Over-education, On-the-Job Search and Job Polarisation in Cyprus

By:

Christiana Charalambidou

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Social Sciences
Department of Economics

July 2018

Abstract

The present thesis is an empirical investigation into three labour market phenomena, namely over-education, on-the-job search and job polarisation in Cyprus.

Chapter 2 uses longitudinal panel data from the EU-SILC for the period 2005-2011 and employs a multitude of both static as well as dynamic probit models to examine the micro and macro determinants, persistence and dynamics of over-education. The main novelty in terms of the determinants of over-education is the inclusion of macro level independent variables to control for both aggregate supply and aggregate demand labour market conditions. These are found to be strongly significant and to have the expected sign. This chapter also disentangles the effect of past over-education experience on the likelihood of current over-education using a Wooldridge (2005) dynamic probit model with Mundlak (1978) corrections. Results demonstrate that over-education is not only a long-run phenomenon for many workers but also that current over-education is largely due to past circumstances of the individual with this state dependence present in all career stages.

Chapter 3 uses pooled cross sectional data from the EU-LFS for the period 2000-2015 to examine the determinants of on-the-job search and to shed light on its relationship with over-education. An econometric complication arises due to the possibility that unobserved heterogeneity could be driving both over-education and on-the job search. In order to overcome this potential endogeneity issue, an Instrumental Variables (IV) approach is implemented, using one of the macro level determinants found to significantly affect the likelihood of over-education in Chapter 2, as an IV for over-education. Results show that there is a strong positive relationship between over-education and on-the-job search both in the Probit and Ordinary Least Squares as well as in the IV regressions. This analysis is also replicated for the UK and Germany with results pointing to the fact that Cyprus behaves more like the more flexible UK labour market rather than the stricter German labour market.

Chapter 4 looks into the phenomenon of job polarisation using EU-LFS data for the period 1999-2014. Jobs are defined as specific occupations within sectors, a methodology called the jobs approach, and are ranked both according to their modal education level as well as by their average wage. The net employment changes are then plotted over time to observe trends in job change. Results demonstrate that job polarisation has taken place in Cyprus but only when jobs are ranked according to wages. Following this finding, the

raw proportions in each job level by age and year, as well as broken down by education are presented so as to observe how the workforce has changed its shares across the various job groups over time. Lastly, in order to examine job mobility of workers displaced from mid-level jobs as a result of routinisation, pseudo cohorts based on age and education are constructed and followed over four distinct periods of time. IV regressions at the cohort level are then ran with results providing evidence of job mobility from mid-level towards low-level and to a lesser extent towards high-level jobs while no evidence of movements out of the labour market is found.

Acknowledgments

I would like to express my greatest gratitude to my first supervisor, Professor Steven McIntosh for his invaluable guidance and support throughout my PhD. I am also beyond grateful for all the knowledge he has passed on to me and for his kind assistance and patience at each and every step of this long process. I would also like to thank my second supervisor, Dr Pamela Lenton.

I would also like to give my greatest heartfelt thanks to my husband for being by my side since the beginning and for all the sacrifices he has made along the way. His understanding and support were invaluable at each step of the way.

I would also like to thank my amazing son, Aris, who came into my life during my Phd and has been part of this adventure on a daily basis. With his laughter, affection and unconditional love he has made this long process more enjoyable and less stressful, helping me see the light at the end of the tunnel.

Last, but by no means least, I would like to thank my parents for supporting me in all my decisions in every possible way they could and the rest of my family for being by my side and for always loving and understanding me.

Notes and Disclaimers

Chapter 2 uses the Cyprus data files from the European Union Statistics on Income and Living Conditions (EU-SILC). I gratefully acknowledge use of the EU-