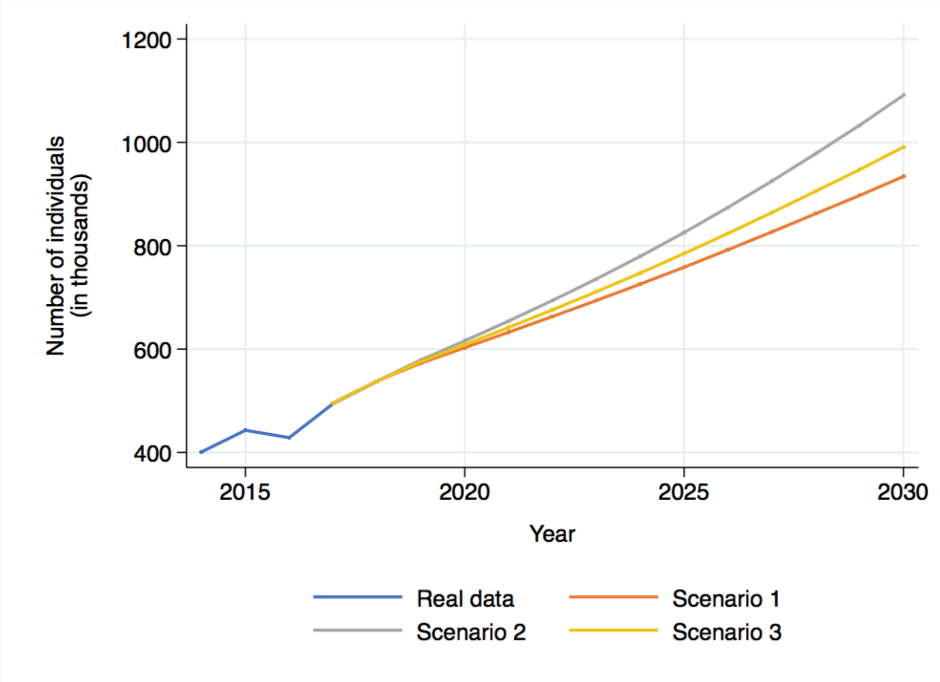
Year	Men			Women		
	HSJ1	HSJ2	HSJ3	HSJ1	HSJ2	HSJ3
2014	73.2	73.2	73.2	67.1	67.1	67.1
2015	71.7	71.7	71.7	66.5	66.5	66.5
2016	73.3	73.3	73.3	67.1	67.1	67.1
2017	71.2	71.2	71.2	67.0	67.0	67.0
2018	69.7	69.7	69.7	64.6	64.6	64.6
2019	69.2	69.2	69.4	63.3	63.3	64.2
2020	68.6	68.7	69.1	62.1	62.1	63.9
2021	68.1	68.1	68.7	60.8	60.9	63.6
2022	67.5	67.6	68.4	59.5	59.6	63.3
2023	66.9	66.9	68.1	58.3	58.3	63.0
2024	66.3	66.3	67.7	57.0	57.0	62.8
2025	65.7	65.7	67.4	55.7	55.7	62.5
2026	65.1	65.0	67.0	54.3	54.4	62.2
2027	64.5	64.3	66.7	53.0	53.1	62.0
2028	63.9	63.6	66.3	51.7	51.7	61.7
2029	63.3	62.9	66.0	50.3	50.3	61.5
2030	62.7	62.2	65.6	49.0	49.0	61.2

Note: Actual data in blue and the projection's results in black
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

5.4.4. Forecast results

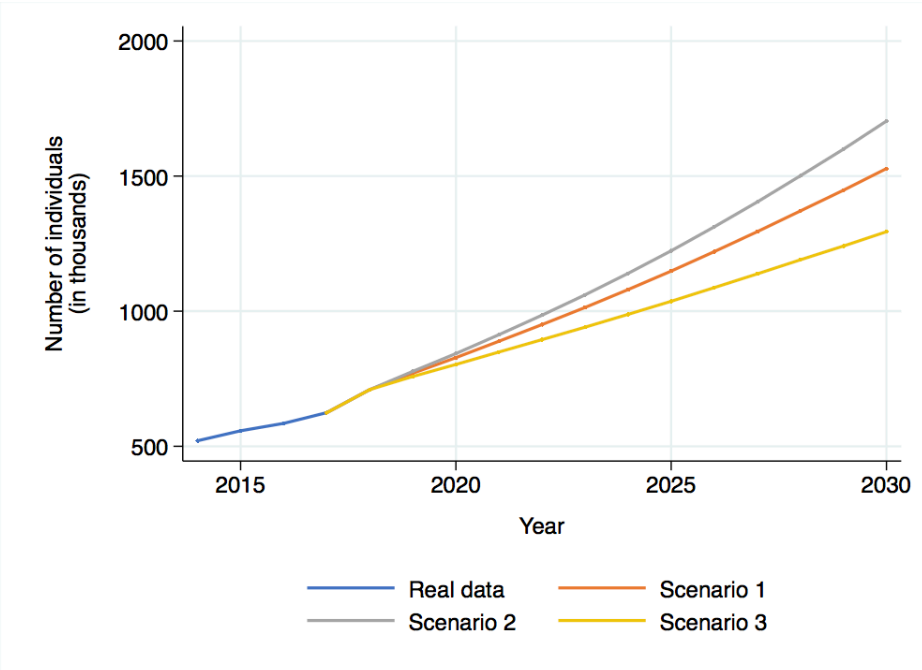
The next figures (Fig. 5.11 to 5.14) draw the results of the projection. Figure 5.11 depicts the number of 25-64-year-old university graduate men who are over-educated, and Figure 5.12 is the same variable but for women. Figures 5.12 and 5.13 display the proportion of 25-64-year-old university graduate over-educated out of the number of university graduates for men and women, respectively.

According to data from the first two figures (5.11 and 5.12), in any scenario, 25-64-year-old university graduate men will succeed in reducing the number of over-educated, and the general trend is upward both for men and for women. The gender differences are marked by the scenarios. For women, the lowest number of over-educated university graduate women is when Scenario 3 is accomplished, whereas for men the best results are for Scenario 2.



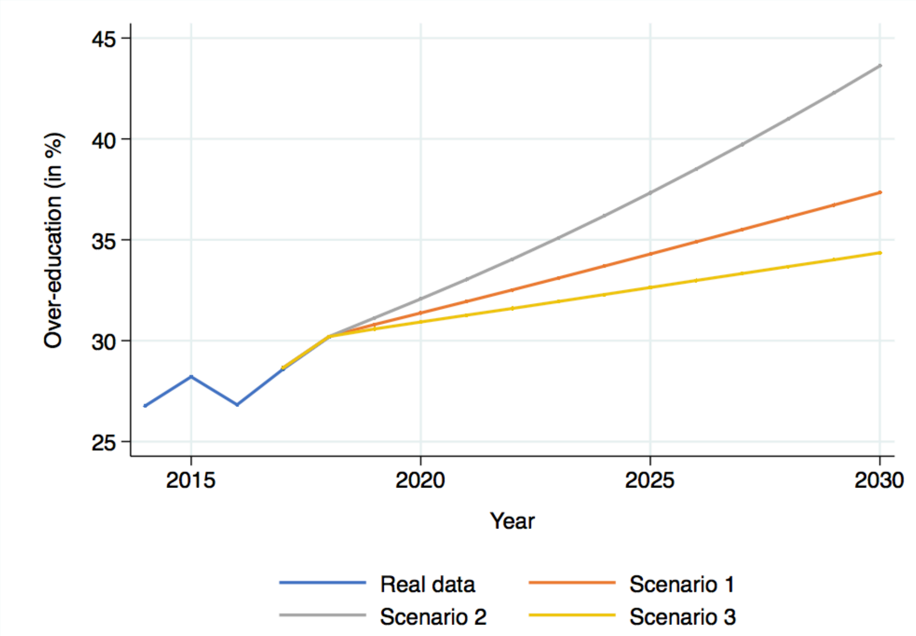
Note: Scenario 1=U1-HSJ1; Scenario 2=U1-HSJ2; Scenario 3=U1-HSJ3
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Figure 5. 11. 25-64-year-old over-educated university graduate men in the Spanish labour market (in thousands)



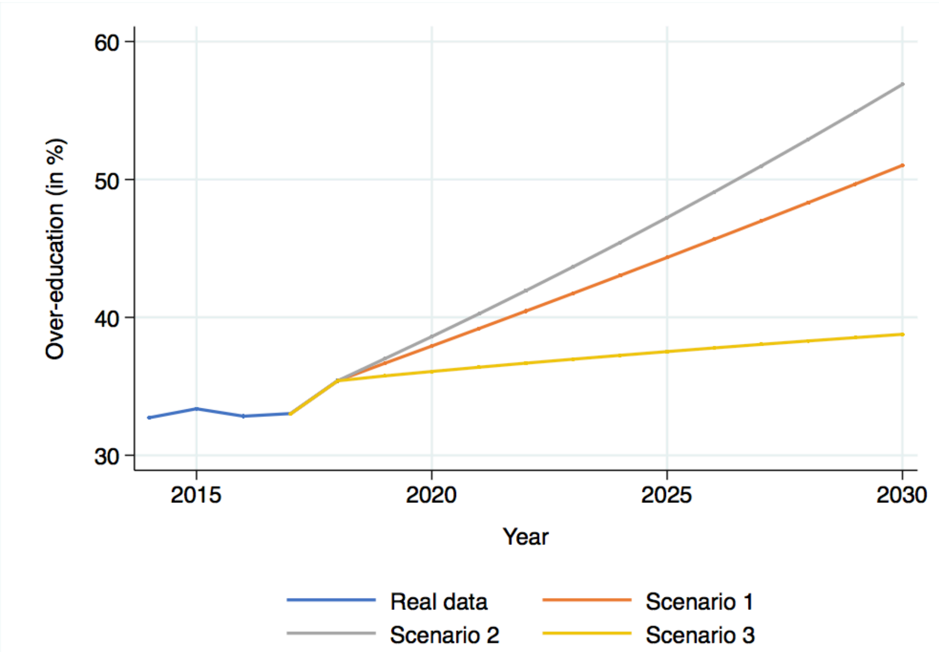
Note: Scenario 1=U1-HSJ1; Scenario 2=U1-HSJ2; Scenario 3=U1-HSJ3
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Figure 5. 12. 25-64-year-old over-educated university graduate women in the Spanish labour market (in thousands)



*Note: Scenario 1=((U1-HSJ1)/U1)*100; Scenario 2=((U1-HSJ2)/U1)*100; Scenario 3=((U1-HSJ3)/U1)*100*
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Figure 5. 13. 25-64-year-old over-educated university graduate men in the Spanish labour market (in percentage)



*Note: Scenario 1 = ((U1-HSJ1)/U1)*100; Scenario 2 = ((U1-HSJ2)/U1)*100; Scenario 3 = ((U1-HSJ3)/U1)*100*
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Figure 5. 14. 25-64-year-old over-educated university graduate women in the Spanish labour market (in percentage)

Better results appear in Figures 5.13 and 5.14, since the slope of the lines, even though still positive, is lower than for Figures 5.11 and 5.12. This indicates that the proportion of over-educated university graduates will grow more slowly than the number of university graduates employed. The reason for these results is that the number of university graduates employed is also expected to grow, as seen in the previous figures and tables.

The lower increase in the proportion of over-educated will occur for women if Scenario 3 is achieved, that is an increase in the number of university graduates employed equal to half of the year-on-year variation in the GDP and an increase in the number of university graduates employed in highly skilled jobs equal to half of the year-on-year variation in R+D investment (Fig. 5.13). However, even this outcome stands for a slight increase in the proportion of university graduate women over-educated. The following table depicts these proportions.

Table 5. 15. Over-educated university graduates out of university graduates employed by gender, year, and scenario. Spain, 2014-2030 (in percentages)

Year	Men			Women		
	Scenario 1	Scenario 2	Scenario 3	Scenario 1	Scenario 2	Scenario 3
2014	26.8	26.8	26.8	32.9	32.9	32.9
2015	28.3	28.3	28.3	33.5	33.5	33.5
2016	26.7	26.7	26.7	32.9	32.9	32.9
2017	28.8	28.8	28.8	33.0	33.0	33.0
2018	30.3	30.3	30.3	35.4	35.4	35.4
2019	30.8	30.8	30.6	36.7	36.7	35.8
2020	31.4	31.3	30.9	37.9	37.9	36.1
2021	31.9	31.9	31.3	39.2	39.1	36.4
2022	32.5	32.4	31.6	40.5	40.4	36.7
2023	33.1	33.1	31.9	41.7	41.7	37.0
2024	33.7	33.7	32.3	43.0	43.0	37.2
2025	34.3	34.3	32.6	44.3	44.3	37.5
2026	34.9	35.0	33.0	45.7	45.6	37.8
2027	35.5	35.7	33.3	47.0	46.9	38.0
2028	36.1	36.4	33.7	48.3	48.3	38.3
2029	36.7	37.1	34.0	49.7	49.7	38.5
2030	37.3	37.8	34.4	51.0	51.0	38.8

Note: Actual data in blue and the projection's results in black
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

A similar tendency is followed for each of the regions (Appendix 4). Nevertheless, it is worthy highlighting here the decrease in the proportion of over-educated male university graduates for some scenarios in the North-East area (Scenario 1 and 2), the region of Madrid (Scenario 3), and the South (Scenario 3). This trend is followed by women in Scenario 3 in the North-East area, the region of Madrid, and the Central area.

On balance, the previous results validate the model and its consistency for the next years. Besides, they indicate that some efforts in the Spanish labour market are needed in order to increase the demand for university graduates, since it is expected that their supply will continue growing. The current Spanish labour market is not offering enough highly skilled jobs. Furthermore, data show that the inadequacy between the supply of and the demand for university graduates is even worse for women, showing less favourable scenarios.

These results show the general education mismatch in the Spanish labour market. However, real data including the field of study will help to obtain a more accurate forecast. It may improve the results regarding the areas of study more demanded by the labour market and whether a variation in the demand for specific fields of study for university degrees should be promoted or, on the contrary, their supply should be reduced. Variations in the supply of university studies may help to promote or discourage the demand for specific fields of study and this could help to reduce the field of study imbalance in the labour market.

5.5. An unexpected external shock

Recently European economies in general, and the Spanish economy in particular, have seemed to follow a favourable course. Despite declines in the real GDP growth rate from 2016 to 2019, it was still positive and, this was expected to continue the coming year. However, the COVID19 pandemic has been a huge shock to all economies world-wide. The International Monetary Fund (IMF) has projected an 8 percent fall in the variation of the Spanish GDP for 2020,[96] whereas the *Banco de España* (the Spanish banking authority) expects an even higher negative effect of the pandemic, a fall of 12.4 percent to 13.6 percent, if the lockdown lasts 12 weeks or more.[97]

It is expected that the pandemic crisis will not last for long, but there is considerable uncertainty regarding its effects on the real economy and how long its effects will impact the productive system. Trying to forecast a new scenario while still inside this highly uncertain situation does not seem very sensible. First, we do not know yet how long this lockdown will last and consequently for how long firms will have to bear its impact on their production system. Besides, it is likely that unlike the last economic crisis of 2008, this one will not affect only those less productive firms. Second, it is likely (and desirable) that this first shock effect which has reduced the number of workers contributing to the national insurance by more than 200,000 in two weeks will reduce its intensity soon. Third, this study focuses on university graduates and whether they are over-educated, which makes even more complicated the feat of trying to measure the effect of this external shock on their future tendency.

Against all these difficulties, it is worth trying to disentangle the effect that the pandemic could have on the previous scenarios projected. This sensitive analysis of the model will help to understand the behaviour of the model, and moreover, it will complete the sixth stage that a simulation exercise should have according to Gilbert and Troitzsch (2005).

Following the same structure as in the previous sections I will follow the three stages of this forecast to discuss the variations that should be included in the previous scenarios due to the health crisis.

Among the three previous scenarios, there was only one hypothesis regarding the number of university graduates (Stage 1). The hypothesis was that the variable will follow the same tendency as in the last two years (U1). It seems sensible to maintain

---

[96] https://www.imf.org/external/spanish/index.htm

[97] This is the sixth week of lockdown in Spain, and it is expected to last at least two more weeks. https://www.bde.es/f/webbde/GAP/Secciones/SalaPrensa/COVID-19/be2002-art1.pdf

this hypothesis as the reasons which were valid in the previous section are still valid here. The number of university graduates out of the 25-64-year-old population is expected not to vary. These days some Principals from some Spanish universities are expecting a fall in next year's enrolments. They assume that the economic crisis that will follow this health crisis will trigger an increase in the number of young people entering the labour market to help their family financially. Therefore, a reduction in university enrolments is expected. However, Spanish data has demonstrated that despite the economic crisis the number of students enrolling at university has still been increasing. Besides, even if there is a reduction in university enrolments for the coming year because of the pandemic and consequent economic crisis, it is expected that they might just be postponing their enrolment. We will have to wait to see which the real impact of this crisis is.

Indeed, in the last years, the tendency in the number of university graduates has been quite stable in all regions and for both genders. Therefore, a sudden change in this tendency seems unlikely. Holding a university degree is still considered the best way to find a 'good' job, even if it is not a highly skilled one. Therefore, only significant changes in university entrance fees or scholarships can, respectively, reduce or increase this tendency, and even in these cases, it will take some years to see the complete effect on the number of university graduates in the Spanish labour market. There are two main reasons: first, the generation effect. A reduction in the number of university students enrolled at university would take at least four years, until the current students finish their degree, to see the effect in the labour market. And, even so, the effect would be minimum, since the older generations holding a university degree are still active in the labour market. Second, it is not clear that these changes in the university entrance policies would have an immediate effect on the number of students enrolled at university. There are other familiar, cultural, and social factors which may affect their decision to study for a university degree. Given this, no change in the enrolment trend seems to be the simplest assumption to make at the moment.

However, the number of university graduates employed (W) and the number of university graduates employed in highly skilled jobs (HSJ) are more likely to see their tendency modified due to COVID19, at least in 2020-2021, and maybe also 2022 and 2023. These two variables would be directly related to the expected negative variation in the GDP predicted by the IMF or the *Banco de España* which is due to firms' shutting down and production stopping.

Furthermore, data show that the services sector and all those jobs which have direct contact with customers have been the most affected by this health crisis. That is, all those jobs that cannot be done via teleworking, which used to be also the less skilled jobs: related to trade and commerce or tourism and hotel business. According to this, it is sensible to think that highly skilled jobs will be less affected by this situation. However, there is also a group of skilled jobs related to the administration of other business, for instance, legal, economic or administrative advisers. The latter are highly skilled jobs which can be affected by the downward demand for their services. In this case, it will be expected that university graduates will face greater difficulties to find a job, especially a highly skilled job.

Other questions to take into account here are for how long the Spanish production system is expected to slow down and the Spanish economy to be losing added value and consequently destroying jobs, even if they are unskilled, skilled, or highly skilled jobs. Furthermore, it is also important to take into account how the Spanish production system and the Spanish economy in general will address the return.

The most likely outcome is that the Spanish labour market will re-establish the previous situation with high proportions of over-educated university graduates. It is difficult to think in changes which improve the situation of a significant proportion of the population who has been working in precarious conditions in recent years. To achieve such an important change in the Spanish labour market, there should be radical changes in the active labour market policies.

However, the most beneficial aspect for the Spanish economy and society will be whether the Spanish policy makers as well as the Spanish society take advantage of this external shock that has shaken Spanish society and the economy, to rebuild and renovate the Spanish production system. Only if there is a skill biased technological change demanding more skilled workers, will these over-educated university graduates find an opportunity to develop and improve their already acquired high skills.

5.6. Conclusions

After the analysis of the past and present of the educational mismatch for university graduates in the Spanish labour market done in the previous chapters, this chapter has used data from the Spanish Labour Force Survey and from the populations' projections of the National Statistical Institute (INE) within others to forecast the near future of this educational mismatch, from 2019 to 2030.

Focusing on the 25-64-year-old population, the chapter describes, calculates, and analyses the results of three possible scenarios which have been constructed using three different stages. In the first stage the number of university graduates are calculated; the second stage calculates the number of university graduates employed; and the last stage calculates the number of university graduates employed in a highly skilled job. Using these results, the number and the proportion of 25-64-year-old university graduates who are over-educated is calculated, by single ages, gender, and region.

The scenarios considered are based in the tendency of the last five years for each of the variables (from 2014 to 2018). Taking this as the initial data, Scenario 1 accounts for a variation in the variables equal to the variation of the last two years; Scenario 2 differs from Scenario 1 in the inclusion of a higher variation in the second stage, the number of university graduates employed, which is expected to follow the year-on-year variation of the GDP by region. Finally, Scenario 3 differs from Scenario 2 in the variation on the third stage of the calculations, the number of university graduates employed in highly skilled jobs, which is supposed to follow the tendency of investment in R+D of the region.

A final discussion, attempting to consider the effect that the current COVID19 pandemic is having and will have on the Spanish economy has also been included. I do not add a new scenario since we do not yet know how long this health crisis will last as there are no consistent real data to project a future tendency. Nevertheless, the inclusion of a discussion regarding the possible problems and solutions seems to be sensible, given the great impact that this pandemic is having in the world's economy in general, and on the Spanish economy, in particular. Besides, the sensitive analysis of the last section is also used to see the behaviour of the forecast model when some of the variables have significant variation.

Both the national results and the regional results of the projection are unfavourable to reduce the education mismatch for university graduates in the Spanish labour market. Slight differences appear between genders. In absolute terms, the number of university graduates who are over-educated is expected to increase in the next years for both genders. However, it is also expected that their proportion out of the number of university graduates who are employed will slowly increase. It might be due to the higher increase in the number of employed university graduates.

Therefore, the results of this projection stand for a constant increase in the proportion of over-educated university graduates in the Spanish labour market, following the same tendency found in Chapter Three. These results suggest the application of important active labour market policies attempting to overcome Spanish labour market behaviour in order to reduce its tendency to be imbalanced.

Throughout this thesis, the negative consequences that over-education has for both the individual and the society have been discussed. Lower productivity and lower wages and satisfaction are only some of these. Therefore, reducing the number of educationally mismatched individuals in the labour market seems appropriate to improve the individual’s conditions and returns in the labour market, but also the social and economic general conditions for the country, especially in Spain where the state is still funding the major part of university studies.

Overall, the main conclusion of this fifth chapter is the importance of reducing the number of over-educated university graduates for the proper functioning of the Spanish labour market. In this sense, policy makers should promote the creation of more skilled jobs. A positive relationship between R+D investment and employment has been seen, therefore, it seems sensible to think that promoting R+D investment both in the public sector, but also in the firms, is a good way to increase the demand for these more skilled workers, and also to create new jobs for highly skilled, skilled, and unskilled workers. In this sense, Section 5.5 discusses how despite the adversity of the health crisis, it could open up the way to a restructuring and a creation of a new productive system more geared to increase the demand for university graduates in appropriate jobs.

"Spain should assume structural reforms focused on making its labour market more dynamic. It should begin with the modernisation of its economic and productive structure."

Jordi Bacaria, Josep M. Coll,
and Elena Sánchez-Montijano (2015, p. 9)

Chapter 6. Conclusions

This chapter compiles first the main contributions of the thesis, to focus then on the main possible policy implications of the outcomes of the thesis. The objective is to extract which results aim to be useful for the policy makers to better understand the Spanish labour market, and to design more adequate policies to reduce the high impact of over-education. A final section suggests some indications regarding the direction of further research in this topic.

6.1. Context

The recent transformation of jobs into more skill demanding ones drives the change of the workforce characteristics. Note, for instance, how some office jobs nowadays require substantial computer-using knowledge while a basic level of literacy and numeracy were enough some decades ago. Robotics is another example of a sector which has developed extensively recently. The introduction of robots to everyday worklife is common nowadays, but the introduction of robots and artificial intelligence implies a transformation of work, workplaces and workers (Hagel, Schwartz and Bersin, 2017).

In this new scenario, with constant skill-biased technological innovations and organisational changes, workers may need higher or more flexible skills in order to adjust to new requirements of the labour market. They should be able, first, to do different tasks, either in the same occupation or in a different one, due to the organisational changes, and secondly, to learn quickly the new requirements of the workplace to adapt to the technological upgrading of skills. If not, either they will be displaced to other less skilled jobs or they will even be excluded from the labour market, increasing the unemployment rate.

University graduates are expected to be flexible workers. They have been acquiring high levels of general and specific skills at university. The specific skills should be useful to develop the task related to the area of study, e.g. law, economics or medicine, as well as the general skills acquired by studying for a long time, such as ability to read, to comprehend what they read, and to summarise. Therefore, university graduates are more likely to be employed in these new jobs.

However, in the last four decades, individuals with higher education levels have also found some difficulties finding a job in the labour market and especially, finding a suitable job where they can use all their skills. In the 1970s in the USA, and later in some European countries, the increment in the number of university graduates was equally neither followed nor preceded by an upgrade of the skills required by the job. The result was the emergence of skill (education) imbalances in the labour market.

This thesis aims at analysing the past, present and future of the education mismatch for university graduates in the Spanish labour market. Focusing on university graduates and using econometrics models, this thesis provides new evidence in three main aspects. First, it analyses the changes in the probability for a university graduate to be over-educated in the Spanish labour market from 1995 to 2018. Second, it estimates the effect that accounting for individuals' heterogeneity has on the probability of being over-educated. And third, the thesis forecasts the future tendency of university graduates to be over-educated in the Spanish labour market.

6.2. Sources, data, methodology and limitations

The focus of this thesis is on the Spanish labour market in general, and in 25-64-year-old university graduates, in particular. The Spanish Labour Force Survey (SLFS) and the Survey of Adults Skills (PIAAC) are the two main sources used in this thesis.

The first source, the SLFS, has been used here to analyse the main characteristics of the Spanish labour force (Chapter Two), to analyse the changes of the probability for a university graduate to be over-educated (Chapter Three), and to forecast the number and proportion of university graduates over-educated in the near future on the Spanish labour market (Chapter Five). Difficulties with these data do not arise either in the analysis of recent trends (Chapter Two) or in the forecast for the near future (Chapter Five), but there were some difficulties with the analysis of the probability of being over-educated (Chapter Three). In the latter case, there were three issues to solve before starting the analysis.

First, how to measure who was over-educated, since in the literature there is neither a single measurement nor a single definition (Green, McIntosh and Vignoles, 1999). As the main empirical literature related to education mismatches, data have finally defined the best way to calculate the degree of over-education. In this case, and given that the SLFS does not provide individual self-assessment and the objective methods required a very good definition of the requirements of the jobs that does not exist in the Spanish case, I followed Verdugo and Verdugo (1989) and used the statistical mean approach. Therefore, educational level was converted first into years of schooling to be used later to measure the degree of education mismatch.

Second, the use of the SLFS does not allow one to consider the possible heterogeneity among individuals. Therefore, some homogeneity among individuals as well as among occupations were assumed in Chapter Three.

The third issue to resolve was how to estimate the probability for university graduates to be over-educated: using a probit model or using the Heckman two-step model (Heckman, 1979). Using a probit model implies that individuals are randomly distributed whereas the Heckman two-step model accounts for a possible sample selection bias. In order to corroborate the differences between the estimated values, I run both models. As expected, the results justify the use of the Heckman two-step model since there is a pre-selection of the employed individuals and an over-estimation of the probit coefficients.

The second source, PIAAC data, is used in Chapter Four to measure first the variation in literacy and numeracy skills across countries, and second, to calculate the effect that accounting for these observable cognitive skills has on the probability of being over-educated in the Spanish labour market.

Pairwise comparisons are used in this fourth chapter to disentangle whether there is variation in these cognitive skills among individuals across countries. The aim is to sort out whether it makes sense to use them to account for individuals' skills heterogeneity.

In this fourth chapter and using the PIAAC data, I use two different measurements of education mismatch. The OECD approach (subjective approach) and the statistical mean approach. And to estimate the probability for a university graduate to be over-educated I use a probit model and a variation of the Heckman two-step model to account for the possible sample selection, as in Chapter Three. Slight differences appear in the proportion of over-educated university graduates between statistical mean approach and subjective approach due to the different nature of the measurements. Statistical mean approach is based on the data whereas the subjective measurement is based on the answers of the individuals and it may be affect by the individuals' expectation regarding to their job. A Pearson Chi-squared test for non-correlation is used to corroborate the similarity among both educational mismatch measurements.

Higher differences between both measurements emerge when the estimated coefficients for the probability of being over-educated are calculated. For instance, the coefficients accounting for the effect of the cognitive skills (literacy and numeracy) are negative for the sample using the statistical mean approach, whereas they are positive for the sample using the subjective method. Again, this result is the consequence of the different nature of both over-education measurements.

The main problems with this second source are caused by the nature of the data. On the one hand, there are ten numeracy and literacy PIAAC scores for each individual. These scores are derived from the background questionnaire and the outcome of the answer to the literacy and numeracy questionnaires which were done in two different stages. From these, ten plausible values using the Item Response Theory (IRT) are derived for each individual. On the other hand, samples were created to be used applying replicated weights by jackknife resampling. As a result, the estimation of the probit and the Heckman two-step model using Stata became more difficult when both issues, plausible values and jackknife resampling, were included in the same model, although it could be solved by creating a subprogram.[98]

Besides, PIAAC data only accounts for cognitive skills, and in particular, for literacy and numeracy proficiency. In some countries, although not for all, it also accounts for problem solving in technology-rich environments and reading components. In this thesis, the skill level of individuals and the skill mismatch are measured using these two cognitive skills.

Unfortunately, other sorts of skills, for instance manual skills or management skills are not measured in this survey. These other skills could have an important

[98] In this sense, the help and assistance of Dr. Jan Paul Heisig was essential.

role depending on the kind of job, and the low or excessive proficiency in them can also affect the skill mismatch in the labour market, and since those would not be considered by PIAAC data, neither are they in the analysis.

Furthermore, PIAAC data contains cross-section data only for one wave, therefore it does not allow an analysis of change over time. However, it was carried out in different countries and a comparison between them is possible. Therefore, even if the focus of the thesis is on the Spanish labour market, an analysis of other European countries (England, France, Italy, and Sweden) is also included to place Spain in the European context. In particular, data from these countries are used in Chapter Two and in Chapter Four.

6.3. The results of the research

This thesis aims to contribute to the analysis of over-education in the Spanish labour market. The analysis of other European labour markets has been useful to place Spain in the European context. The main results of this thesis can be summarised as follows:

First, the thesis analyses recent trends in the educational level and the occupational structure of the Spanish labour market. The analysis of the data from 1995 to 2018 allows one to follow the period of growth of the supply of and demand for university studies, and therefore, the increase in the educational level of the labour force and the more demanding skills in the labour market.

Second, the thesis concludes that *the probability for university graduates to be over-educated in the Spanish labour market has substantially increased from 1995 to 2018.* The analysis using both probit and the Heckman two-step model confirms that the Spanish labour market has not been able to generate sufficient skilled jobs either for those university graduate workers already available in the labour force or for these new more educated workers entering the Spanish labour market which has been saturated by the increasing number of university graduates.

Third, the analysis of the likelihood of being over-educated also detected a positive relationship between the economic crisis and the probability of being over-educated in the Spanish labour market. However, these results cannot confirm whether this positive relationship is due to a lower supply of highly skilled jobs or a higher propensity for university graduates to accept a job even if it does not require a university degree. More detailed data accounting for the motivations of the individuals to accept a job will help to better understand the reason behind this positive relationship between the economic crisis and the probability of being over-educated.

Fourth, the comparison between probit and the Heckman two-step model confirms the sample selection and the effect that accounting for the probability of being employed has on the probability of being over-educated. That is, when the Heckman two-step model is applied, the probabilities of being over-educated decrease for both genders and for all years. And, contrary to expectations, the probability of being employed has a higher effect in reducing the probability of being over-educated for men than for women.

Fifth, according to the data 25-64-year-old women are more likely to be over-educated than men in the Spanish labour market. This outcome is true for both models probit and the Heckman two-step model and it is in concordance with previous studies.

Sixth, contrary to what might be expected, including literacy or numeracy proficiency scores to account for cognitive skills heterogeneity does not have significant effects on the probability of being over-educated. The analysis of the data demonstrate that there is a negative relationship between higher cognitive skills and the probability of being over-educated, however, the value of this relationship is very low, and slightly affect the predicted probability of being over-educated. Therefore, two are the main conclusions for these results: first, it is important to account for individuals' skills heterogeneity; and second, it would be appropriate to find more accurate measurements to test individuals' skills, since, at least in the Spanish labour market, the literacy and numeracy proficiency skills do not seem to have enough power to explain the variation in the probability of being over-educated among university graduates.

Seventh, substantial changes in the Spanish production system will be recommended in order to reduce the education mismatch in the Spanish labour market. According to the results of the forecast, it would not be easy for the Spanish labour market to reduce the educational mismatch. All the scenarios considered get an upward trend in the number and the proportion of over-educated university graduates, for both genders, men and women.

Overall, the results of the thesis advise against the possible lengthening of the educational mismatch in the Spanish labour market. Even some policies to increase the demand for highly skilled jobs are applied.

6.4. Policy implications

In the matter of policy implications, the findings of this research suggest some changes in the current educational as well as labour market policies.

The results of this thesis basically recommend improving the education match for university graduates in the Spanish labour market. This is, a better use of the abilities and the knowledge of the individuals should be promoted. Therefore, policies on the supply and the demand side are needed to diminish the educational imbalance of the labour market.

In relation to the demand side, action is needed to *increase the demand for university graduate workers*. This is needed so that the labour market accommodates as soon as possible the university graduates that it has not been able to accommodate so far, but also, the ones that are still graduating from universities. In this sense, *the policies should be oriented to foster high-skill demanding industries.* One of the main problems in the Spanish labour market is the large number of small businesses and even self-employed, especially after their promotion with the labour market reform of 2012. This makes the creation and development of high-skill demanding industries more complicated. One possible action to shift this situation would be the promotion of start-ups or business incubators from the universities or encouraging the newly-created businesses to be part of a technologic cluster, to take advantage of the synergies created by their counterparts. Another potential option is to increase public and private R+D investment. It is expected that it would increase the demand for university graduate workers, as much for doing research as for implementing their results in the productive system, in non-routine jobs which need autonomy and which require decision-making skills.

Moreover, the results of the dissimilarity index reveal a noticeable level of gender occupational segregation. According to this, since 1995 there has been an improvement in the gender occupational distribution in the Spanish labour market. The dissimilarity index was 0.6817 in 1995 and 0.5583 in 2018. However, in order to have a balanced gender occupation distribution, it would be necessary for more than half the women to change their current occupational category.

Regarding the supply side, the results of this thesis point out the difficulty for university graduates to match their degree to a highly skilled job in the Spanish labour market. Results from Chapter Four do suggest an improvement in the employability of university graduates. In this sense, more detailed data including field of study is needed in order to be able to do a longitudinal analysis of the fields of study mismatch, and even, a projection of its future, to be able to highlight the differences in the probability of being over-educated regarding the field of study.

6.5. Further research

Overall, further research should be geared to updating the data on education mismatch in the Spanish labour market. In this sense, the continuous rounds of PISA and PIAAC are providing important data to evaluate the trend of the more cognitive skills among the students of compulsory education, but also among the adult population. However, actual and current individual data including larger skills proficiency, rather than only literacy and numeracy, would be essential to improve knowledge about individuals' skills heterogeneity. This would be especially necessary in the Spanish case where, as has been demonstrated in

Chapter Four, literacy and numeracy proficiency do not perfectly measure individuals' skills heterogeneity. Furthermore, recent literature has pointed out the increase in the demand for other sorts of skills, for instance more social and socioemotional abilities, which, in fact, are usually supposed to be more related to women (CEDEFOP, 2018; Matias Cortes, Jaimovich and Siu, 2016).

The final objective of improving the data available is to regularly measure the supply of and demand for skills of the labour market in order to diminish the labour market mismatches because, as has been demonstrated, education mismatch is inefficient both either for individuals and for society.

A perfect balance between the supply and the demand side is merely impossible, since there is a time lag between them. The lack or deficit of graduates in one area today, when the students are entering the university system, cannot be the same as it would be in four years' time, when the university graduates are entering the labour market. However, a better and continuous knowledge of the actual skills lacking or excessive in the labour market would be beneficial for politicians to design new strategies in order to reduce the imbalance, but also for future post-compulsory students to enrol in those areas or education levels which are expected to be more required or needed in the future labour market. In this sense, updating the data as well as the scenarios of the projections is also a very useful practice to foresee the changes and the main trends of the variables. The more detailed and up-to-date the data, the more useful the analysis and the forecast for policy makers.

References

Acemoglu, D. (1998). 'Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality'.*The Quarterly Journal of Economics,* 113 (4), 1055-1089.

Acemoglu, D. (2002a). 'Directed Technical Change'. *Journal of Economic Literature*, 40 (1), 7-72.

Acemoglu, D. (2002b). 'Technical Change, Inequality, and the Labor Market'. *American Economic Association,* 40 (1)*,* 7-72.

Aguilar Ramos, M. I. and García Crespo, D. (2008). 'Desajuste Educativo y Salarios en España: Nueva Evidencia con Datos de Panel'. *Estadística Española,* 50 (168)*,* 393-426.

Alba-Ramírez, A. (1993). 'Mismatch in the Spanish Labor Market. Overeducation?'. *The Journal of Human Resources,* XXVIII (2)*,* 259-278.

Alba-Ramírez, A. and Blázquez, M. (2003). 'Types of Job Match, Overeducation and Labour Mobility in Spain'. In F. Büchel, A. de Grip and A. Mertens (Eds), *Overeducation in Europe: Current Issues in Theory and Policy* (pp. 65-92). Cheltenham, UK: Edward Elgar Publishing.

Alba-Ramírez, A. and San Segundo, M. J. (1995). 'The Returns to Education in Spain'. *Economics of Education Review,* 14 (2)*,* 155-166.

Albert, C. (1998). *La Demanda de Educación Superior en España: 1977-1994*. Salamanca: CIDE. MEC.

Albert, C. (2008). 'Evolución de la Demanda de Educación Universitaria en España: 1977-2007'. *Temas Actuales de Economía,* (2)*,* 135-165.

Albert, C. and Davia, M. A. (2007). Understanding the Effect of Education on Health across European Countries. In C. D. Dávila Quintana, S. Rodríguez Feijoó, M. Tejera Gil, Y. Santana Jiménez, J. Á. Gil Jurado and A. Rodríguez Caro (eds), *XII Jornadas de la Asociación de Economía de la Educación (pp. 271-280)*. Las Palmas de Gran Canaria: Asociación Española de Economía de la Educación (AEDE) and Delta Publicaciones.

Albert, C., Juárez, J. P., Sánchez, R. and Toharia, L. (2000). 'La transición de la escuela al mercado de trabajo en España: años noventa'. *Papeles de Economía Española,* (86)*,* 42-58.

Albert, C., Pablo, J., Sánchez, R. and Toharia, L. (2003). 'Del Sistema Educativo al Mercado de Trabajo: Un Análisis de Flujos'. *Revista de Educación,* (330)*,* 137-155.

Albert, C., Toharia, L. and Davia, M. A. (2008). 'To Find or not to Find a First "Significant" Job'. *Revista de Economía Aplicada,* XVI (46)*,* 37-59.

Allen, J., Levels, M. and van der Velden, R. (2013). 'Skill Mismatch and Skill Use in Developed Countries: Evidence from the PIAAC Study'. *ROA Research Memorandum,* (017).

Allen, J. and Van der Velden, R. (2001). 'Education Mismatches versus Skill Mismatches: Effects on Wages, Job Satisfaction, and On-the-Job Search'. *Oxford Economic Papers,* (3)*,* 434-452.

Aparicio, A. (2010). High-School Dropouts and Transitory Labor Market Shocks: The Case of the Spanish Housing Boom (Vol. 5139). Bonn - Germany: IZA.

Arrow, K. J. (1973). 'Higher Education as a Filter'. *Journal of Public Economics,* (2)*,* 193-216.

Bacaria, J., Coll, J. M. and Sánchez-Montijano, E. (2015). 'The Labour Market in Spain: Problems, Challenges and Future Trends'. [Online]. *Policy Brief* (2015/01). Available at: https://www.social-inclusion-monitor.eu/uploads /tx_itao_download/EZ_Policy_Brief_SIM_Europe_ENG_2015_03.pdf. [Last accessed January 2021].

Badillo-Amador, L., García-Sánchez, A. and Vila, L. E. (2005). 'Mismatches in the Spanish Labor Market: Education vs. Competence Match'. *International Advances in Economic Research,* 11 (1)*,* 93-109.

Badillo Amador, L., López Nicolás, Á. and Vila, L. E. (2008). The Consequences on Job Satisfaction of Job-Worker Educational and Skill Mismatches in the Spanish Labour Market: a Panel Analysis, *FEDEA - Documentos de trabajo (2008-32)*: FEDEA.

Barceinas, F., Oliver, J., Raymond, J. L. and Roig, J. L. (2000). 'Los Rendimientos de la Educación y la Inserción Laboral en España'. *Papeles de Economía Española,* (86)*,* 128-148.

Barceinas, F., Oliver, J., Raymond, J. L. and Roig, J. L. (2001). 'Spain'. In C. Harmon, I. Walker and N. Westergaard-Nielsen (Eds), *Education and Earnings in Europe. A Cross Country Analysis of Returns to Education* (pp. 234-264). Cheltenham: Edward Elgar.

Barone, C. and Ortiz, L. (2011). 'Overeducation Among European University Graduates: A Comparative Analysis of Its Incidence and the Importance of Higher Education Differentiation'. *Higher Education,* (61)*,* 325-337.

Battu, H., Belfield, C. R. and Sloane, P. J. (1999). 'Overeducation Among Graduates: a Cohort Review'. *Education Economics,* 7 (1).

Becker, G. S. (1962). 'Investment in Human Capital: A Theoretical Analysis'. *The Journal of Political Economy,* LXX. October supplement (5, part 2)*,* 9-49.

Becker, G. S. (1964). *Human Capital. A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: National Bureau of Economic Research.

Belfield, C. R. (2000). *Economic Principles for Education. Theory and Evidence*. Cheltenham: Edward Elgar.

Beneito, P., Ferri, J., Moltó, M. L. and Uriel, E. (1996). Desajuste Educativo y Formación Laboral Especializada: Efectos sobre los Rendimientos Salariales. Valencia: IVIE.

Berg, I. E. (1973). *Education and Jobs. The Great Training Robbery*. Great Britain: Penguin Education.

Bertrand-Cloodt, D., Cörvers, F., Kriechel, B. and Thor, J. V. (2011). Flexible Jobs among Recent Graduates. In A. Caparrós Ruiz (ed), *XX Jornadas de la Asociación de Economía de la Educación* (pp. 625-648). Málaga: Asociación de Economía de la Educación.

Bertrand, Y. (1998). *Théories Contemporaines de l'Éducation*. Montréal: Chronique Sociale. Éditions Nouvelles.

Black, K. (2009). *Business Statistics: Contemporary Decision Making*. USA: John Wiley & Sons.

Blau, D. F., Ferber, M. A. and Winkler, A. E. (2006). *The Economics of Women, Men, and Work*. Upper Saddle River, New Jersey: Pearson Prentice Hall.

Boll, C., Leppin, J., Rossen, A. and Wolf, A. (2016). *Magnitude and Impact Factors of the Gender Pay Gap in EU Countries*. Luxembourg: European Comission.

Bordas, I. and Montané, M. (1993). 'Spain'. *European Journal of Teacher Education,* 16 (1)*,* 51-60.

Borghans, L. and de Grip, A. (2000). *The Overeducated Worker? The Economics of Skill Utilization*. Cheltenham, UK: Edward Elgar.

Bowman, M. J. (1970). 'Education and Economic Growth'. In R. L. Johns and others (Ed.), *Economic Factors Affecting the Financing of Education in the Decade Ahead*. Gainesville, Florida: National Educational Finance Project.

Brown, P. (2001). 'Skill Formation in the Twenty-First Century'. In P. Brown, A. Green and H. Lauder (Eds), *High Skills. Globalization, Competitiveness, and Skill Formation* (pp.1-55). New York: Oxford University Press.

Brown, P., Green, A. and Lauder, H. (2001). *High Skills. Globalization, Competitiveness, and Skill Formation*. New York: Oxford University Press.

Büchel, F. and Mertens, A. (2000). 'Overeducation, Undereducation, and the Theory of Career Mobility'. [Online]. *IZA DP* (195). Available at: https://www.iza.org/publications/dp/195/overeducation-undereducation-and-the-theory-of-career-mobility. [Last accessed January 2021].

Cabrales, A., Dolado, J. J. and Mora, R. (2014). 'Dual Labour Markets and (Lack of) On-the-Job Training: PIAAC Evidence from Spain and Other EU Countries'. [Online]. *IZA DP* (8649). Available at: http://ftp.iza.org/dp8649 .pdf. [Last accessed January 2021].

Calvo, J. L. (1988). 'Rendimientos del Capital Humano en Educación en España'. *Investigaciones Económicas (segunda época),* XII (3)*,* 473-782.

Carreras, A. and Tafunell, X. (2006). *Historia Económica de la España Contemporanea*. Barcelona: Critica.

Carreras, A. and Tafunell, X. (2016). *Historia Económica de la España Contemporánea (1789-2009)*. Barcelona: Crítica.

CEDEFOP. (2018). *Insights into Skill Shortages and Skill Mismatch. Learning from Cedefop's European Skills and Jobs Survey*. Luxemburg: Publications Office of the European Union.

Clogg, C. and Shockey, J. W. (1984). 'Mismatch Between Occupation and Schooling: A Prevalence Measure, Recent Trends and Demographic Analysis'. *Demography,* 21 (2)*,* 24.

Consejo de Universidades. (1987). El Mercado de Trabajo de los Titulados Universitarios en España. MEC (ed). Madrid.

Cook, J., Roberts, J. and Waylen, G. (2000). *Towards a Gendered Political Economy*. London: Palgrave Macmillan.

Chevalier, A. (2003). 'Measuring Over-education'. *Economica,* (70)*,* 509-531.

Chevalier, A. and Lindley, J. (2007). 'Over-Education and the Skills of UK Graduates'. [Online]. *CEE DP*. Available at: http://cee.lse.ac.uk/ceedps /ceedp79.pdf [Last accessed January 2021].

Choi de Mendizabal, Á. and Calero Martínez, J. (2018). 'Early School Dropout in Spain: Evolution During the Great Recession'. In M. Á. Malo and A. Moreno Mïnguez (Eds), *European Youth Labour Markets* (pp, 143-156)*. Problems and Policies*: Springer.

De Juan Asenjo, O. and López Santiago, L.A. (2004). 'Cambio técnico y cambio ocupacional en la economía Española (1980-2000)'. *Cuadernos de Economía,* 27, 3-32.

De Miguel, J. M. (1998). *Estructura y Cambio Social en España*. Madrid: Alianza Editorial.

Desjardins, R. and Warnke, A. J. (2012). 'Ageing and Skills: A Review and Analysis of Skill Gain and Skill Loss Over the Livespan and Over Time'. [Online]. *OECD Working Paper* (72). Available at: https://www.oecd-ilibrary.org/education/ageing-and-skills_5k9csvw87ckh-en. [Last accessed January 2021].

Doeringer, P. and Piore, M. J. (1971). *Internal Labour Markets and Manpower Analysis*. New York D.C.: Heath and Company.

Dolado, J. J., Felgueroso, F. and Jimeno, J. F. (2000). 'La inserción laboral de los titulados universitarios en España'. *Papeles de Economía Española,* (86)*,* 78-88.

Dolado, J. J., Jansen, M. and Jimeno, J. F. (2002). 'A Matching Model of Crowding-Out and On-the-Job Search (with an Application to Spain)'. *IZA DP,* (612).

Dolton, P. J. and Vignoles, A. (2000). 'The Incidence and Effects of Overeducation in the U.K. Graduate Labour Market'. *Economics of Education Review,* (19)*,* 179-198.

Dore, R. (1997). *The Diploma Disease. Education, Qualification and Development*. (2nd Ed.). London: Institute of Education. University of London.

Duncan, G. J. and Hoffman, S. D. (1981). 'The Incidence of Wage Effects of Overeducation'. *Economics of Education Review,* 1 (1)*,* 75-86.

Duncan, O. D. and Duncan, B. (1955). 'A Methodological Analysis of Segregation Indexes'. *American Sociological Review,* 20 (2)*,* 210-217.

Egido Gálvez, I. (2006). 'La Política Universitaria en la España Democrática: Logros y Carencias Después de Treinta Años'. *Tendencias Pedagógicas,* (11)*,* 207-222.

Eicher, J.-C. and Lévy-Garboua, L. (1979). *Économique de l'Éducation*. Paris: Éconómica.

Eichhorst, W. and Marx, P. (2016). *Non-Standard Employment in Post-Industrial Labour Markets. An Occupational Perspective*. Cheltenham (UK) and Northampton (MA, USA): Edward Elgar Publishing.

Englader, S. and Gurney, A. (1994). 'Productivity Growth: Medium-term Trends'. [Online]. *OECD, Economics Studies* (22). Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.411.7513&rep=rep1&type=pdf. [Last accessed January 2021].

Esping-Andersen, G. and Regini, M. (1999). *Why Deregulate Labour Markets?* Barcelona and Milano: Oxford University Press.

European Comission. *Eurostat*. [Online]. Available at: http://ec.europa.eu/eurostat/data/database. [Last accessed January 2021].

Felgueroso, F., Gutiérrez Domènech, M. and Jiménez-Martín, S. (2013). ¿Por qué el abandono escolar se ha mantenido tan elevado en España en las últimas dos décadas? El papel de la Ley de Eduación (LOGSE) [Online]. *Colección Estudios Económicos* (02-2013): FEDEA. Available at:

https://documentos.fedea.net/pubs/ee/2013/02-2013.pdf. [Last accessed January 2021].

Fernández, C. (2004). 'Job Match Quality throughout the Business Cycle in the Spanish Labour Market'. [Online]. *Documento de trabajo* (2004-01). Available at: http://documentos.fedea.net/pubs/dt/2004/dt-2004-01.pdf. [Last accessed January 2021].

Fisher, I. (1965 [1906]). *The Nature of Capital and Income*. New York: Augustus M. Kelley Publisher.

Flisi, S., Goglio, V., Meroni, E., Rodrigues, M. and Vera-Toscano, E. (2014). 'Occupational Mismatch in Europe: Understanding Overeducation and Overskilling for Policy Making'. [Online]. *JRC Science and Policy Reports* (Report EUR 26618 EN). Available at: http://publications.jrc.ec.europa.eu/repository/bitstream/JRC89712/occupational%20mismatch%20in%20europe.pdf. [Last accessed January 2021].

Freeman, R. and Schettkat, R. (2001). 'Skill Compression, Wage Differentials, and Employment: Germany vs the US'. *Oxford Economic Papers,* (3)*,* 582-603.

Freeman, R. B. (1976). *The Over-Educated American*. New York; London: Academic Press.

García-Mainar, I. and Montuenga-Gómez, V. M. (2005). 'Education Returns of Wage Earners and Self-Employed Workers: Portugal vs. Spain'. *Economics of Education Review,* (24)*,* 161-170.

García-Montalvo, J. (1995). 'Empleo y Sobrecualificación en el Caso Español '. *FEDEA - Documento de trabajo, 95-20*.

García-Montalvo, J. (2001). *Formación y Empleo de los Graduados de Enseñanza Superior en España y en Europa*. Valencia: Bancaja.

García-Montalvo, J. (2009). 'La Inserción Laboral de los Universitarios. El Fenómeno de la Sobrecualificación en España'. *Papeles de Economía Española,* (119)*,* 172-186.

García-Pérez, I. and Muñoz Bullón, F. (2007). 'Transitions into Permanent Employment in Spain: an Empirical Analysis for Young Workers'. [Online]. *Working Papers* (07-38. Business Economics Series 08). Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.597.6996&rep=rep1&type=pdf. [Last accessed January 2021].

García-Serrano, C. and Malo, M. A. (2013). 'Beyond the Contract Type Segmentation in Spain'. [Online]. *Cornell University ILR School - International Publications*. Available at: https://digitalcommons.ilr.cornell.edu/intl/278. [Last accessed January 2021].

García, P. and Gómez, R. (1994). 'Elaboración de Series Históricas de Empleo a Partir de la Encuesta de Población Activa (1964-1992)'. [Online]. *Documento de trabajo* (9409). Available at: https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/94/Fich/dt9409.pdf. [Last accessed January 2021].

García Serrano, C. and Malo, M. A. (1996). 'Desajuste Educativo y Movilidad Laboral en España'. *Revista de Economía Aplicada,* IV (11)*,* 105-131.

Gardiner, J. (2000). 'Gender and Family in the Formation of Human Capital'. In J. Cook, J. Roberts and G. Waylen (Eds), *Towards a Gendered Poitical Economy*. London: Palgrave Macmillan.

Garrido Medina, L. (2004). 'Demografía Longitudinal de la Ocupación'. *ICE,* (Mayo-Junio)*,* 105-142.
Gilbert, N. and Troitzsch, K. G. (2005). *Simulation for the Social Scientist. Second Edition*. New York: Open University Press.
Goldin, C. and Katz, L. F. (2007). 'The Race between Education and Techonology: The Evolution of U.S. Educational Wage Differentials, 1890 to 2005'. [Online]. *NBER Working Paper Series*. Available at: https://ideas.repec.org/p/nbr/nberwo/12984.html [Last accessed January 2021]]
Goldin, C. and Katz, L. F. (2008). *The Race between Education and Technology*. Massachusetts and England: The Belknap Press of Harvard University Press.
Green, F. (2012). 'Employee Involvement, Technology and Evolution in Job Skills: A Task-Based Analysis'. *Industrial and Labor Relations Review,* 65 (1)*,* 36-67.
Green, F. (2013). *Skills and Skilled Work. An Economic and Social Analysis*. Oxford, UK: Oxford University Press.
Green, F. and Henseke, G. (2016a). 'The Changing Graduate Labour Market: Analysis Using a New Indicator of Graduate Jobs'. [Online]. *IZA Journal of Labor Policy*, 5*,* 25. Available at: https://izajolp.springeropen.com/articles/10.1186/s40173-016-0070-0. [Last accessed January 2021]
Green, F. and Henseke, G. (2016b). *Should Governments of OECD Countries Worry about Graduate Over-Education?* London: UCL Institute of Education.
Green, F. and McIntosh, S. (2007). 'Is There a Genuine Under-Utilization of Skills amongst the Over-Qualified?'. *Applied Economics,* (39)*,* 427-439.
Green, F., McIntosh, S. and Vignoles, A. (1999). 'Overeducation and Skills: Clarifying the Concepts'. [Online]. *CEPDP* (435). Available at: http://cep.lse.ac.uk/pubs/download/dp0435.pdf. [Last accessed January 2021].
Green, F., McIntosh, S. and Vignoles, A. (2002). 'The Utilization of Education and Skills: Evidence from Britain'. *The Manchester School,* 70 (6)*,* 792-811.
Greenwood, C., Jenkins, A. and Vignoles, A. (2007). *Returns to Qualifications in England Updating the Evidence Base on Level 2 and Level 3 Vocational Qualifications*. [Online]. Available at: http://cee.lse.ac.uk/cee%20dps/ceedp89.pdf. [Last accessed January 2021].
Groot, W. and Maasen Van Den Brink, H. (1997). 'Allocation and the Returns to Over-education in the UK'. *Education Economics,* 5 (2)*,* 169-183.
Gonzalez Pandiella, A. (2014). 'Moving Towards a More Dynamic Busness Sector in Spain'. OECD Economics Department Working Papers, 1173. [Online]. Available at: https://www.oecd-ilibrary.org/docserver/5jxszm2k7fnw-en.pdf?expires=1612691620&id=id&accname=guest&checksum=1AB48EFC8F104F1F8847B5648A6D2ECD. [Last accessed January 2021].
Gutiérrez Domènech, M. (2008). 'The Impact of the Labour Market on the Timing of Marriage and Birth in Spain'. *Journal of Population Economics,* 21 (1)*,* 83-110.

Hagel, J., Schwartz, J. and Bersin, J. (2017). 'Navegando el Futuro del Trabajo ¿Podemos Hacer que las Empresas, los Trabajadores y las Instituciones Sociales Apunten en una Misma Dirección?'. [Online]. *Deloitte Review* (Edición 21. Julio 2017), 27-45. Available at: https://www2.deloitte.com/content/dam/insights/us/articles/3944_Navigating-the-FoW/DUP_Navigating-the-FoW-ES.pdf. [Last accessed January 2021].

Halaby, C. N. (1994). 'Overeducation and Skill Mismatch'. *Sociology of Education,* 67 (January), 47-59.

Hanushek, E. and Kimko, D. (2000). 'Schooling, labor force quality, and the growth of nations'. *American Economic Review,* (90), 1184-1208.

Hanushek, E. A. and Woessmann, L. (2007). 'The Role of School Improvement in Economic Development'. [Online]. *NBER Working Paper Series* (12832). Available at: https://www.nber.org/papers/w12832. [Last accessed January 2021]

Hanushek, E. A. and Woessmann, L. (2015). *The Knowledge Capital of Nations: Education and The Economics of Growth*. Cambridge, Massachussets: MIT Press.

Hartmann, H. (1976). 'Capitalism, Patriarchy, and Job Segregation by Sex'. *Signs,* 1 (3. Women and the Workplace: The Implications of Occupational Segregation), 137-169.

Hartog, J. (2000). 'Over-education and Earnings: Where Are We, Where Should We Go?'. *Economics of Education Review,* (19), 131-147.

Hartog, J. and Oosterbeek, H. (1988). 'Education, Allocation and Earnings in the Netherlands: Overschooling?'. *Economics of Education Review,* 7 (2), 185-194.

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

Heckman, J. J., Stixrud, J. and Urzua, S. (2006). 'The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behaviour'. [Online]. *NBER Working Paper Series,* 45. Available at: http://www.nber.org/papers/w12006. [Last accessed January 2021].

Henseke, G. and Green, F. (2016). *"Graduate Jobs" in OECD Countries: Analysis Using A New Indicator Based on High Skills Use*. Paris: OECD Publishing.

Hernández Lahiguera, L. and Serrano Martínez, L. (2013). 'Economic Effects of Education in Spain'. *Programme for the International Assessment of Adult Competences. 2013,* Volum II: Secondary Research (Spanish Report), 65-86.

Iglesias, J. (2005). *Capital Humano y Señalización*. Unpublished Research work. Departamento de Economía Aplicada. Universitat Autònoma de Barcelona, Bellaterra (Barcelona). Available at: http://www.ecap.uab.es/secretaria/docrecerca/jiglesias.pdf [Last accessed January 2021].

Instituto Nacional de Estadística (INE). Available at: http://www.ine.es. [Last accessed January 2021].

IRDAC. (1991). 'Preámbulo: enseñanza y formación europea, recursos humanos y competitividad', *La insuficiencia de cualificaciones en Europa. Dictamen IRDAC*: Comisión Europea.

Jimeno, J. F., Lacuesta, A. and Villanueva, E. (2013). 'Education, Work Experience and Cognitive Skills: A First Approximation of the PIAAC

Results'. *Programme for the International Assessment of Adult Competences. 2013,* Volume II: Secondary Research (Spanish Report)*,* 87-114.

Jovanovic, B. (1979a). 'Firm-especific Capital and Turnover'. *Journal of Political Economy,* 87 (6)*,* 1246-1260.

Jovanovic, B. (1979b). 'Job Matching and the Theory of Turnover'. *Journal of Political Economy,* 87 (5, part 1)*,* 972-990.

Keller, G. (2015). *Statistics for Management and Economics, Abreviated*. Boston: Cengage Learning.

Kennedy, P. (2014). *A Guide to Econometrics (6th ed.)*. Cambridge, MA: Wiley-Blackwell.

Lawrence, E. M. (2017). 'Why Do College Graduates Behave More Healthfully than Those Who Are Less Educated?'. *Journal of Health and Social Behaviour,* 58 (3)*,* 291-306.

Levels, M., Van der Velden, R. and Allen, J. (2013). 'Educational Mismatches and Skills: New Empirical Test of Old Hypotheses'. *Investigaciones de Economía de la Educación,* Track 4: Education and Labour Market (9)*,* 655-680.

Lévy-Garboua, L. (1979). 'Marché du Travail et Marché de l'Enseignement Superieur'. In J.-C. Eicher and L. Lévy-Garboua (Eds), *Économique de l'Éducation* (pp. 178-210). Paris: Économica.

Lindley, J. and McIntosh, S. (2015). 'Growth in Within Graduate Wage Inequality: The Role of Subjects, Cognitive Skills Dispersion and Occupational Concentration'. *Labour Economics,* (37)*,* 101-111.

Lynch, K. and Feeley, M. (2009). *Gender and Education (and Employment). Gendered Imperatives and their Implications for Women and Men. Lessons from Research for Policy Makers*: NESSE and European Comission.

Malo, M. A. (2015). *Labour Market Measures in Spain 2008-2013: The Crisis and Beyond*. Geneva: International Labour Organization.

Malthus, T. R. (1968 [1836]). *Principles of Political Economy Considered with a View to their Practical Application*. (2nd edition Ed.). New York: Augustus M. Kelley Publishers.

Maluquer de Motes, J. (2014). *La Economía Española en Perspectiva Histórica*. Barcelona: Pasado y Presente.

Mañe Vernet, F. (2001). *Cambio Tecnológico y Cualificaciones en la Industria Española: Una Aproximación Estructural*. Unpublished PhD. Universitat Autònoma de Barcelona, Bellaterra (Barcelona).

Marcenaro-Gutierrez, O. and Navarro Gómez, M. L. (2001). 'Un Análisis Microeconómico de la Demanda de Educación Superior en España'. *Estudios de Economía Aplicada,* (19)*,* 69-86.

Marcenaro-Gutierrez, O., Vignoles, A. and Coulon, A. d. (2007). *The Value of Basic Skills in the British Labour Market*. [Online]. Available at: http://cee.lse.ac.uk/cee%20dps/ceedp77.pdf. [Last accessed January 2021].

Marqués Perales, I. (2015). *La Movilidad Social en España*. Madrid: Catarata.

Marshall, A. (1920 [1890]). *Principles of Economics. An Introductory Volume. "Natura non facit saltum"*. (Eighth edition Ed.). London: Macmillan.

Martínez García, J. S. (2013). 'Over-Qualification of University Graduates and Social Mobility'. *Programme for the International Assessment of Adult Competences. 2013,* Volume II: Secondary Research (Spanish Report)*,* 115-138.

Matias Cortes, G., Jaimovich, N. and Siu, H. E. (2016). 'The End of Men and Rise of Women in the High-Skilled Labor Market'. [Online]. *Working Papers*. Available at: http://www.econ.queensu.ca/files/other/Siu%20paper%20endofmen20160920.pdf. [Last accessed January 2021].

Meroni, E. C. and Vera-Toscano, E. (2017). 'The Persistence of Overeducation among Recent Graduates'. *Labour Economics,* (48)*,* 120-143.

Mill, J. S. (1968 [1848]). *Principles of Political Economy with Some of their Applications to Social Philosophy*. London, Ontario: University of Toronto Press. Routledge & Kegan Paul.

Mincer, J. (1962). 'On-the-job Training: Costs, Returns, and Some Implications'. *Journal of Political Economy,* 70 (5, Part 2: Investment in Human Beings)*,* 50-79.

Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: Columbia University Press.

Mingat, A. and Eicher, J.-C. (1982). 'Higher Education and Employment Markets in France'. *Higher Education,* (11)*,* 211-220.

Molina, O. and López-Roldán, P. (2015). 'Occupational Growth and Non-Standard Employment in the Spanish Service Sector: From Upgrading to Polarisation'. In W. Eichhorst and P. Marx (Eds), *Non-Standard Employment in Post-Industrial Labour Markets. An Occupational Perspective* (pp. 110-149). Cheltenham, UK. Northampton, MA, USA: Edward Elgar Publishing.

Montt, G. (2015). 'The Causes and Consequences of Field-of-study Mismatch: An Analysis Using PIAAC'. [Online]. *OECD Social, Emploment and Migration Working Papers* (167). Available at: https://www.oecd-ilibrary.org/docserver/5jrxm4dhv9r2-en.pdf?expires=1612712439&id=id&accname=guest&checksum=6D3C71F1C5354C1E0FC850C86B556A5F. [Last accessed January 2021]

Mora, J.-G., García-Montalvo, J. and García-Aracil, A. (2000). 'Higher Education and Graduate Employment in Spain'. *European Journal of Education,* 35 (2)*,* 229-237.

Muñoz de Bustillo Llorente, R. and Antón Pérez, J.-I. (2011). 'From the Highest Employment Growth to the Deepest Fall: Economic Crisis and Labour Inequalities in Spain'. In D. Vaughan-Whitehead (Ed.), *Work Inequalities in the Crisis* (pp. 393-444). Cheltenham (UK) and Northampton (MA, USA): Edward Elgar Publishing and International Labour Office.

Murillo, I. P., Rahona, M. and Salinas, M. M. (2010). Incidencia del Desajuste Educativo en el Rendimiento Privado de la Educación en España. In M. J. Mancebón-Torrubia, D. Pérez Ximénez de Embún, J. M. Gómez Sancho and G. Giménez Esteban (eds), *XIX Jornadas de la Asociación de Economía de la Educación*. Zaragoza, Spain: Asociación de Economía de la Educación.

Myro, R. (2005). 'Crecimiento Económico y Cambio Estructural'. In J. L. García Delgado and R. Myro (Eds), *Lecciones de Economía Española* (pp. 53-86). Navarra: Thomson Civitas.

National Center for Education Statistics. *Adult Literacy and Lifeskills Survey (ALL)*. [Online]. Available at: http://nces.ed.gov/surveys/all. [Last accessed January 2021].

National Center for Education Statistics. *International Adult Literacy Survey (IALS)*. [Online]. Available at: http://nces.ed.gov/surveys/ials/. [Last accessed January 2021].

Nicolau, R. (2005). 'Población, Salud y Actividad'. In A. Carreras and X. Tafunell (Eds), *Estadísticas Históricas de España. Siglos XIX y XX* (Vol. I, pp. 77-154). Bilbao: Fundación BBVA.

Nieto, S. (2014). *Essays on Overeducation: Evidence from Spain*. Unpublished PhD. Universitat de Barcelona, Barcelona. Available at: http://diposit.ub.edu/dspace/bitstream/2445/64545/1/SNV_PHD_THESIS.pdf [Last accessed January 2021].

Nieto, S. and Ramos, R. (2011). ¿La sobreeducación de los padres afecta al rendimiento académico de sus hijos? In A. Caparrós Ruiz (ed), *XX Jornadas de la Asociación de Economía de la Educación*. Málaga: Asociación de Economía de la Educación.

Nuñez, C. E. (1993). 'Alfabetización y Desarrollo Económico en España: una Visión a Largo Plazo'. In C. E. Nuñez and G. Tortella (Eds), *La Maldición Divina. Ignorancia y Atraso Económico en Perspectiva Histórica* (pp. 223-236). Madrid: Alianza Editorial.

OECD. (2013a). *OECD Skills Outlook 2013. First Results from the Survey of Adult Skills*: OECD.

OECD (2013b). *PIAAC. Programa Internacional para la Evaluación de las Competencias de la Población Adulta. 2013 . Volumen I. Informe Español*. Madrid: Ministerio de Educación, Cultura y Deporte.

OECD. (2013c). *The Survey of Adult Skills. Reader's companion*: OECD.

OECD. (2016). *Skills Matter. Further Results from the Survey of Adults Skills*.

Oliver, J., Ramos, X. and Raymond, J. L. (1999). 'Not All University Degrees Yield the Same Return: Private and Social Returns to Higher Education for Males in Spain'. [Online]. *Working papers. Departament d'Economia Aplicada. Universitat Autònoma de Barcelona* (99.04). Available at: http://www.ecap.uab.es/RePEc/doc/wpdea9904.pdf. [Last accessed January 2021].

Oliver, J., Ramos, X. and Raymond, J. L. (2001). 'Capital Humano y Desigualdad en España: 1985-1996'. [Online]. *Working papers. Departament d'Economia Aplicada. Universitat Autònoma de Barcelona* (11). Available at: http://www.ecap.uab.es/RePEc/doc/wp0111.pdf. [Last accessed January 2021].

Oliver, J., Raymond, J. L. and Sala, H. (2001). 'Necesidad de Formación en el Mercado de Trabajo Español: Composición del Empleo y Estructura Productiva'. [Online]. *Working papers. Departament d'Economia Aplicada. Universitat Autònoma de Barcelona* (01.17). Available at: http://www.ecap.uab.es/RePEc/doc/Wp0117.pdf. [Last accessed January 2021].

Pablo-Romero, M. P. and Gómez-Calero, M. P. (2008). 'Análisis por Provincias de la Contribución del Capital Humano a la Producción en la Década de los Noventa'. *Investigaciones Económicas,* XXXII (1)*,* 27-52.

Parellada, M., Duch, N. and Alvarez, M. (2009). 'Job Supply and Demand for University Graduates in Spain. A (Relative) Mismatch Perspective'. *Industry & Higher Education,* 23 (1, February 2009)*,* 29-38.

Pérez Infante, J. I. (2006). *Las Estadísticas del Mercado de Trabajo en España*. Madrid: Ministerio de Trabajo y Asuntos Sociales.

Pérez Infante, J. I. (2008). 'Luces y Sombras del Mercado de Trabajo en España. Problemas, Situación y Tendencias'. *Economia Industrial,* (367)*,* 35-49.

Pigou, A. C. (1929). *A Study in Public Finance*. (2nd Ed.). London: Macmillan.

Polavieja, J. G. (2003). 'Temporary Contracts and Labour Market Segmentation in Spain: An Employment-Rent Approach'. *European Sociological Review,* 19 (5)*,* 501-517.

Polavieja, J. G. (2005). 'The Incidence of Temporary Employment in Advanced Economies: Why is Spain Different?'. *European Sociological Review,* 22 (1)*,* 61-78.

Pryor, F. L. and Schaffer, D. L. (1999). *Who's Not Working and Why. Employment, Cognitive Skills, Wages, and The Changing U.S. Labor Market*. Cambridge, UK: Cambridge University Press.

Psacharopoulos, G. (1984). 'The Contribution of Education to Economic Growth'. In J. W. Kendrick (Ed.), *International Comparisons of Productivity and Causes of the Slowdown* (pp. 340-346). Cambridge: Ballinger Publishing.

Psacharopoulos, G. (1996). 'La contribución de la educación al crecimiento económico: comparaciones internacionales'. In E. Oroval Planas (Ed.), *Economía de la Educación* (pp. 161-182). Barcelona: Ariel Educación.

Quintini, G. (2011). 'Over-Qualified or Under-Skilled: A Review of Existing Literature'. [Online]. *OECD Social, Emploment and Migration Working Papers* (121). Available at: http://dx.doi.org/10.1787/5kg58j9d7b6d-en. [Last accessed January 2021].

Rahona López, M. M. (2008). *Tesis: La Educación Universitaria en España y la Inserción Laboral de los Graduados en la Década de los Noventa. Un enfoque comparado*. Universidad Autónoma de Madrid, Madrid.

Recio, A. (1994). 'Flexibilidad laboral y desempleo en España (reflexiones al filo de la reforma laboral)'. *Cuadernos de Relaciones Laborales,* (5)*,* 57-74.

Recio, A. (1997). *Trabajo, Personas, Mercados. Manual de Economía Laboral*. Barcelona: Icaria: FUHEMM.

Reich, M. (1984). 'Segmented Labour: Time Series Hypothesis and Evidence'. *Cambridge Journal of Economics,* 8*,* 63-81.

Reich, M., Gordon, D. M. and Edwards, R. C. (1973). 'Dual Labor Markets: A Theory of Labor Market Segmentation'. *American Economic Review,* 63 (2)*,* 359-365.

Requena, M. (2016). El Ascenso Social, *Observatorio Social de "La Caixa"* (Vol. Dossier. La Educación como Ascensor Social, pp. 19-28).

Reques Velasco, P. D. (2007). *Atlas Digital de la España Universitaria: Bases para la Planificación Estratégica de la Enseñanza Superior*. Santander: Universidad de Cantabria y Banco de Santander.

Riviére Gómez, J. (2002). *Estructura Ocupacional, Formación y Mercado de Trabajo en España*. Unpublished PhD. Universidad Nacional de Educación a Distancia (UNED), Madrid.

Rumberger, R. W. (1981). 'The Rising Incidence of Overeducation in the U.S. Labor Market'. *Economics of Education Review,* (1), 293-314.

Rumberger, R. W. (1987a). 'The Impact of Surplus Schooling on Productivity and Earnings'. *The Journal of Human Resources,* 22 (1), 24-50.

Rumberger, R. W. (1987b). 'The Potential Impact of Technology on the Skill Requirements of Future Jobs'. In G. Burke and R. W. Rumberger (Eds), *The Future Impact of Technology on Work and Education* (pp. 74-95). London: The Stanford Series on Education & Public Policy. The Flamer Press.

San Segundo, M. J. (1997). 'Educación e Ingresos en el Mercado de Trabajo Español'. *Cuadernos Económicos del ICE,* (63), 105-123.

Sanromá, E. and Ramos, R. (2004). Overeducation and Local Labour Markets in Spain, *44th Congress of the European Regional Science Association*. Porto.

Sarasúa, C. (2018). 'Women's Work and Structural Change: Occupational Structure in Eighteenth-Century Spain'. [Online]. *The Economic History Review*, 00 (0), 1-29. Available at: https://doi.org/10.1111/ehr.12733. [Last accessed January 2021].

Schultz, T. W. (1961). 'Investment in Human Capital'. *The American Economic Review,* LI (1), 1-17.

Schultz, T. W. (1981). *Investing in People. The Economics of Population Quality*. Los Angeles: University of California Press Berkeley and Los Angeles.

Sen, A. (1987). *Commodities and Capabilities*. New York: Oxford University Press.

Sen, A. (1997). 'Editorial: Human Capital and Human Capability'. *World Development,* 25 (12), 1959-1961.

Sicherman, N. (1987). "Over-education" in the Labor Market, *Working Paper*. Center for the Study of the Economy and the State. University of Chicago.

Sicherman, N. and Galor, O. (1990). 'A Theory of Career Mobility'. *Journal of Political Economy,* 98 (1), 169-192.

Sloane, P. J. (2002). 'How Ado About Nothing? What Does the Over-Education Literature Really Tell Us?'. [Online]. *International Conference on Over-education in Europe: What Do We Know?* Available at: http://200.6.99.248/~bru487cl/files/libros/Tendencias/pdf/sloane.pdf. [Last accessed January 2021].

Smith, A. (1961 [1776]). *The Wealth of Nations*. London: University Paperbacks by Methuen.

Solow, R. M. (1957). 'Technical Change and the Aggregate Production Function'. *The Review of Economics and Statistics,* 39 (3), 312-320.

Spence, A. M. (1973). 'Job Market Signalling'. *The Quarterly Journal of Economics,* 87 (3), 355-374.

Stiglitz, J. E. (1975). 'The Theory of "Screening", Education, and the Distribution of Income'. *The American Economic Review,* 65 (3), 283-301.

Teichler, U. E. (2007). *Careers of University Graduates. Views and Experiences in Comparative Perspective*. The Netherlands: Springer.

Tight, M. (2009). *The Development of Higher Education in the United Kingdom since 1945*. London: The Society for Research into Higher Education and Open University Press.

Tohari, L. (1997). *Labour Market Studies. Spain.* European Commission. Employment & Social Affairs. Employment & Labour Market. Series nº1. [Online]. Available at: http://repositori.uji.es/xmlui/bitstream/handle/10234/72364/Labour_market_Studies_Spain.pdf?sequence=1&isAllowed=y. [Last Accessed: January 2021]

Toharia, L. (1998). *El Mercado de Trabajo en España*. Madrid: McGraw-Hill.

Toharia, L. (2003). 'El Mercado de Trabajo en España, 1978-2003'. *ICE,* (811).

Toharia, L. (2008). 'Activité, Emploi et Chômage en Espagne (1976-2007): Une Perspective Évolutive et Comparée'. *Travail et Emploi,* (115)*,* 13-27.

Tomlinson, S. (2009). *Education in a Post-Welfare Society*. Glasgow: Open University Press. McGraw-Hill Education.

Torrejón, S. (2019). *El Cambio Estructural del Empleo en España a lo Largo del Ciclo Económico (1995-2014)*. Madrid: Consejo Superior de Investigaciones Científicas.

Trevis Certo, S., Busenbark, J. R., Woo, H.-S. and Semadeni, M. (2016). 'Sample Selection Bias and Heckman Models in Strategic Management Research'. *Strategic Management Journal,* (37)*,* 2639-2657.

Tsang, M. C. and Levin, H. M. (1985). 'The Economics of Overeducation'. *Economics of Education Review,* 4 (2)*,* 93-104.

Tsang, M. C., Rumberger, R. W. and Levin, H. M. (1991). 'The Impact of Surplus Schooling on Worker Productivity'. *Industrial Relations,* 30 (2)*,* 209-228.

Van der Velden, R. K. W. and Van Smoorenburg, M. S. M. (1997). 'The Measurement of Overeducation and Undereducation: Self-Report vs. Job-Analyst Method'. [Online]. *ROA-RM* (1997/2E)*,* 13

Vaughan-Whitehead, D. (2011). *Work Inequalities in the Crisis*. Cheltenham (UK) and Northampton (MA, USA): Edward Elgar Publishing Limited and International Labour Office.

Verdugo, R. R. and Verdugo, N. T. (1989). 'The Impact of Surplus Schooling on Earnings'. *The Journal of Human Resources,* XXIV (4, Autumn)*,* 629-643.

Verhaest, D. and Omey, E. (2006). 'The Impact of Overeducation and its Measurement'. *Social Indicators Research,* (77)*,* 419-448.

Vila, L. E. and Mora, J.-G. (1996). 'Educación e ingresos de los trabajadores en España: evolución en los años ochenta'. In J. Grao and A. Ipiña (Eds), *Economía de la Educación. Temas de estudio e investigación* (pp. 233-257). Bilbao: Servicio Central de Publicaciones del Gobierno Vasco.

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.

Weiss, A. (1995). 'Human Capital vs. Signalling Explanations of Wages'. *Journal of Economic Prospectives,* 9 (4)*,* 22.

Willis, P. (2005). *Aprendiendo a Trabajar. Como los Chicos de Clase Obrera Consiguen Trabajos de Clase Obrera*. Madrid: Akal Universitaria.

Wolf, A. (2002). *Does Education Matter? Myths about Education and Economic Growth*. London: Penguin Books.

Appendix 1

Table A1. 1. Educational Classification used by SLFS from 1995 to 1999

Code	Spanish description	English description	Education level
A	**ANALFABETOS Y SIN ESTUDIOS**	**ILLITERATE AND WITHOUT ANY STUDIES**	
01	Analfabetos	Illiterate	
02	Sin estudios	Without any studies	
B	**ESTUDIOS PRIMARIOS, E.G.B.(1), CICLO INICIAL, MEDIO O PRIMERA ETAPA Y EQUIVALENTE**	**PRIMARY EDUCATION, INITIAL OR SECOND STAGE OF PRIMARY EDUCATION OR FIRST STAGE OR EQUIVALENT**	
03	Estudios primarios y E.G.B.(1), ciclos inicial, medio o primera etapa o equivalente	Primary education, initial or second stage of primary education or first stage or equivalent	
C	**ENSEÑANZA GENERAL: 2º GRADO, 1ER CICLO. BACHILLER ELEMENTAL, E.G.B.(1), CICLO SUPERIOR, SEGUNDA ETAPA Y EDUCACIÓN SECUNDARIA (NUEVO SISTEMA)**	**GENERAL EDUCATION: SECOND YEAR, FIRST STAGE. TWO YEARS OF SECONDARY EDUCATION, PRIMARY EDUCATION, SUPERIOR STAGE, SECOND STAGE AND SECONDARY EDUCATION (NEW SYSTEM)**	**Compulsory or less**
04	Bachiller elemental o equivalente; E.G.B(1), ciclo superior o 2ª etapa, y E.S.O. (2) (nuevo sistema)	Two years of secondary education or equivalent; Primary education, second or superior stage and Secondary education (new system).	
05	Certificado de escolaridad o cualquier otro certificado que en el futuro acredite la asistencia a la escuela durante el período de escolaridad obligatoria (habiendo superado un mínimo de cursos)	Schooling certificate or any other certificate that is evidence in the future of school attendance during the compulsory education period (having passed a minimum number of courses)	
E	**ENSEÑANZA GENERAL: SEGUNDO GRADO, SEGUNCO CICLO. BACHILLERATO SUPERIOR, CON O SIN PREUNIVERSITARIO, B.U.P.(4), CON O SIN C.O.U.(5) Y BACHILLERATO (NUEVO SISTEMA)**	**GENERAL EDUCATION: SECOND STAGE. KEY STAGE 5 OR SIXTH FORM, WITH OR WITHOUT PREUNIVERSITY COURSE**	
08	Bachillerato superior, B.U.P.(4) y Bachillerato	Key Stage 5 or Sixth Form	**Academic**
M	**OTRAS ENSEÑANZAS SUPERIORES**	**OTHER HIGHER EDUCATION**	
27	Estudios superiores de dos o tres años de duración no equivalentes a diplomado universitario	Two or three-year-long superior studies not equivalent to 3-year degrees	
29	Estudios superiores de al menos cuatro años de duración no equivalentes a licenciado universitario	At least four-year-long Superior Studies which are not equivalent to 5-year degrees	

Code	Spanish description	English description	Educational level
D	**ESTUDIOS TÉCNICO-PROFESIONALES DE PRIMER GRADO**	**TECHNICAL AND VOCATIONAL STUDIES FIRST STAGE**	**Vocational**
06	F.P.(3) de primer grado o equivalente	First stage of vocational Education and Training or equivalent	
07	Otras enseñanzas técnico-profesionales de primer grado	Other first stage of technical and vocational teachings	
F	**ENSEÑANZA TÉCNICO-PROFESIONAL DE GRADO MEDIO**	**TECHNICAL AND VOCATIONAL STUDIES MIDDLE STAGE**	
09	Módulo 2 de F.P.(3)	2nd Module of Vocational Education and Training	
G	**ENSEÑANZA TÉCNICO-PROFESIONAL DE SEGUNDO GRADO**	**TECHNICAL AND VOCATIONAL STUDIES SECOND STAGE**	
10	F.P.II(6) o equivalentes académicamente	2nd stage of Vocational Education and Training or academically equivalents	
11	Estudios de artes aplicadas y oficios artísticos	Arts and artistic jobs studies	
12	Otras enseñanzas regladas equivalentes laboralmente o similares a la F.P.II (6)	Other formal trainings equivalent in relation to the labour market to 2nd stage of Vocational Education and Training	
H	**ENSEÑANZA TÉCNICO-PROFESIONAL DE GRADO SUPERIOR**	**TECHNICAL AND VOCATIONAL STUDIES SUPERIOR STAGE**	
13	Módulo 3 de F.P.(3)	3rd Module of Vocational Education and Training	

Code	Spanish description	English description	Education level
J	**CARRERAS UNIVERSITARIAS DE CICLO CORTO**	**3-YEAR DEGREE**	
15	Área de Ingeniería y Tecnología	Engineering and Technology Area	
16	Área de Ciencias Médicas y de la Salud	Medical and Health Sciences Area	
17	Área de Humanidades	Humanities Area	
18	Área de Ciencias Sociales y Jurídicas	Social Sciences and Law Area	
19	Área de Ciencias Exactas y Naturales	Exact Sciences and Biology Area	
K	**TRES CURSOS APROBADOS, SIN DERECHO A TITULACIÓN, DE UNA CARRERA DE CICLO LARGO**	**THREE PASSED COURSES OF A 5-YEAR DEGREE WITH NO RIGHT TO A TITLE**	
20	Tres cursos aprobados (o primer ciclo), sin derecho a titulación, de una carrera de ciclo largo	Three passed courses (or first cycle) of a 5-year degree with no right to a title	
L	**CARRERAS UNIVERSITARIAS DE CICLO LARGO**	**5-YEAR DEGREES**	
21	Área de Ingeniería y Tecnología	Engineering and Technology Area	
22	Área de Ciencias Médicas y de la Salud	Medical and Health Sciences Area	**University**
23	Área de Humanidades	Humanities Area	
24	Área de Ciencias Sociales y Jurídicas	Social Sciences and Law Area	
25	Área de Ciencias Exactas y Naturales	Exact Sciences and Biology Area	
M	**OTRAS ENSEÑANZAS SUPERIORES**	**OTHER HIGHER EDUCATION**	
26	Estudios equivalentes a todos los efectos a diplomado universitario	Equivalent studies to any 3-year degrees	
28	Estudios equivalentes a todos los efectos a licenciado universitario	Equivalent studies to any 5-year degree	
N	**DOCTORES**	**PhD (Doctorate)**	
30	Doctores en Ingeniería, Tecnología, Arquitectura, Ciencias médicas y de la Salud, Exactas y Naturales	PhD (Doctorate) in Engineering, Technology, Architecture, Medical and Health Sciences, Exact Sciences and Biology	
31	Doctores en Humanidades y Ciencias Sociales	PhD (Doctorate) in Humanities and Social Sciences	

Notes:
(1) Educación General Básica (2) Educación Secundaria Obligatoria (3) Formación Profesional (4) Bachillerato Unificado Polivalente (5) Curso de Orientación a la Universidad (6) Formación Profesional de Segundo Grado
Source: INE and my own translation based on information of the Spanish Embassy in United Kingdom (http://www.mecd.gob.es/dms-static/a8de9034-5c52-4347-b80a-69c78f0652d5/consejerias-exteriores/reino-unido/estudiar/cuadro-comparativo-inglaterra.pdf)

Table A1. 2. Educational Classification used by SLFS from 2000 to 2010

Code	Spanish description	English description	Education level
80	Analfabetos	Illiterate	Compulsory or less
11	Estudios primarios incompletos	Incomplete primary studies	Compulsory or less
12	Estudios primarios completos	Complete primary studies	Compulsory or less
21	Programas para la formación e inserción laboral que no precisan de una titulación académica de la 1ª etapa de secundaria para su realización	Schemes for vocational trainings and employability that do not require any academic diploma of the 1st stage of secondary education to be carried out	Compulsory or less
22	Primera etapa de secundaria sin título	First stage of secondary education with no title	Compulsory or less
23	Primera etapa de secundaria con título	First stage of secondary education with a title	Compulsory or less
31	Programas para la formación e inserción laboral que precisan de una titulación de estudios secundarios de 1ª etapa para su realización	Schemes for vocational trainings and employability that require a diploma of the 1st stage of secondary education to be carried out	Vocational
33	Enseñanzas de grado medio de formación profesional específica, artes plásticas y diseño y deportivas	Teachings of middle grade of specific vocational education and training, plastic arts and design and sport	Vocational
34	Enseñanzas de grado medio de música y danza	Teachings of middle grade of specific vocational education and training of music and dance	Vocational
36[(1)]	Garantía social / Iniciación profesional	Social Guarantee / professional introduction	Vocational
41	Programas para la formación e inserción laboral que precisan de una titulación de estudios secundarios de 2ª etapa para su realización	Schemes for vocational trainings and employability that require an academic diploma of the 2nd stage of secondary education to be carried out	Vocational
51	Enseñanzas de grado superior de formación profesional específica y equivalentes: artes plásticas y diseño y deportivas	Certificate of Higher Education in specific vocational education and training and equivalent: plastic arts and design and sport	Vocational
53	Programas que precisan de una titulación de formación profesional de grado superior para su realización (más de 300 horas o 6 meses)	Schemes that require a Certificate of Higher Education in specific vocational education and training (moret han 300 hours or 6 month)	Vocational

Code	Spanish description	English description	Education level
32	Enseñanzas de bachillerato	Key Stage5 of sixth form teachings	**Academic**
52	Títulos propios de las universidades que no sean de postgrado (dos años o más)	Degree diploma proper to universities (2 years or more)	
50[2]	Grado	Degree (4 years)	**University**
54	Enseñanzas universitarias de 1º ciclo y equivalentes o personas que han aprobado 3 cursos completos de una licenciatura o créditos equivalentes	3-years degree or equivalent or people who passed 3 full courses of a 5-year degree or equivalent credits	
55	Enseñanzas universitarias de 1º y 2º ciclo, de solo 2º ciclo y equivalentes	1st and 2nd cycles of university studies (3+2), only 2nd cycle university studies and equivalent	
56	Programas oficiales de especialización profesional	Official schemes in specialization in vocational training	
61	Doctorado universitario	University PhD (doctorate)	

Notes:

(1) New category included the three month of 2003. “It gathers the people with a basic vocational education which allow their participation in the labour market, but who have not reached the objectives of the compulsory secondary education (ESO) and, therefore, they cannot be classified.” (INE, my own translation).

(2) New category included in 2009

Source: INE and my own translation based on information of the Spanish Embassy in United Kingdom (http://www.mecd.gob.es/dms-static/a8de9034-5c52-4347-b80a-69c78f0652d5/consejerias-exteriores/reino-unido/estudiar/cuadro-comparativo-inglaterra.pdf)

Appendix 2

Table A2. 1. Educational Classification used by SLFS from 1995 to 1999 and correspondence in years of education

Code	Spanish description	English description	Years of education
A	**ANALFABETOS Y SIN ESTUDIOS**	**ILLITERATE AND WITHOUT ANY STUDIES**	
01	Analfabetos	Illiterate	0
02	Sin estudios	Without any studies	
B	**ESTUDIOS PRIMARIOS, E.G.B.(1), CICLO INICIAL, MEDIO O PRIMERA ETAPA Y EQUIVALENTE**	**PRIMARY EDUCATION, INITIAL OR SECOND STAGE OF PRIMARY EDUCATION OR FIRST STAGE OR EQUIVALENT**	
03	Estudios primarios y E.G.B.(1), ciclos inicial, medio o primera etapa o equivalente	Primary education, initial or second stage of primary education or first stage or equivalent	5
C	**ENSEÑANZA GENERAL: 2º GRADO, 1ER CICLO. BACHILLER ELEMENTAL, E.G.B.(1), CICLO SUPERIOR, SEGUNDA ETAPA Y EDUCACIÓN SECUNDARIA (NUEVO SISTEMA)**	**GENERAL EDUCATION: SECOND YEAR, FIRST STAGE. TWO YEARS OF SECONDARY EDUCATION, PRIMARY EDUCATION, SUPERIOR STAGE, SECOND STAGE AND SECONDARY EDUCATION (NEW SYSTEM)**	
04	Bachiller elemental o equivalente; E.G.B(1), ciclo superior o 2ª etapa, y E.S.O. (2) (nuevo sistema)	Two years of secondary education or equivalent; Primary education, second or superior stage and Secondary education (new system).	8
05	Certificado de escolaridad o cualquier otro certificado que en el futuro acredite la asistencia a la escuela durante el período de escolaridad obligatoria (habiendo superado un mínimo de cursos)	Schooling certificate or any other certificate that is evidence in the future of school attendance during the compulsory education period (having passed a minimum number of courses)	
D	**ESTUDIOS TÉCNICO-PROFESIONALES DE PRIMER GRADO**	**TECHNICAL AND VOCATIONAL STUDIES FIRST STAGE**	
06	F.P.(3) de primer grado o equivalente	First stage of vocational Education and Training or equivalent	10
07	Otras enseñanzas técnico-profesionales de primer grado	Other first stage of technical and vocational teachings	
E	**ENSEÑANZA GENERAL: SEGUNDO GRADO, SEGUNCO CICLO. BACHILLERATO SUPERIOR, CON O SIN PREUNIVERSITARIO, B.U.P.(4), CON O SIN C.O.U.(5) Y BACHILLERATO (NUEVO SISTEMA)**	**GENERAL EDUCATION: SECOND STAGE. KEY STAGE 5 OR SIXTH FORM, WITH OR WITHOUT PREUNIVERSITY COURSE**	
08	Bachillerato superior, B.U.P.(4) y Bachillerato	Key Stage 5 or Sixth Form	11

Code	Spanish description	English description	Years of education
F	**ENSEÑANZA TÉCNICO-PROFESIONAL DE GRADO MEDIO**	**TECHNICAL AND VOCATIONAL STUDIES MIDDLE STAGE**	
09	Módulo 2 de F.P.[3]	2nd Module of Vocational Education and Training	13
G	**ENSEÑANZA TÉCNICO-PROFESIONAL DE SEGUNDO GRADO**	**TECHNICAL AND VOCATIONAL STUDIES SECOND STAGE**	
10	F.P.II[6] o equivalentes académicamente	2nd stage of Vocational Education and Training or academically equivalents	
11	Estudios de artes aplicadas y oficios artísticos	Arts and artistic jobs studies	13
12	Otras enseñanzas regladas equivalentes laboralmente o similares a la F.P.II [6]	Other formal trainings equivalent in relation to the labour market to 2nd stage of Vocational Education and Training	
H	**ENSEÑANZA TÉCNICO-PROFESIONAL DE GRADO SUPERIOR**	**TECHNICAL AND VOCATIONAL STUDIES SUPERIOR STAGE**	
13	Módulo 3 de F.P.[3]	3rd Module of Vocational Education and Training	15
J	**CARRERAS UNIVERSITARIAS DE CICLO CORTO**	**3-YEAR DEGREE**	
15	Área de Ingeniería y Tecnología	Engineering and Technology Area	
16	Área de Ciencias Médicas y de la Salud	Medical and Health Sciences Area	
17	Área de Humanidades	Humanities Area	15
18	Área de Ciencias Sociales y Jurídicas	Social Sciences and Law Area	
19	Área de Ciencias Exactas y Naturales	Exact Sciences and Biology Area	
K	**TRES CURSOS APROBADOS, SIN DERECHO A TITULACIÓN, DE UNA CARRERA DE CICLO LARGO**	**THREE PASSED COURSES OF A 5-YEAR DEGREE WITH NO RIGHT TO A TITLE**	
20	Tres cursos aprobados (o primer ciclo), sin derecho a titulación, de una carrera de ciclo largo	Three passed courses (or first cycle) of a 5-year degree with no right to a title	15
L	**CARRERAS UNIVERSITARIAS DE CICLO LARGO**	**5-YEAR DEGREES**	
21	Área de Ingeniería y Tecnología	Engineering and Technology Area	
22	Área de Ciencias Médicas y de la Salud	Medical and Health Sciences Area	
23	Área de Humanidades	Humanities Area	17
24	Área de Ciencias Sociales y Jurídicas	Social Sciences and Law Area	
25	Área de Ciencias Exactas y Naturales	Exact Sciences and Biology Area	

Code	Spanish description	English description	Years of education
M	**OTRAS ENSEÑANZAS SUPERIORES**	**OTHER HIGHER EDUCATION**	
26	Estudios equivalentes a todos los efectos a diplomado universitario	Equivalent studies to any 3-year degrees	15
27	Estudios superiores de dos o tres años de duración no equivalentes a diplomado universitario	Two or three-year-long superior studies not equivalent to 3-year degrees	
28	Estudios equivalentes a todos los efectos a licenciado universitario	Equivalent studies to any 5-year degree	17
29	Estudios superiores de al menos cuatro años de duración no equivalentes a licenciado universitario	At least four-year-long Superior Studies which are not equivalent to 5-year degrees	16
N	**DOCTORES**	**PhD (Doctorate)**	
30	Doctores en Ingeniería, Tecnología, Arquitectura, Ciencias médicas y de la Salud, Exactas y Naturales	PhD (Doctorate) in Engineering, Technology, Architecture, Medical and Health Sciences, Exact Sciences and Biology	21
31	Doctores en Humanidades y Ciencias Sociales	PhD (Doctorate) in Humanities and Social Sciences	

Notes:
(1) Educación General Básica
(2) Educación Secundaria Obligatoria
(3) Formación Profesional
(4) Bachillerato Unificado Polivalente
(5) Curso de Orientación a la Universidad
(6) Formación Profesional de Segundo Grado
Source: INE and my own translation based on information of the Spanish Embassy in United Kingdom (http://www.mecd.gob.es/dms-static/a8de9034-5c52-4347-b80a-69c78f0652d5/consejerias-exteriores/reino-unido/estudiar/cuadro-comparativo-inglaterra.pdf)

Table A2. 2. Educational Classification used by SLFS from 2000 to 2010 and correspondence in years of education

Code	Spanish description	English description	Years of education
80	Analfabetos	Illiterate	0
11	Estudios primarios incompletos	Incomplete primary studies	5
12	Estudios primarios completos	Complete primary studies	6
21	Programas para la formación e inserción laboral que no precisan de una titulación académica de la 1ª etapa de secundaria para su realización	Schemes for vocational trainings and employability that do not require any academic diploma of the 1st stage of secondary education to be carried out	7
22	Primera etapa de secundaria sin título	First stage of secondary education with no title	8
23	Primera etapa de secundaria con título	First stage of secondary education with a title	
31	Programas para la formación e inserción laboral que precisan de una titulación de estudios secundarios de 1ª etapa para su realización	Schemes for vocational trainings and employability that require a diploma of the 1st stage of secondary education to be carried out	9
32	Enseñanzas de bachillerato	Key Stage5 of sixth form teachings	12
33	Enseñanzas de grado medio de formación profesional específica, artes plásticas y diseño y deportivas	Teachings of middle grade of specific vocational education and training, plastic arts and design and sport	
34	Enseñanzas de grado medio de música y danza	Teachings of middle grade of specific vocational education and training of music and dance	
36(1)	Garantía social / Iniciación profesional	Social Guarantee / professional introduction	10
41	Programas para la formación e inserción laboral que precisan de una titulación de estudios secundarios de 2ª etapa para su realización	Schemes for vocational trainings and employability that require an academic diploma of the 2nd stage of secondary education to be carried out	13
50(2)	Grado	Degree (4 years)	16
51	Enseñanzas de grado superior de formación profesional específica y equivalentes: artes plásticas y diseño y deportivas	Certificate of Higher Education in specific vocational education and training and equivalent: plastic arts and design and sport	14
52	Títulos propios de las universidades que no sean de postgrado (dos años o más)	Degree diploma proper to universities (2 years or more)	
53	Programas que precisan de una titulación de formación profesional de grado superior para su realización (más de 300 horas o 6 meses)	Schemes that require a Certificate of Higher Education in specific vocational education and training (moret han 300 hours or 6 month)	
54	Enseñanzas universitarias de 1º ciclo y equivalentes o personas que han aprobado 3 cursos completos de una licenciatura o créditos equivalentes	3-years degree or equivalent or people who passed 3 full courses of a 5-year degree or equivalent credits	15

Code	Spanish description	English description	Years of education
55	Enseñanzas universitarias de 1º y 2º ciclo, de solo 2º ciclo y equivalentes	1st and 2nd cycles of university studies (3+2), only 2nd cycle university studies and equivalent	17
56	Programas oficiales de especialización profesional	Official schemes in specialization in vocational training	19
61	Doctorado universitario	University PhD (doctorate)	21

Notes:

(1) New category included the three month of 2003. “It gathers the people with a basic vocational education which allow their participation in the labour market, but who have not reached the objectives of the compulsory secondary education (ESO) and, therefore, they cannot be classified.” (INE, my own translation).

(2) New category included in 2009

Source: INE and my own translation based on information of the Spanish Embassy in United Kingdom (http://www.mecd.gob.es/dms-static/a8de9034-5c52-4347-b80a-69c78f0652d5/consejerias-exteriores/reino-unido/estudiar/cuadro-comparativo-inglaterra.pdf)

Table A2. 3. National Classification of Economic Activities (CNAE – 2009)

Code	Description	New Code
A	**Agriculture, forestry and fishing**	
01	Crop and animal production, hunting and related service activities	
02	Forestry and logging	0
03	Fishing and aquaculture	
B	**Mining and quarrying**	
05	Mining of coal and lignite	
06	Extraction of crude petroleum and natural gas	
07	Mining of metal ores	1
08	Other mining and quarrying	
09	Mining support service activities	
C	**Manufacturing**	
10	Manufacture of food products	
11	Manufacture of beverages	
13	Manufacture of textiles	
14	Manufacture of wearing apparel	
15	Manufacture of leather and related products	
16	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	
17	Manufacture of paper and paper products	
18	Printing and reproduction of recorded media	
19	Manufacture of coke and refined petroleum products	
20	Manufacture of chemicals and chemical products	
21	Manufacture of basic pharmaceutical products and pharmaceutical preparations	2
22	Manufacture of rubber and plastic products	
23	Manufacture of other non-metallic mineral products	
24	Manufacture of basic metals	
25	Manufacture of fabricated metal products, except machinery and equipment	
26	Manufacture of computer, electronic and optical products	
27	Manufacture of electrical equipment	
28	Manufacture of machinery and equipment n.e.c.	
29	Manufacture of motor vehicles, trailers and semi-trailers	
30	Manufacture of other transport equipment	
31	Manufacture of furniture	
32	Other manufacturing	
33	Repair and installation of machinery and equipment	

Code	Description	New Code
D	**Electricity, gas, steam and air conditioning supply**	
35	Electricity, gas, steam and air conditioning supply	3
E	**Water supply; sewerage; waste managment and remediation activities**	
36	Water collection, treatment and supply	
37	Sewerage	
38	Waste collection, treatment and disposal activities; materials recovery	3
39	Remediation activities and other waste management services	
F	**Construction**	
41	Construction of buildings	
42	Civil engineering	4
43	Specialised construction activities	
G	**Wholesale and retail trade; repair of motor vehicles and motorcycles**	
45	Wholesale and retail trade and repair of motor vehicles and motorcycles	
46	Wholesale trade, except of motor vehicles and motorcycles	5
47	Retail trade, except of motor vehicles and motorcycles	
H	**Transporting and storage**	
49	Land transport and transport via pipelines	
50	Water transport	
51	Air transport	6
52	Warehousing and support activities for transportation	
53	Postal and courier activities	
I	**Accommodation and food service activities**	
55	Accommodation	7
56	Food and beverage service activities	
J	**Information and communication**	
58	Publishing activities	
59	Motion picture, video and television programme production, sound recording and music publishing activities	
60	Programming and broadcasting activities	8
61	Telecommunications	
62	Computer programming, consultancy and related activities	
63	Information service activities	
K	**Financial and insurance activities**	
64	Financial service activities, except insurance and pension funding	
65	Insurance, reinsurance and pension funding, except compulsory social security	9
66	Activities auxiliary to financial services and insurance activities	

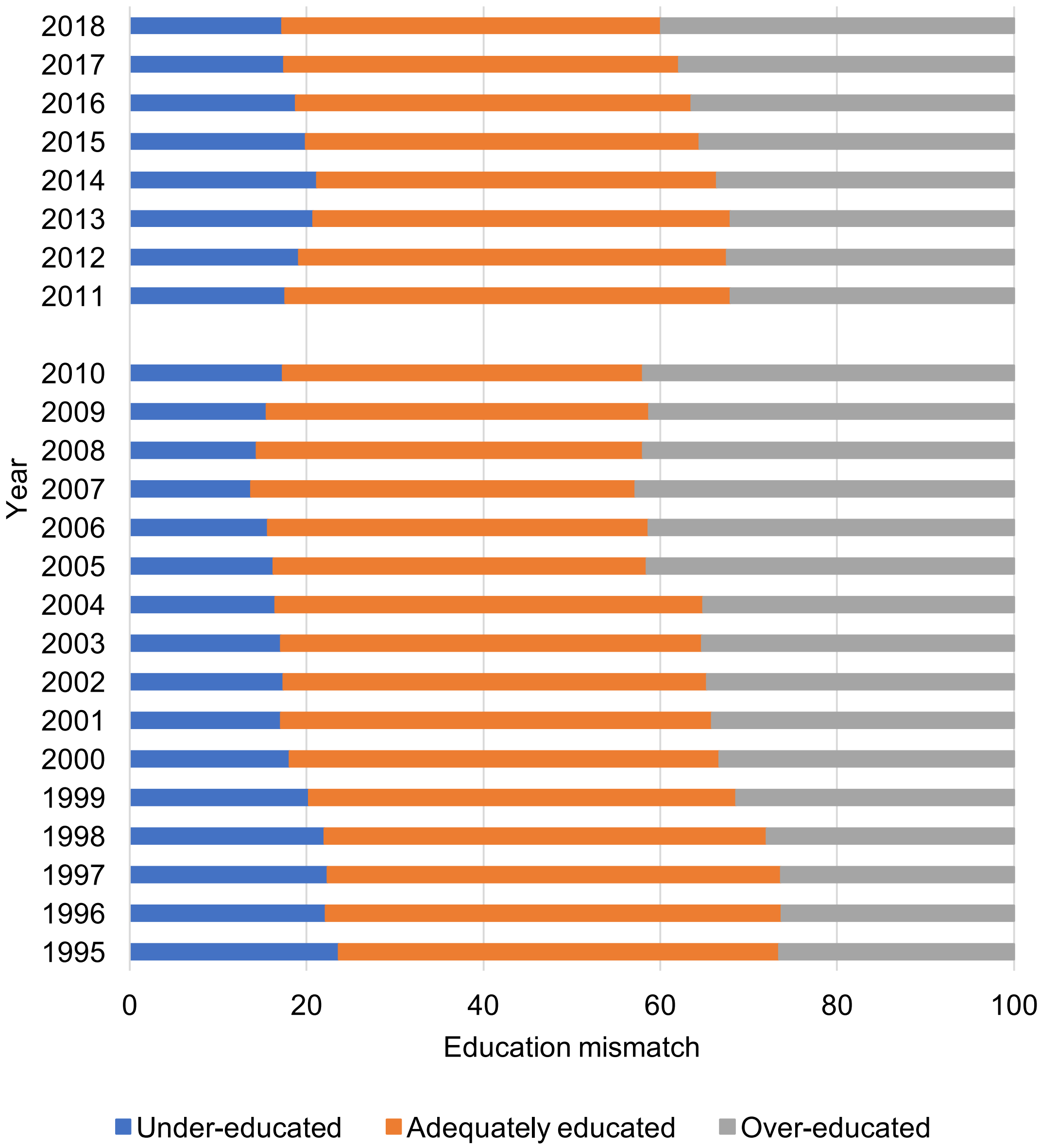
Code	Description	New Code
L	**Real estate activities**	
68	Real estate activities	10
M	**Professional, scientific and technical activities**	
69	Legal and accounting activities	
70	Activities of head offices; management consultancy activities	
71	Architectural and engineering activities; technical testing and analysis	
72	Scientific research and development	11
73	Advertising and market research	
74	Other professional, scientific and technical activities	
75	Veterinary activities	
N	**Administrative and support service activities**	
77	Rental and leasing activities	
78	Employment activities	
79	Travel agency, tour operator and other reservation service and related activities	12
80	Security and investigation activities	
81	Services to buildings and landscape activities	
82	Office administrative, office support and other business support activities	
O	**Public administration and defence; compulsory social security**	
84	Public administration and defence; compulsory social security	13
P	**Education**	
85	Education	14
Q	**Human health activities**	
86	Hospital activities	
87	Residential care activities	15
88	Social work activities without accommodation	
R	**Arts, entertainment and recreation**	
90	Creative, arts and entertainment activities	
91	Libraries, archives, museums and other cultural activities	16
92	Gambling and betting activities	
93	Sports activities and amusement and recreation activities	
S	**Other services activities**	
94	Activities of membership organisations	
95	Repair of computers and personal and household goods	17
96	Other personal service activities	
T	**producing activities of households for own use**	
97	Activities of households as employers of domestic personnel	18
98	Undifferentiated goods	
U	**Activities of extraterritorial organisations and bodies**	
99	Activities of extraterritorial organisations and bodies	19

Souce: https://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736177032&menu=ultiDatos&idp=1254735976614

Table A2. 4. Education mismatch in the Spanish labour market by gender. Three-year degree university graduates included (in percentages).

		Men	Women	Total
1995-2010	**Under-educated**	12.9	21.7	17.2
	Adequately educated	50.6	40.7	45.8
	Over-educated	36.5	37.5	37.0
2011-2018	**Under-educated**	15.5	21.9	19.0
	Adequately educated	49.4	42.8	45.8
	Over-educated	35.1	35.3	32.5

Source: SLFS (2nd quarters), INE



Source: SLFS (2nd quarters), INE

Figure A2. 1. Education mismatch among university graduates in the Spanish labour market. Three-year degree university graduates. Spain, 1995-2018 (in percentages).

Table A2. 5. Estimated coefficients for the probit model. Three-year degree university graduates included. Spain, 1995-2010

Y=(over-educated)

Independent variables (Reference category)		Coef.	Robust Std. Err.	Sig. Level
Woman (Man)		0.0656	0.0083	***
Age group (25-34)	**35-44**	-0.0080	0.0101	n.s.
	45-54	-0.1164	0.0114	***
	55-64	-0.1703	0.0158	***
Not Spanish (Spanish)		0.6526	0.0289	***
Activity sector (Agriculture, forestry, and fishing)	**Manufacturing**	-0.0709	0.0463	n.s.
	Construction	-0.1838	0.0490	***
	Sales	0.4927	0.0465	***
	Information and communication	-0.4593	0.0538	***
	Financial and insurance activities	0.4838	0.0484	***
	Real estate activities	0.4142	0.0703	***
	Professional, scientific, technical	-0.3212	0.0469	***
	Public administration	-0.7507	0.0450	***
	Other services	0.0099	0.0478	n.s.
Region (Madrid)	**North-West**	0.1048	0.0162	***
	North-East	0.1131	0.0151	***
	Centre	0.1173	0.0147	***
	East	-0.0048	0.0148	n.s.
	South	0.0478	0.0148	**
	Canary Islands	0.0838	0.0207	***
Year (1995)	**1996**	-0.0113	0.0280	n.s.
	1997	-0.0271	0.0276	n.s.
	1998	-0.0063	0.0272	n.s.
	1999	0.0480	0.0267	n.s.
	2000	0.1003	0.0265	***
	2001	0.0862	0.0264	**
	2002	0.0913	0.0264	**
	2003	0.0900	0.0257	***
	2004	0.0992	0.0257	***
	2005	0.2195	0.0263	***
	2006	0.2173	0.0254	***
	2007	0.2253	0.0251	***
	2008	0.2542	0.0252	***
	2009	0.2464	0.0252	***
	2010	0.2430	0.0250	***
Constant		0.0565	0.0505	n.s.
Number of observations		**170,336**		
Pseudo R2		**0.1200**		
Wald Chi-squared (35)		**14444.02**	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Note: Individual sample weights considered.
Source: SLFS (2nd quarters), INE

Table A2. 6. Estimated coefficients for the probit model. Three-year degree university graduates included. Spain, 2011-2018

Y=(over-educated)

Independent variables (Reference category)		Coef.	Robust Std. Err.	Sig. Level
Woman (Man)		-0.0389	0.0103	***
Age group (25-34)	**35-44**	-0.0793	0.0133	***
	45-54	-0.1193	0.0145	***
	55-64	-0.2080	0.0177	***
Not Spanish (Spanish)		0.6294	0.0255	***
Activity sector (Agriculture, forestry, and fishing)	**Manufacturing**	-0.5268	0.0635	***
	Construction	-0.5335	0.0692	***
	Sales	0.0853	0.0630	n.s.
	Information and communication	-0.9241	0.0680	***
	Financial and insurance activities	-0.5045	0.0654	***
	Real estate activities	-0.2128	0.0834	*
	Professional, scientific, technical	-0.5326	0.0638	***
	Public administration	-0.7452	0.0620	***
	Other services	-0.2546	0.0648	***
Region (Madrid)	**North-West**	0.0519	0.0183	**
	North-East	0.0313	0.0187	n.s.
	Centre	-0.0336	0.0184	n.s.
	East	-0.1135	0.0182	***
	South	-0.0645	0.0187	**
	Canary Islands	0.0512	0.0277	n.s.
Year (2011)	**2012**	0.0266	0.0214	n.s.
	2013	0.0614	0.0215	**
	2014	0.0993	0.0215	***
	2015	0.1428	0.0217	***
	2016	0.1733	0.0206	***
	2017	0.2269	0.0204	***
	2018	0.2585	0.0202	***
Constant		0.1319	0.0654	*
Number of observations		**112,981**		
Pseudo R2		**0.0646**		
Wald Chi-squared (27)		**4971.38**	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Note: Individual sample weights considered.
Source: SLFS (2nd quarters), INE

Table A2. 7. Estimated coefficients for the selection equation [3.2] of the Heckman two-step model. Three-year degree university graduates included. Spain, 1995-2010

Y=(employed)

Independent variables (Reference category)		Coef.	Robust Std. Err.	Sig. Level
Woman (Man)		-0.2745	0.0085	***
Age group (25-34)	**35-44**	0.1659	0.0110	***
	45-54	0.1021	0.0130	***
	55-64	-0.7393	0.0140	***
Not Spanish (Spanish)		-0.7089	0.0212	***
Person of reference	**Partner**	-0.2640	0.0107	***
	Children	-0.6665	0.0117	***
	Children-in-law	-0.3748	0.0556	***
	Grand-children	-0.9583	0.0546	***
	Parents	-0.3748	0.0921	***
	Other relatives	-0.3581	0.0369	***
	People working in the housework	6.6693	.	.
	Other people	0.0251	0.0571	n.s.
Region (Madrid)	**North-West**	-0.2067	0.0148	***
	North-East	-0.0860	0.0145	***
	Centre	-0.1735	0.0137	***
	East	0.0408	0.0145	**
	South	-0.2348	0.0137	***
	Canary Islands	-0.1103	0.0198	***
Year (1995)	**1996**	0.0276	0.0228	n.s.
	1997	0.0604	0.0225	**
	1998	0.0813	0.0224	***
	1999	0.1227	0.0224	***
	2000	0.1964	0.0225	***
	2001	0.2392	0.0222	***
	2002	0.2525	0.0226	***
	2003	0.2993	0.0223	***
	2004	0.3231	0.0221	***
	2005	0.3484	0.0229	***
	2006	0.3515	0.0230	***
	2007	0.4191	0.0227	***
	2008	0.4003	0.0227	***
	2009	0.3129	0.0223	***
	2010	0.2697	0.0222	***
Constant		1.1387	0.0222	***
Number of observations		**219,573**		
Pseudo R2		**0.0859**		

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Note: Individual sample weights considered.
Source: SLFS (2nd quarters), INE

Table A2. 8. Estimated coefficients for the selection equation [3.2] of the Heckman two-step model Three-year degree university graduates included. Spain, 2011-2018

(Y=Employed)

Independent variables (Reference category)		Coef.	Robust Std. Err.	Sig. Level
Woman (Man)		-0.2233	0.0099	***
Age group (25-34)	**35-44**	0.0653	0.0139	***
	45-54	-0.0174	0.0155	n.s.
	55-64	-0.6561	0.0161	***
Not Spanish (Spanish)		-0.7094	0.0199	***
Person of reference	**Partner**	-0.0006	0.1126	n.s.
	Children	-0.6540	0.1475	***
	Children-in-law	-0.2582	0.0697	***
	Grand-children	-0.6564	0.1158	***
	Parents	-0.2995	0.0775	***
	Other relatives	-0.4018	0.0436	***
	People working in the housework	6.6828	.	.
	Other people	0.0162	0.0668	n.s.
Region (Madrid)	**North-West**	-0.1839	0.0174	***
	North-East	-0.0217	0.0183	n.s.
	Centre	-0.1551	0.0173	***
	East	-0.0372	0.0175	n.s.
	South	-0.2674	0.0174	***
	Canary Islands	-0.2558	0.0254	***
Year (2011)	**2012**	-0.0656	0.0195	**
	2013	-0.1368	0.0195	***
	2014	-0.1221	0.0196	***
	2015	-0.0658	0.0195	**
	2016	-0.0093	0.0192	n.s.
	2017	0.0289	0.0192	n.s.
	2018	0.0738	0.0192	***
Constant		1.3285	0.0231	***
Number of observations		**146,143**		
Pseudo R2		**0.0727**		

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Note: Individual sample weights considered.
Source: SLFS (2nd quarters), INE

Table A2. 9. Estimated coefficients for the outcome equation [3.1] of the Heckman two-step model. Three-year degree university graduates included. Spain, 1995-2010
Y=(over-educated | employed)

Independent variables (Reference category)		Coef.	Robust Std. Err.	Sig. Level
Woman (Man)		-0.0043	0.0282	n.s.
Age group (25-34)	35-44	0.0696	0.0337	*
	45-54	-0.0372	0.0365	n.s.
	55-64	-0.2595	0.0304	***
Not Spanish (Spanish)		0.4796	0.0963	***
Activity sector (Agriculture, forestry, and fishing)	Manufacturing	-0.0631	0.0436	n.s.
	Construction	-0.1720	0.0471	***
	Sales	0.4646	0.0508	***
	Information and communication	-0.4331	0.0539	***
	Financial and insurance activities	0.4611	0.0506	***
	Real estate activities	0.3937	0.0688	***
	Professional, scientific, technical	-0.3077	0.0455	***
	Public administration	-0.7100	0.0506	***
	Other services	0.0142	0.0448	n.s.
Region (Madrid)	North-West	0.0522	0.0268	n.s.
	North-East	0.0890	0.0181	***
	Centre	0.0767	0.0221	**
	East	0.0052	0.0151	n.s.
	South	-0.0033	0.0242	n.s.
	Canary Islands	0.0567	0.0229	*
Year (1995)	1996	-0.0039	0.0270	n.s.
	1997	-0.0134	0.0272	n.s.
	1998	0.0102	0.0269	n.s.
	1999	0.0679	0.0264	*
	2000	0.1312	0.0270	***
	2001	0.1257	0.0283	***
	2002	0.1351	0.0289	***
	2003	0.1437	0.0306	***
	2004	0.1582	0.0316	***
	2005	0.2790	0.0306	***
	2006	0.2790	0.0307	***
	2007	0.3005	0.0338	***
	2008	0.3251	0.0322	***
	2009	0.3017	0.0286	***
	2010	0.2919	0.0273	***
Constant		-0.1467	0.0897	n.s.
Number of observations		219,573		
Selected observations		170,336		
Wald Chi-squared (35)		11576.55	***	
athrho		0.5671	*	
rho		0.5132	*	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Note: Individual sample weights considered.
Source: SLFS (2nd quarters), INE

Table A2. 10. Estimated coefficients for the outcome equation [3.1] of the Heckman two-step model Three-year degree university graduates included. Spain, 2011-2018

Y=(over-educated | employed)

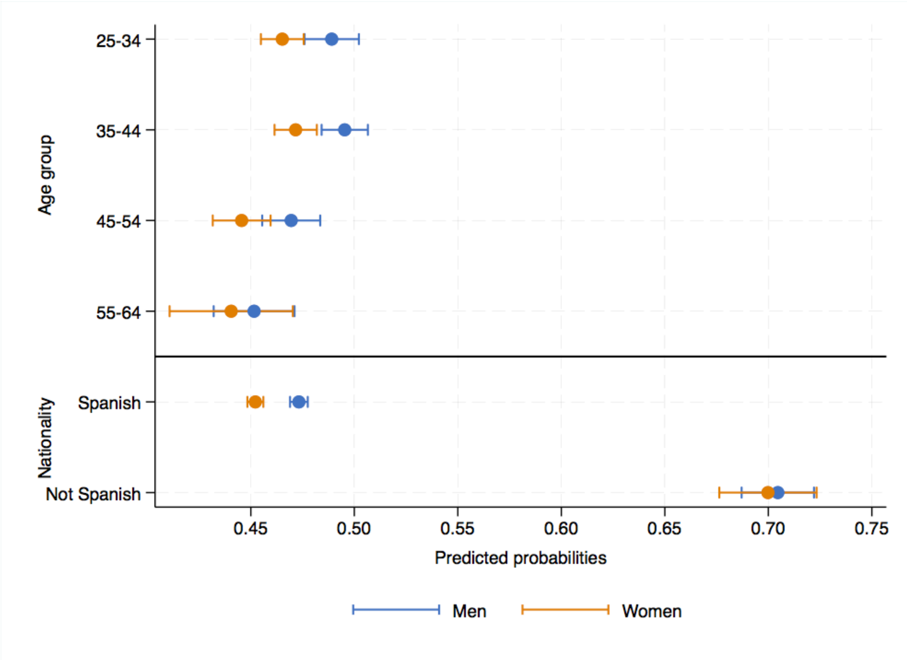
Independent variables (Reference category)		Coef.	Robust Std. Err.	Sig. Level
Woman (Man)		-0.0918	0.0094	***
Age group (25-34)	35-44	0.0189	0.0123	n.s.
	45-54	-0.0289	0.01345	*
	55-64	-0.306	0.0158	***
No Spanish (Spanish)		0.2424	0.0272	***
Activity sector (Agriculture, forestry, and fishing)	Manufacturing	-0.4360	0.0511	***
	Construction	-0.4339	0.0568	***
	Sales	0.0688	0.0506	n.s.
	Information and communication	-0.7914	0.0559	***
	Financial and insurance activities	-0.4145	0.0530	***
	Real estate activities	-0.1795	0.0679	**
	Professional, scientific, technical	-0.4445	0.0514	***
	Public administration	-0.6360	0.0497	***
	Other services	-0.1957	0.0521	***
Region (Madrid)	North-West	-0.0181	0.0168	n.s.
	North-East	0.0202	0.0172	n.s.
	Centre	-0.0770	0.0169	***
	East	-0.1079	0.0167	***
	South	-0.1458	0.0172	***
	Canary Islands	-0.0581	0.0256	*
Year (2011)	2012	0.0051	0.0196	n.s.
	2013	0.0119	0.0197	n.s.
	2014	0.0488	0.0197	*
	2015	0.1021	0.0200	***
	2016	0.1449	0.0189	***
	2017	0.2021	0.0187	***
	2018	0.2420	0.0186	***
Constant		-0.1069	0.0533	*
Number of observations		146,143		
Selected observations		112,981		
Wald Chi-squared (27)		4717.48	***	
athrho		1.4626	***	
rho		0.8982	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Note: Individual sample weights considered.
Source: SLFS (2nd quarters), INE

Table A2. 11. Predicted probabilities for a 25-64-year-old university graduate to be over-educated by year. Three-year degree university graduates included. Spain, 1995-2018

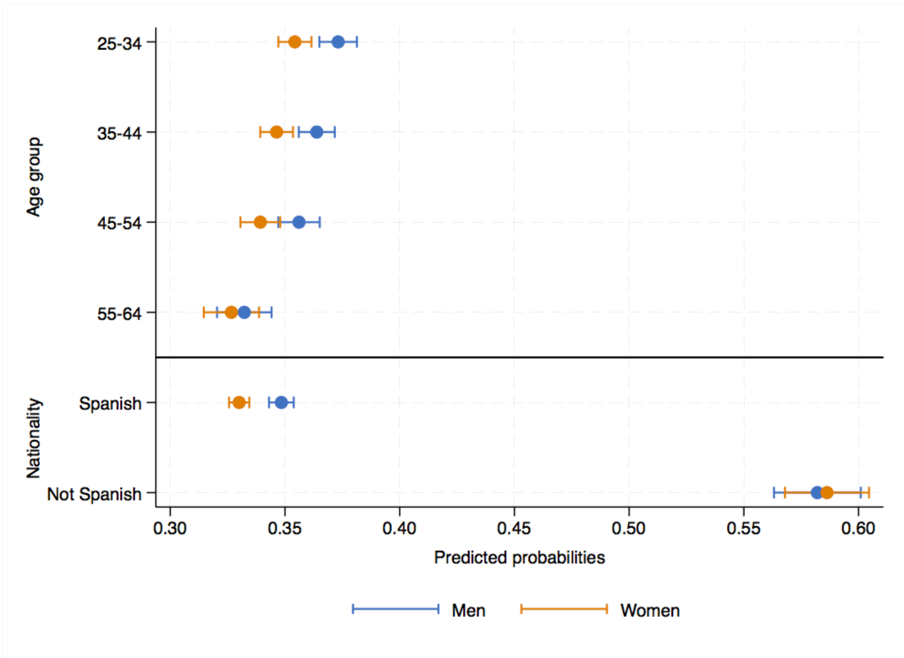
	Men			Women		
	Pred. Pr.	Std. Err.		Pred. Pr.	Std. Err.	
1995	0.4126	*0.0070*	***	0.4345	*0.0072*	***
1996	0.4615	*0.0094*	***	0.4794	*0.0095*	***
1997	0.4565	*0.0093*	***	0.4739	*0.0094*	***
1998	0.4635	*0.0092*	***	0.4807	*0.0092*	***
1999	0.4811	*0.0090*	***	0.4978	*0.0091*	***
2000	0.4990	*0.0089*	***	0.5149	*0.0090*	***
2001	0.4949	*0.0089*	***	0.5101	*0.0090*	***
2002	0.4970	*0.0089*	***	0.5121	*0.0089*	***
2003	0.4977	*0.0087*	***	0.5122	*0.0086*	***
2004	0.5013	*0.0088*	***	0.5155	*0.0085*	***
2005	0.5399	*0.0089*	***	0.5540	*0.0086*	***
2006	0.5397	*0.0086*	***	0.5538	*0.0083*	***
2007	0.5435	*0.0087*	***	0.5567	*0.0081*	***
2008	0.5520	*0.0086*	***	0.5655	*0.0081*	***
2009	0.5481	*0.0084*	***	0.5628	*0.0081*	***
2010	0.5466	*0.0083*	***	0.5619	*0.0080*	***
2011	0.3191	*0.0053*	***	0.3081	*0.0051*	***
2012	0.3689	*0.0072*	***	0.3608	*0.0072*	***
2013	0.3784	*0.0074*	***	0.3721	*0.0073*	***
2014	0.3901	*0.0074*	***	0.3837	*0.0074*	***
2015	0.4033	*0.0076*	***	0.3960	*0.0075*	***
2016	0.4127	*0.0072*	***	0.4043	*0.0071*	***
2017	0.4293	*0.0071*	***	0.4205	*0.0070*	***
2018	0.4390	*0.0071*	***	0.4294	*0.0070*	***

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Note: The probability of being over-educated is conditional on the probability of being employed as employee.
Source: SLFS (2nd quarters), INE



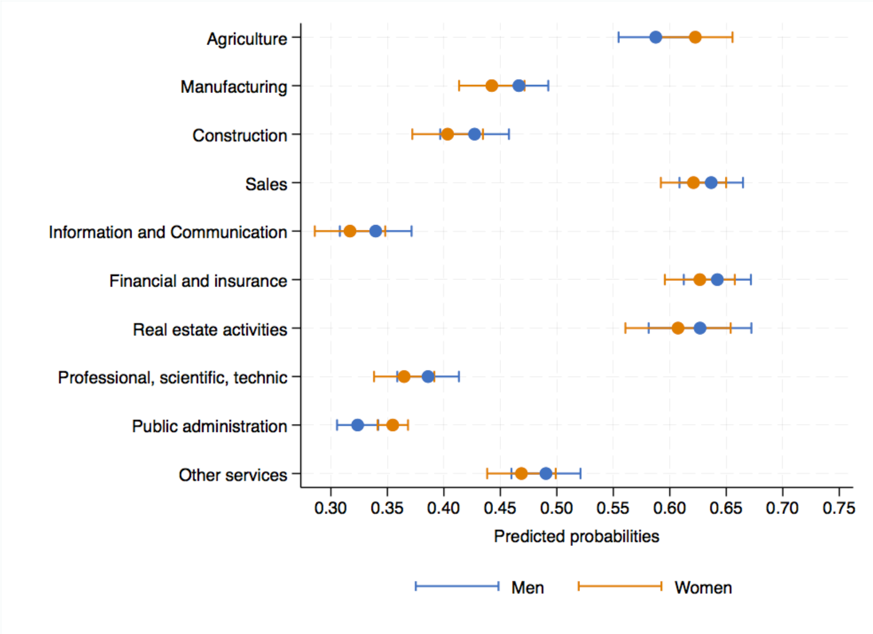
Note: The probability of being over-educated is conditional on the probability of being employed as employee.
Source: SLFS (2nd quarters), INE

Figure A2. 2. Predicted probabilities for a university graduate to be over-educated by age-group, nationality and gender. Three-year degree university graduates included. Spain, 1995-2010



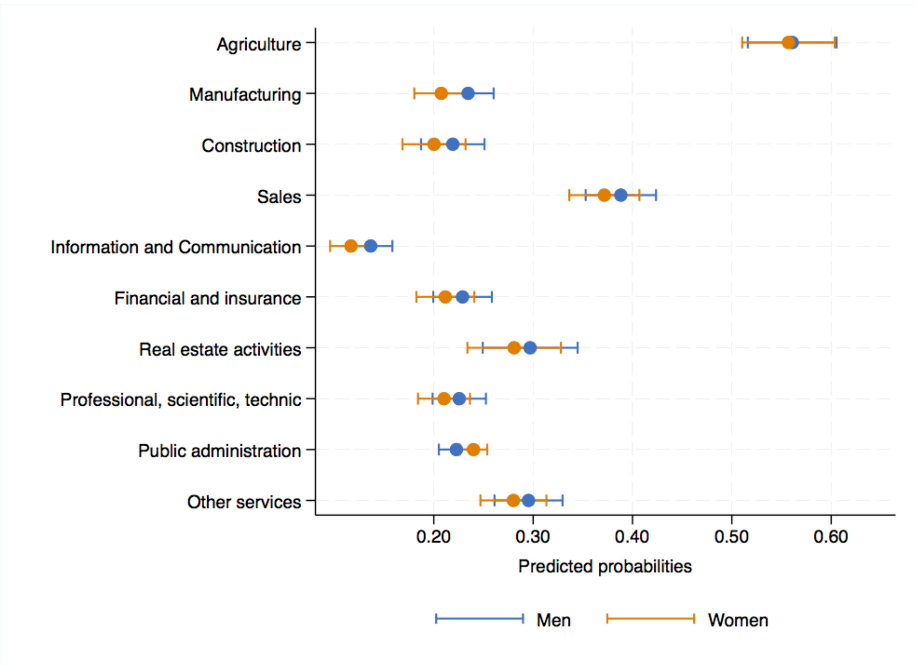
Note: The probability of being over-educated is conditional on the probability of being employed as employee.
Source: SLFS (2nd quarters), INE

Figure A2. 3. Predicted probabilities for a university graduate to be over-educated by age-group, nationality, and gender. Three-year degree university graduates included. Spain, 2011-2018



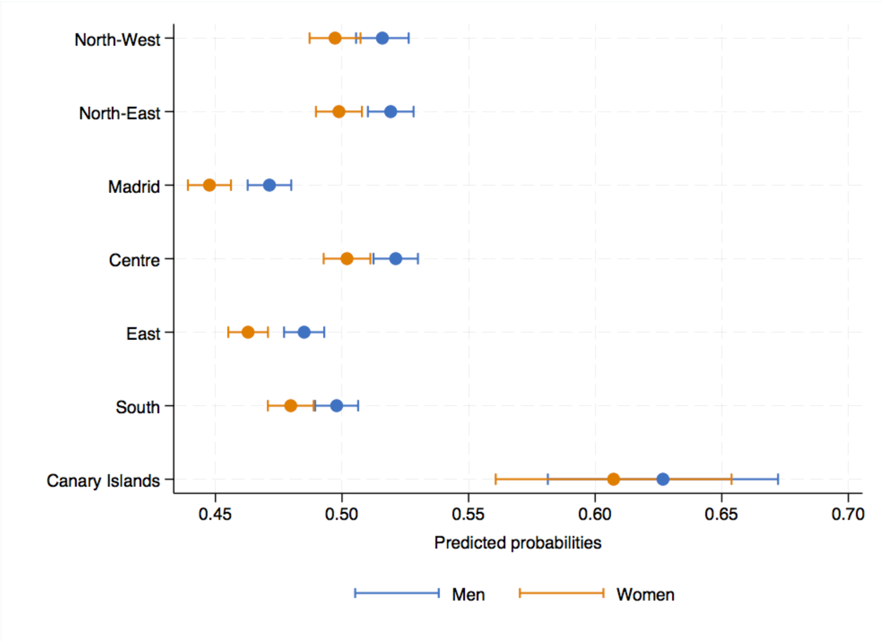
Note: The probability of being over-educated is conditional on the probability of being employed as employee.
Source: SLFS (2nd quarters), INE

Figure A2. 4. Predicted probabilities for a university graduate to be over-educated by activity sector and gender. Three-year degree university graduates included. Spain, 1995-2010



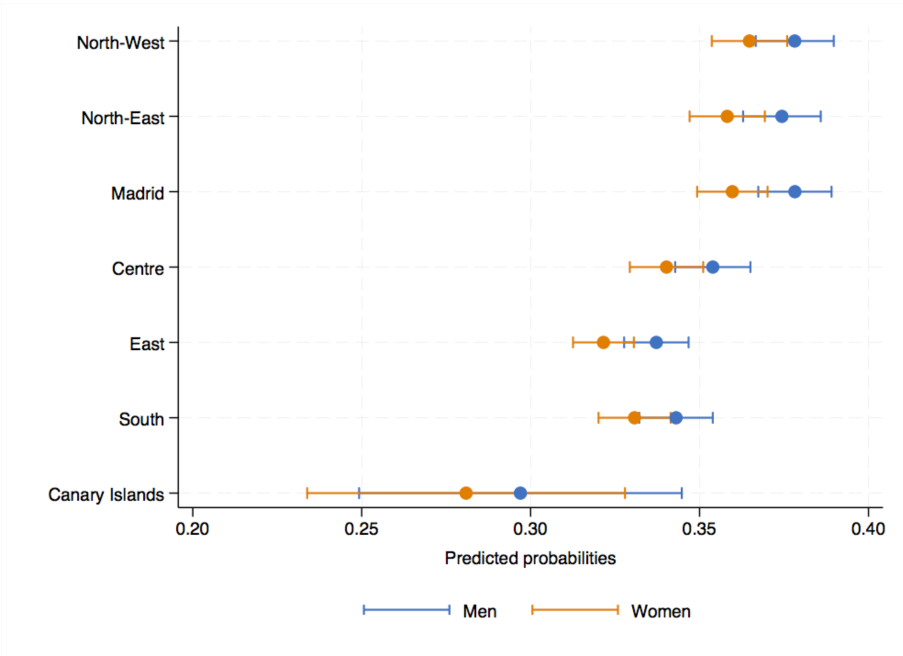
Note: The probability of being over-educated is conditional on the probability of being employed as employee.
Source: SLFS (2nd quarters), INE

Figure A2. 5. Predicted probabilities for a university graduate to be over-educated by activity sector and gender. Three-year degree university graduates included. Spain, 2011-2018



Note: The probability of being over-educated is conditional on the probability of being employed as employee.
Source: SLFS (2nd quarters), INE

Figure A2. 6. Predicted probabilities for a university graduate to be over-educated by region and gender. Three-year degree university graduates included. Spain, 1995-2010



Note: For the Heckman two-step model the probability of being over-educated is conditional on the probability of being employed as employee.
Source: SLFS (2nd quarters), INE

Figure A2. 7. Predicted probabilities for a university graduate to be over-educated by region and gender. Three-year degree university graduates included. Spain, 2011-2018

Appendix 3

Table A3. 1. Estimated coefficients of the probit model. Baseline specification and subjective approach

Independent variables (reference category)		Coef.	Jackknife Std. Err.	Sig. Level
Woman (Man)		0.1170	0.0579	*
Age-group (25-34)	**35-44**	0.0166	0.0806	n.s.
	45-54	0.0223	0.0715	n.s.
	55-64	-0.3914	0.0897	***
University graduates (Lower educational levels)		0.1026	0.0644	n.s.
Public		-0.0806	0.0772	n.s.
Region (Madrid)	**North-West**	0.0209	0.1186	n.s.
	North-East	0.0715	0.0910	n.s.
	Central Area	0.0754	0.1044	n.s.
	East	0.1548	0.0963	n.s.
	South	0.1190	0.0871	n.s.
	Canary Islands	0.2042	0.1688	n.s.
Constant		-0.4630	0.0916	***
Number of observations		**2,119**		
F(12,68)		**3.03**	**	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Table A3. 2. Estimated coefficients of the probit model. Literacy specification and subjective approach

Independent variables (reference category)		Coef.	Jackknife Std. Err.	Sig. Level
Woman (Man)		0.1362	0.0588	*
Age-group (25-34)	**35-44**	0.0182	0.0810	n.s.
	45-54	0.0435	0.0748	n.s.
	55-64	-0.3358	0.0963	**
University graduates (Lower educational levels)		0.0331	0.0732	n.s.
Literacy		0.0018	0.0009	*
Public (Private)		-0.0924	0.0775	n.s.
Region (Madrid)	**North-West**	0.0096	0.1191	n.s.
	North-East	0.0814	0.0923	n.s.
	Central Area	0.0719	0.1049	n.s.
	East	0.1527	0.0973	n.s.
	South	0.1223	0.0870	n.s.
	Canary Islands	0.2318	0.1671	n.s.
Constant		-0.9371	0.2489	***
Number of observations		**2,119**		
F(13,67)		**3.69**	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Table A3. 3. Estimated coefficients of the probit model. Numeracy specification and subjective approach

Independent variables (reference category)		Coef.	Jackknife Std. Err.	Sig. Level
Woman (Man)		0.1301	0.0606	*
Age-group (25-34)	**35-44**	0.0170	0.0807	n.s.
	45-54	0.0317	0.0749	n.s.
	55-64	-0.3685	0.0961	***
University graduates (Lower educational levels)		0.0751	0.0723	n.s.
Numeracy		0.0007	0.0008	n.s.
Public (Private)		-0.0859	0.0775	n.s.
Region (Madrid)	**North-West**	0.0200	0.1186	n.s.
	North-East	0.0754	0.0913	n.s.
	Central Area	0.0763	0.1045	n.s.
	East	0.1551	0.0965	n.s.
	South	0.1249	0.0868	n.s.
	Canary Islands	0.2213	0.1671	n.s.
Constant		-0.6597	0.2525	*
Number of observations		**2,119**		
F(13,67)		**3.33**	***	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Table A3. 4. Estimated coefficients of the Heckman two-step model. Baseline specification and subjective approach

Independent variables (Reference category)			Coef.	Jackknife Std. Err.	Sig. Level
Outcome equation	Woman (Man)		0.1749	0.1297	n.s.
	Age-group (25-34)	35-44	0.0119	0.0798	n.s.
		45-54	0.0415	0.0791	n.s.
		55-64	-0.2269	0.4226	n.s.
	University graduates (Less than university-education)		-0.0182	0.2610	n.s.
	Public sector (Private sector)		-0.0811	0.0753	n.s.
	Region (Madrid)	North-West	0.0336	0.1294	n.s.
		North-East	0.0690	0.0904	n.s.
		Centre	0.1080	0.1199	n.s.
		East	0.1698	0.0976	n.s.
		South	0.1762	0.1455	n.s.
		Canary Islands	0.3060	0.2637	n.s.
	Constant		-0.3100	0.3598	n.s.
Selection equation	Woman (Man)		-0.3512	0.0449	***
	Age-group (25-34)	35-44	0.0635	0.0693	n.s.
		45-54	-0.0591	0.0654	n.s.
		55-64	-0.7336	0.0682	***
	University graduates (Less than university-education)		0.7099	0.0580	***
	Good-excellent health (Poor-fair health)		-0.1863	0.0572	**
	Region (Madrid)	North-West	-0.1201	0.0944	n.s.
		North-East	0.0003	0.0940	n.s.
		Centre	-0.1902	0.0784	*
		East	-0.1173	0.0912	n.s.
		South	-0.3697	0.0784	***
		Canary Islands	-0.6008	0.1356	***
	Constant		0.7384	0.0839	***
Number of observations			**3,571**		
F (12, 68)			**2.86**	***	
athrho			**-0.3316**	**n.s.**	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Table A3. 5. Estimated coefficients of the Heckman two-step model. Literacy specification and subjective approach

	Independent variables (Reference category)		Coef.	Std. Err. (1)	Sig. Level
Outcome equation	Woman (Man)		0.1687	0.1960	n.s.
	Age-group (25-34)	35-44	0.0141	0.0798	n.s.
		45-54	0.0495	0.0884	n.s.
		55-64	-0.2494	0.5069	n.s.
	University graduates (Less than university-education)		-0.0251	0.3095	n.s.
	Literacy		0.0014	0.0024	n.s.
	Public sector (Private sector)		-0.0934	0.0731	n.s.
	Region (Madrid)	North-West	0.0155	0.1376	n.s.
		North-East	0.0782	0.1043	n.s.
		Centre	0.0934	0.1607	n.s.
		East	0.1631	0.1161	n.s.
		South	0.1565	0.2157	n.s.
		Canary Islands	0.2895	0.3675	n.s.
	Constant		-0.7343	1.1005	n.s.
Selection equation	Woman (Man)		-0.3384	0.0529	***
	Age-group (25-34)	35-44	0.0740	0.0708	n.s.
		45-54	-0.0125	0.0737	n.s.
		55-64	-0.6339	0.0802	***
	University graduates (Less than university-education)		0.5708	0.0705	***
	Numeracy		0.0034	0.0007	***
	Good-excellent health (Poor-fair health)		-0.1575	0.0634	*
	Region (Madrid)	North-West	-0.1410	0.0905	n.s.
		North-East	0.0080	0.0908	n.s.
		Centre	-0.2189	0.0884	*
		East	-0.1365	0.0846	n.s.
		South	-0.3626	0.0805	***
		Canary Islands	-0.5670	0.1348	***
	Constant		-0.1354	0.1958	n.s.
Number of observations			3,571		
F (13, .)			.		
athrho			-0.207	n.s.	

(1) The estimation of jackknife standard errors were not possible for the literacy specification when the subjective education mismatch method was used.
*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Table A3. 6. Estimated coefficients of the Heckman two-step model. Numeracy specification and subjective approach

Equation	Independent variables (Reference category)		Coef.	Jackknife Std. Err.	Sig. Level
Outcome equation	Woman (Man)		0.1825	0.1060	n.s.
	Age-group (25-34)	35-44	0.0091	0.0798	n.s.
		45-54	0.0389	0.0730	n.s.
		55-64	-0.2073	0.3637	n.s.
	University graduates (Less than university-education)		-0.0207	0.1719	n.s.
	Numeracy		-0.0003	0.0020	n.s.
	Public sector (Private sector)		-0.0868	0.0740	n.s.
	Region (Madrid)	North-West	0.0342	0.1288	n.s.
		North-East	0.0681	0.0907	n.s.
		Centre	0.1125	0.1168	n.s.
		East	0.1720	0.0968	n.s.
		South	0.1813	0.1377	n.s.
		Canary Islands	0.3148	0.2414	n.s.
	Constant		-0.2250	0.8488	n.s.
Selection equation	Woman (Man)		-0.2975	0.0478	***
	Age-group (25-34)	35-44	0.0739	0.0699	n.s.
		45-54	0.0105	0.0681	n.s.
		55-64	-0.5972	0.0739	***
	University graduates (Less than university-education)		0.5158	0.0589	***
	Numeracy		0.0046	0.0006	***
	Good-excellent health (Poor-fair health)		-0.1561	0.0579	**
	Region (Madrid)	North-West	-0.1169	0.0933	n.s.
		North-East	0.0116	0.0938	n.s.
		Centre	-0.2016	0.0801	*
		East	-0.1259	0.0900	n.s.
		South	-0.3292	0.0783	***
		Canary Islands	-0.5076	0.1335	***
	Constant		-0.4623	0.1673	**
Number of observations			**3,571**		
F (13, .)			**.**		
athrho			**-0.3850**	**n.s.**	

*Significance levels: *p<0.05, **p<0.01, ***p<0.001, n.s. not significant.*
Source: PIAAC (http://www.oecd.org/skills/piaac/publicdataandanalysis/)

Appendix 4

Table A4. 1. Forecast's results for 25-64-year-old university graduate men living in the North-West area

A. University graduates who are over-educated (in thousands)

Year	Pop.	Stage1	Stage 2		Stage 3		
		U1	W1	W2	HSJ1	HSJ2	HSJ3
2014	1,230	151	121	121	91	91	91
2015	1,217	153	125	125	93	93	93
2016	1,203	155	128	128	91	91	91
2017	1,191	163	138	138	98	98	98
2018	1,178	170	144	144	102	102	102
2019	1,166	176	149	150	105	106	106
2020	1,153	182	154	156	108	110	110
2021	1,139	187	158	162	111	114	114
2022	1,125	192	163	168	114	118	118
2023	1,111	196	167	173	117	122	122
2024	1,097	201	171	179	120	125	126
2025	1,083	205	175	184	122	129	130
2026	1,070	208	179	189	125	132	134
2027	1,057	212	182	195	127	136	137
2028	1,045	216	186	200	130	139	141
2029	1,033	219	190	205	132	142	145
2030	1,022	223	194	211	134	146	149

Note: Pop. is population, U1 is the number of university graduates (hypothesis 1); Wi is the number of university graduates employed (hypothesis i); HSJj is the number of university graduates employed in highly skilled jobs (hypothesis j).
Actual data in blue and the projection's results in black.

B. University graduates who are over-educated (in percentages)

Year	HSJ/W		
	S1	S2	S3
2014	24.2	24.2	24.2
2015	25.5	25.5	25.5
2016	28.8	28.8	28.8
2017	29.2	29.2	29.2
2018	29.5	29.5	29.5
2019	29.5	29.5	29.4
2020	29.6	29.5	29.4
2021	29.6	29.6	29.4
2022	29.7	29.7	29.4
2023	29.8	29.7	29.4
2024	29.9	29.9	29.4
2025	30.0	30.0	29.3
2026	30.2	30.1	29.3
2027	30.4	30.3	29.3
2028	30.5	30.5	29.3
2029	30.7	30.7	29.3
2030	30.9	30.9	29.3

Note: Si refers to Scenario i. Actual data in blue and the projection's results in black.
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Table A4. 2. Forecast's results for 25-64-year-old university graduate women living in the North-West area

A. University graduates who are over-educated (in thousands)

Year	Pop.	Stage1	Stage 2		Stage 3		
		U1	W1	W2	HSJ1	HSJ2	HSJ3
2014	1,247	192	148	148	104	104	104
2015	1,236	194	147	147	104	104	104
2016	1,224	210	168	168	114	114	114
2017	1,213	209	172	172	111	111	111
2018	1,203	209	168	168	113	113	113
2019	1,188	203	164	166	109	111	111
2020	1,177	199	161	166	107	110	109
2021	1,165	194	157	164	104	108	107
2022	1,151	189	154	162	101	107	105
2023	1,138	183	150	160	98	105	103
2024	1,124	178	146	158	96	103	101
2025	1,111	173	142	156	93	101	99
2026	1,099	168	138	154	90	99	97
2027	1,087	163	135	152	88	97	95
2028	1,075	158	131	150	85	95	93
2029	1,063	154	128	148	82	94	92
2030	1,051	149	124	145	80	92	90

Note: Pop. is population, U1 is the number of university graduates (hypothesis 1); Wi is the number of university graduates employed (hypothesis i); HSJj is the number of university graduates employed in highly skilled jobs (hypothesis j)
Actual data in blue and the projection's results in black.

B. University graduates who are over-educated (in percentages)

Year	HSJ/W		
	S1	S2	S3
2014	29.9	29.9	29.9
2015	29.5	29.5	29.5
2016	32.2	32.2	32.2
2017	35.5	35.5	35.5
2018	33.1	33.1	33.1
2019	33.3	33.4	33.6
2020	33.5	33.8	34.2
2021	33.8	34.1	34.7
2022	34.0	34.4	35.2
2023	34.2	34.8	35.7
2024	34.4	35.1	36.1
2025	34.6	35.4	36.5
2026	34.8	35.7	36.9
2027	35.0	36.0	37.3
2028	35.2	36.4	37.7
2029	35.4	36.7	38.0
2030	35.6	36.9	38.3

Note: Si refers to Scenario i. Actual data in blue and the projection's results in black.
Source: SLFS (2^{nd} quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Table A4. 3. Forecast's results for 25-64-year-old university graduate men living in the North-East area

A. University graduates who are over-educated (in thousands)

Year	Pop.	Stage1	Stage 2		Stage 3		
		U1	W1	W2	HSJ1	HSJ2	HSJ3
2014	1,251	194	169	169	122	122	122
2015	1,241	193	163	163	116	116	116
2016	1,224	194	169	169	116	116	116
2017	1,215	192	167	167	120	120	120
2018	1,212	192	171	171	118	118	118
2019	1,206	190	170	172	118	119	118
2020	1,197	188	169	172	118	121	119
2021	1,188	185	168	172	118	121	119
2022	1,178	182	166	172	118	122	118
2023	1,168	179	165	172	118	123	118
2024	1,157	177	163	172	118	124	118
2025	1,148	174	162	172	117	124	118
2026	1,139	171	160	172	117	125	118
2027	1,130	169	159	172	117	126	118
2028	1,122	166	158	171	117	127	118
2029	1,113	164	156	171	117	127	118
2030	1,105	162	155	171	117	128	118

Note: Pop. is population, U1 is the number of university graduates (hypothesis 1); Wi is the number of university graduates employed (hypothesis i); HSJj is the number of university graduates employed in highly skilled jobs (hypothesis j)
Actual data in blue and the projection's results in black.

B. University graduates who are over-educated (in percentages)

Year	HSJ/W		
	S1	S2	S3
2014	27.8	27.8	27.8
2015	28.6	28.6	25.6
2016	31.5	31.5	31.5
2017	28.5	28.5	28.5
2018	30.8	30.8	30.8
2019	30.4	30.4	31.0
2020	29.9	29.9	31.1
2021	29.4	29.5	31.1
2022	28.9	29.0	31.2
2023	28.4	28.5	31.2
2024	27.9	28.0	31.2
2025	27.4	27.6	31.3
2026	26.9	27.1	31.3
2027	26.4	26.6	31.3
2028	25.9	26.1	31.3
2029	25.4	25.7	31.2
2030	24.9	25.2	31.2

Note: Si refers to Scenario i. Actual data in blue and the projection's results in black.
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Table A4. 4. Forecast's results for 25-64-year-old university graduate women living in the North-East area

A. University graduates who are over-educated (in thousands)

Year	Pop.	Stage1	Stage 2		Stage 3		
		U1	W1	W2	HSJ1	HSJ2	HSJ3
2014	1,235	223	182	182	118	118	118
2015	1,228	232	188	188	125	125	125
2016	1,218	237	193	193	130	130	130
2017	1,214	255	211	211	143	143	143
2018	1,211	262	229	229	151	151	151
2019	1,208	274	247	242	160	157	163
2020	1,203	285	265	255	169	163	175
2021	1,198	296	284	268	177	168	187
2022	1,192	307	302	281	186	173	200
2023	1,186	316	321	293	194	178	212
2024	1,180	326	340	306	202	183	225
2025	1,173	335	359	318	210	187	238
2026	1,167	345	379	331	218	192	252
2027	1,162	354	399	344	226	196	266
2028	1,156	363	419	357	234	201	281
2029	1,150	371	440	369	241	205	296
2030	1,144	380	460	382	249	209	311

Note: Pop. is population, U1 is the number of university graduates (hypothesis 1); Wi is the number of university graduates employed (hypothesis i); HSJj is the number of university graduates employed in highly skilled jobs (hypothesis j)
Actual data in blue and the projection's results in black.

B. University graduates who are over-educated (in percentages)

Year	HSJ/W		
	S1	S2	S3
2014	34.9	34.9	34.9
2015	33.4	33.4	33.4
2016	32.4	32.4	32.4
2017	32.2	32.2	32.2
2018	34.0	34.0	34.0
2019	35.2	35.1	32.7
2020	36.4	36.2	31.5
2021	37.5	37.2	30.2
2022	38.6	38.2	28.9
2023	39.6	39.2	27.6
2024	40.6	40.1	26.3
2025	41.5	41.1	25.0
2026	42.5	42.0	23.7
2027	43.4	42.9	22.5
2028	44.3	43.7	21.2
2029	45.2	44.6	20.0
2030	46.0	45.5	18.7

Note: Si refers to Scenario i. Actual data in blue and the projection's results in black.
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Table A4. 5. Forecast's results for 25-64-year-old university graduate men living in the Madrid area

A. University graduates who are over-educated (in thousands)

Year	Pop.	Stage1	Stage 2		Stage 3		
		U1	W1	W2	HSJ1	HSJ2	HSJ3
2014	1,786	454	394	394	295	295	295
2015	1,781	475	415	415	309	309	309
2016	1,781	476	413	413	314	314	314
2017	1,784	484	425	425	333	333	333
2018	1,801	489	441	441	333	333	333
2019	1,828	505	465	466	350	351	354
2020	1,835	514	482	484	362	364	370
2021	1,841	523	500	503	375	377	386
2022	1,846	531	517	522	387	390	403
2023	1,850	540	535	540	399	403	420
2024	1,854	548	553	559	412	416	437
2025	1,858	556	572	579	425	430	455
2026	1,863	565	591	599	438	444	473
2027	1,868	574	610	620	451	458	492
2028	1,872	582	629	640	464	472	512
2029	1,876	591	649	661	478	487	532
2030	1,879	599	668	682	491	501	552

Note: Pop. is population, U1 is the number of university graduates (hypothesis 1); Wi is the number of university graduates employed (hypothesis i); HSJj is the number of university graduates employed in highly skilled jobs (hypothesis j)
Actual data in blue and the projection's results in black.

B. University graduates who are over-educated (in percentages)

Year	HSJ/W		
	S1	S2	S3
2014	25.1	25.1	25.1
2015	25.6	25.6	25.6
2016	24.1	24.1	24.1
2017	21.7	21.7	21.7
2018	24.5	24.5	24.5
2019	24.7	24.7	24.1
2020	24.9	24.9	23.6
2021	25.0	25.0	23.2
2022	25.2	25.2	22.8
2023	25.4	25.4	22.3
2024	25.6	25.6	21.9
2025	25.7	25.7	21.4
2026	25.9	25.9	21.0
2027	26.1	26.1	20.5
2028	26.2	26.2	20.1
2029	26.4	26.4	19.6
2030	26.6	26.6	19.1

Note: Si refers to Scenario i. Actual data in blue and the projection's results in black.
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Table A4. 6.Forecast's results for 25-64-year-old university graduate women living in the Madrid area

A. University graduates who are over-educated (in thousands)

Year	Pop.	Stage1	Stage 2		Stage 3		
		U1	W1	W2	HSJ1	HSJ2	HSJ3
2014	1,891	496	386	386	263	263	263
2015	1,887	538	428	428	292	292	292
2016	1,891	545	437	437	299	299	299
2017	1,898	534	445	445	302	302	302
2018	1,919	534	444	444	293	293	293
2019	1,945	540	459	460	296	297	305
2020	1,956	539	467	469	296	297	313
2021	1,968	539	476	478	296	297	321
2022	1,978	537	483	486	294	296	329
2023	1,987	535	490	494	292	295	336
2024	1,997	533	497	502	291	293	344
2025	2,006	531	504	509	288	291	351
2026	2,016	529	511	517	286	289	359
2027	2,025	527	518	524	283	287	366
2028	2,034	524	524	531	280	284	374
2029	2,041	521	529	537	276	281	380
2030	2,047	517	534	543	272	277	387

Note: Pop. is population, U1 is the number of university graduates (hypothesis 1); Wi is the number of university graduates employed (hypothesis i); HSJj is the number of university graduates employed in highly skilled jobs (hypothesis j)
Actual data in blue and the projection's results in black.

B. University graduates who are over-educated (in percentages)

Year	HSJ/W		
	S1	S2	S3
2014	32.0	32.0	32.0
2015	31.8	31.8	31.8
2016	31.6	31.6	31.6
2017	32.2	32.2	32.2
2018	34.2	34.2	34.2
2019	35.4	35.4	33.7
2020	36.6	36.6	33.3
2021	37.9	37.9	32.9
2022	39.1	39.1	32.4
2023	40.4	40.3	32.0
2024	41.6	41.6	31.5
2025	42.8	42.8	31.1
2026	44.1	44.1	30.6
2027	45.3	45.3	30.1
2028	46.5	46.5	29.7
2029	47.8	47.8	29.2
2030	49.0	49.0	28.8

Note: Si refers to Scenario i. Actual data in blue and the projection's results in black.
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Table A4. 7. Forecast's results for 25-64-year-old university graduate men living in the Central area

A. University graduates who are over-educated (in thousands)

Year	Pop.	Stage1	Stage 2		Stage 3		
		U1	W1	W2	HSJ1	HSJ2	HSJ3
2014	1,595	165	139	139	100	100	100
2015	1,584	174	146	146	104	104	104
2016	1,568	174	136	136	101	101	101
2017	1,553	174	148	148	103	103	103
2018	1,538	178	154	154	102	102	102
2019	1,531	180	156	157	101	101	103
2020	1,516	179	158	159	100	100	105
2021	1,500	179	159	160	98	99	106
2022	1,481	178	159	162	96	98	106
2023	1,462	177	160	163	94	96	107
2024	1,442	176	160	163	92	94	108
2025	1,423	174	160	164	90	92	108
2026	1,403	173	160	165	88	91	109
2027	1,385	172	161	165	86	89	109
2028	1,366	170	161	166	84	87	109
2029	1,347	169	160	166	82	85	110
2030	1,329	167	160	167	80	83	110

Note: Pop. is population, U1 is the number of university graduates (hypothesis 1); Wi is the number of university graduates employed (hypothesis i); HSJj is the number of university graduates employed in highly skilled jobs (hypothesis j)
Actual data in blue and the projection's results in black.

B. University graduates who are over-educated (in percentages)

Year	HSJ/W		
	S1	S2	S3
2014	27.6	27.6	27.6
2015	29.1	29.1	29.1
2016	26.2	26.2	26.2
2017	30.5	30.5	30.5
2018	34.0	34.0	34.0
2019	35.4	35.4	34.0
2020	36.8	36.8	34.1
2021	38.1	38.2	34.1
2022	39.5	40.0	34.1
2023	40.8	40.9	34.1
2024	42.2	42.3	34.1
2025	43.5	43.7	34.1
2026	45.0	45.0	34.1
2027	46.2	46.3	34.1
2028	47.5	47.7	34.0
2029	48.8	49.0	34.0
2030	50.2	50.3	33.9

Note: Si refers to Scenario i. Actual data in blue and the projection's results in black.
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Table A4. 8. Forecast's results for 25-64-year-old university graduate women living in the Central area

A. University graduates who are over-educated (in thousands)

Year	Pop.	Stage1	Stage 2		Stage 3		
		U1	W1	W2	HSJ1	HSJ2	HSJ3
2014	1,514	199	144	144	92	92	92
2015	1,507	201	149	149	96	96	96
2016	1,496	200	154	154	102	102	102
2017	1,486	220	167	167	115	115	115
2018	1,476	225	176	176	117	117	117
2019	1,469	234	181	183	121	122	121
2020	1,458	242	185	188	124	126	124
2021	1,448	250	189	194	127	130	127
2022	1,434	257	192	199	130	134	129
2023	1,419	262	195	203	132	137	132
2024	1,404	267	197	207	134	140	134
2025	1,387	272	199	210	136	143	136
2026	1,371	276	201	214	138	146	138
2027	1,355	279	202	217	140	149	139
2028	1,338	283	204	220	141	152	141
2029	1,320	285	204	222	143	154	142
2030	1,304	287	205	225	144	157	143

Note: Pop. is population, U1 is the number of university graduates (hypothesis 1); Wi is the number of university graduates employed (hypothesis i); HSJj is the number of university graduates employed in highly skilled jobs (hypothesis j)
Actual data in blue and the projection's results in black.

B. University graduates who are over-educated (in percentages)

Year	HSJ/W		
	S1	S2	S3
2014	36.4	36.4	36.4
2015	35.4	35.4	35.4
2016	34.1	34.1	34.1
2017	31.5	31.5	31.5
2018	33.6	33.6	33.6
2019	33.4	33.4	33.9
2020	33.1	33.2	34.3
2021	32.9	33.0	34.6
2022	32.6	32.7	34.8
2023	32.3	32.5	35.1
2024	32.0	32.2	35.3
2025	31.6	31.9	35.5
2026	31.3	31.6	35.7
2027	31.0	31.3	35.8
2028	30.6	31.0	36.0
2029	30.3	30.7	36.1
2030	29.9	30.4	36.2

Note: Si refers to Scenario i. Actual data in blue and the projection's results in black.
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Table A4. 9. Forecast's results for 25-64-year-old university graduate men living in the East area

A. University graduates who are over-educated (in thousands)

Year	Pop.	Stage1	Stage 2		Stage 3		
		U1	W1	W2	HSJ1	HSJ2	HSJ3
2014	3,828	506	414	414	295	295	295
2015	3,793	517	448	448	312	312	312
2016	3,774	543	478	478	335	335	335
2017	3,768	570	507	507	355	355	355
2018	3,790	615	540	540	365	365	365
2019	3,817	651	571	590	381	395	390
2020	3,826	684	599	639	393	422	412
2021	3,827	715	625	687	404	449	434
2022	3,829	747	651	738	415	475	456
2023	3,829	778	677	790	424	501	478
2024	3,828	809	703	843	433	527	500
2025	3,829	840	728	899	441	552	522
2026	3,830	872	755	956	448	577	544
2027	3,831	903	781	1,016	455	602	567
2028	3,832	934	806	1,076	461	626	590
2029	3,829	965	832	1,138	466	649	612
2030	3,828	996	857	1,202	470	671	635

Note: Pop. is population, U1 is the number of university graduates (hypothesis 1); Wi is the number of university graduates employed (hypothesis i); HSJj is the number of university graduates employed in highly skilled jobs (hypothesis j)
Actual data in blue and the projection's results in black.

B. University graduates who are over-educated (in percentages)

Year	HSJ/W		
	S1	S2	S3
2014	28.8	28.8	28.8
2015	30.3	30.3	30.3
2016	29.8	29.8	29.8
2017	30.1	30.1	30.1
2018	32.4	32.4	32.4
2019	33.3	33.1	34.0
2020	34.3	33.9	35.5
2021	35.3	34.7	36.9
2022	36.3	35.6	38.2
2023	37.3	36.6	39.5
2024	38.4	37.5	40.8
2025	39.5	38.6	41.9
2026	40.6	39.6	43.1
2027	41.7	40.7	44.2
2028	42.9	41.8	45.2
2029	44.0	43.0	46.2
2030	45.2	44.1	47.2

Note: Si refers to Scenario i. Actual data in blue and the projection's results in black.
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Table A4. 10. Forecast's results for 25-64-year-old university graduate women living in the East area

A. University graduates who are over-educated (in thousands)

Year	Pop.	Stage1	Stage 2		Stage 3		
		U1	W1	W2	HSJ1	HSJ2	HSJ3
2014	3,809	588	461	461	313	313	313
2015	3,787	576	465	465	299	299	299
2016	3,786	626	509	509	340	340	340
2017	3,789	676	556	556	368	368	368
2018	3,815	760	627	627	397	397	397
2019	3,849	809	670	687	414	425	427
2020	3,867	854	709	745	427	450	455
2021	3,884	899	749	805	440	475	485
2022	3,899	943	789	867	451	499	514
2023	3,911	987	828	931	461	521	544
2024	3,923	1,032	869	998	470	543	574
2025	3,935	1,076	911	1,067	478	563	605
2026	3,948	1,122	953	1,139	485	583	638
2027	3,959	1,167	996	1,212	490	601	670
2028	3,971	1,212	1,040	1,289	495	618	704
2029	3,977	1,256	1,082	1,365	497	632	737
2030	3,985	1,300	1,125	1,445	498	645	771

Note: Pop. is population, U1 is the number of university graduates (hypothesis 1); Wi is the number of university graduates employed (hypothesis i); HSJj is the number of university graduates employed in highly skilled jobs (hypothesis j)
Actual data in blue and the projection's results in black.

B. University graduates who are over-educated (in percentages)

Year	HSJ/W		
	S1	S2	S3
2014	32.2	32.2	32.2
2015	35.6	35.6	35.6
2016	33.2	33.2	33.2
2017	33.8	33.8	33.8
2018	36.7	36.7	36.7
2019	38.2	38.1	37.8
2020	39.7	39.5	38.8
2021	41.2	41.0	39.8
2022	42.8	42.5	40.8
2023	44.4	44.1	41.6
2024	46.0	45.6	42.5
2025	47.6	47.2	43.3
2026	49.2	48.8	44.0
2027	50.8	50.4	44.7
2028	52.4	52.1	45.4
2029	54.1	53.7	46.0
2030	55.7	55.4	46.7

Note: Si refers to Scenario i. Actual data in blue and the projection's results in black.
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Table A4. 11. Forecast's results for 25-64-year-old university graduate men living in the South area

A. University graduates who are over-educated (in thousands)

Year	Pop.	Stage1	Stage 2		Stage 3		
		U1	W1	W2	HSJ1	HSJ2	HSJ3
2014	2,785	273	217	217	163	163	163
2015	2,784	289	229	229	166	166	166
2016	2,780	315	251	251	176	176	176
2017	2,777	338	285	285	190	190	190
2018	2,776	331	273	273	190	190	190
2019	2,803	342	287	286	199	198	201
2020	2,801	351	298	295	206	204	209
2021	2,798	359	309	305	212	209	217
2022	2,790	366	320	314	218	214	226
2023	2,781	373	330	323	224	219	234
2024	2,772	380	341	332	230	225	242
2025	2,762	387	352	341	236	230	251
2026	2,752	394	362	350	243	235	259
2027	2,741	401	373	359	249	240	268
2028	2,728	407	384	368	254	244	276
2029	2,712	413	394	376	260	249	285
2030	2,700	419	404	385	266	254	293

Note: Pop. is population, U1 is the number of university graduates (hypothesis 1); Wi is the number of university graduates employed (hypothesis i); HSJj is the number of university graduates employed in highly skilled jobs (hypothesis j)
Actual data in blue and the projection's results in black.

B. University graduates who are over-educated (in percentages)

Year	HSJ/W		
	S1	S2	S3
2014	24.8	24.8	24.8
2015	27.4	27.4	27.4
2016	29.6	29.6	29.6
2017	33.3	33.3	33.3
2018	30.2	30.2	30.2
2019	30.7	30.7	29.8
2020	31.0	31.0	29.2
2021	31.4	31.4	28.7
2022	31.8	31.7	28.1
2023	32.1	32.0	27.6
2024	32.4	32.3	27.1
2025	32.7	32.7	26.5
2026	33.1	33.0	26.0
2027	33.4	33.3	25.4
2028	33.7	33.6	25.0
2029	34.0	33.9	24.4
2030	34.3	34.2	23.8

Note: Si refers to Scenario i. Actual data in blue and the projection's results in black.
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Table A4. 12. Forecast's results for 25-64-year-old university graduate women living in the South area

A. University graduates who are over-educated (in thousands)

Year	Pop.	Stage1	Stage 2		Stage 3		
		U1	W1	W2	HSJ1	HSJ2	HSJ3
2014	2,756	313	223	223	149	149	149
2015	2,759	330	240	240	162	162	162
2016	2,760	353	260	260	172	172	172
2017	2,760	378	280	280	190	190	190
2018	2,766	397	292	292	187	187	187
2019	2,786	421	307	311	192	195	196
2020	2,789	442	319	328	195	201	204
2021	2,790	462	330	345	199	207	212
2022	2,787	481	341	362	201	212	219
2023	2,781	500	351	378	203	217	226
2024	2,775	518	362	394	204	221	233
2025	2,767	536	371	411	205	225	240
2026	2,761	554	381	427	206	229	247
2027	2,753	571	391	444	207	233	254
2028	2,742	587	399	460	207	236	260
2029	2,728	602	407	475	206	238	266
2030	2,716	617	415	490	206	241	272

Note: Pop. is population, U1 is the number of university graduates (hypothesis 1); Wi is the number of university graduates employed (hypothesis i); HSJj is the number of university graduates employed in highly skilled jobs (hypothesis j)
Actual data in blue and the projection's results in black.

B. University graduates who are over-educated (in percentages)

Year	HSJ/W		
	S1	S2	S3
2014	33.3	33.3	33.3
2015	32.2	32.2	32.2
2016	33.7	33.7	33.7
2017	32.0	32.0	32.0
2018	36.1	36.1	36.1
2019	37.4	37.5	37.0
2020	38.7	38.8	37.8
2021	39.9	40.1	38.6
2022	41.1	41.4	39.4
2023	42.4	42.7	40.1
2024	43.5	43.9	40.8
2025	44.7	45.1	41.5
2026	45.9	46.3	42.2
2027	47.0	47.5	42.8
2028	48.2	48.7	43.4
2029	49.3	49.8	44.0
2030	50.4	51.0	44.6

Note: Si refers to Scenario i. Actual data in blue and the projection's results in black.
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Table A4. 13. Forecast's results for 25-64-year-old university graduate men living in the Canary Islands

A. University graduates who are over-educated (in thousands)

Year	Pop.	Stage1	Stage 2		Stage 3		
		U1	W1	W2	HSJ1	HSJ2	HSJ3
2014	641	57	46	46	31	31	31
2015	644	53	43	43	24	24	24
2016	640	76	53	53	30	30	30
2017	644	73	59	59	33	33	33
2018	654	73	59	59	32	32	32
2019	673	75	62	62	33	33	35
2020	680	74	62	63	31	31	35
2021	686	74	62	64	29	29	35
2022	692	73	63	64	27	28	36
2023	697	73	63	65	25	26	36
2024	702	72	63	66	23	24	36
2025	707	71	63	66	21	22	36
2026	710	70	63	67	19	20	37
2027	712	69	63	67	17	18	37
2028	714	68	63	67	15	15	37
2029	714	67	62	67	12	13	37
2030	715	65	62	67	10	11	36

Note: Pop. is population, U1 is the number of university graduates (hypothesis 1); Wi is the number of university graduates employed (hypothesis i); HSJj is the number of university graduates employed in highly skilled jobs (hypothesis j)
Actual data in blue and the projection's results in black.

B. University graduates who are over-educated (in percentages)

Year	HSJ/W		
	S1	S2	S3
2014	32.7	32.7	32.7
2015	43.7	43.7	43.7
2016	43.4	43.4	43.4
2017	43.8	43.8	43.8
2018	45.6	45.6	45.6
2019	47.1	47.2	44.0
2020	50.3	50.5	44.2
2021	53.6	53.8	44.4
2022	56.9	57.2	44.5
2023	60.2	60.5	44.7
2024	63.4	63.8	44.8
2025	66.7	67.1	44.9
2026	70.0	70.4	45.0
2027	73.4	73.7	45.1
2028	76.7	77.0	45.2
2029	80.0	80.4	45.3
2030	83.4	83.7	45.4

Note: Si refers to Scenario i. Actual data in blue and the projection's results in black.
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)

Table A4. 14. Forecast's results for 25-64-year-old university graduate women living in the Canary Islands

A. University graduates who are over-educated (in thousands)

Year	Pop.	Stage1	Stage 2		Stage 3		
		U1	W1	W2	HSJ1	HSJ2	HSJ3
2014	634	70	45	45	29	29	29
2015	638	76	54	54	32	32	32
2016	637	85	60	60	38	38	38
2017	641	85	61	61	38	38	38
2018	652	88	67	67	37	37	37
2019	669	92	73	72	38	37	40
2020	678	95	78	75	37	36	43
2021	687	99	83	78	37	35	45
2022	695	102	88	82	36	33	48
2023	703	105	93	85	35	32	51
2024	710	108	99	88	33	30	53
2025	716	111	104	92	32	28	56
2026	721	114	109	95	29	26	59
2027	725	116	114	98	27	23	62
2028	728	119	120	101	24	21	65
2029	730	121	125	104	21	18	67
2030	732	123	130	107	18	16	70

Note: Pop. is population, U1 is the number of university graduates (hypothesis 1); Wi is the number of university graduates employed (hypothesis i); HSJj is the number of university graduates employed in highly skilled jobs (hypothesis j)
Actual data in blue and the projection's results in black.

B. University graduates who are over-educated (in percentages)

Year	HSJ/W		
	S1	S2	S3
2014	35.6	35.6	35.6
2015	39.8	39.8	39.8
2016	37.1	37.1	37.1
2017	37.7	37.7	37.7
2018	44.7	44.7	44.7
2019	48.4	48.3	43.8
2020	52.1	51.9	42.9
2021	55.7	55.5	42.1
2022	59.3	59.1	41.2
2023	62.8	62.5	40.4
2024	66.2	66.0	39.5
2025	69.6	69.4	38.7
2026	73.0	72.8	37.8
2027	76.3	76.1	37.0
2028	79.6	79.4	36.1
2029	82.9	82.7	35.3
2030	86.2	86.0	34.5

Note: Si refers to Scenario i. Actual data in blue and the projection's results in black.
Source: SLFS (2nd quarters for data from 2014 to 2018), University statistics, and Population's projections from INE for population from 2019 top 2030)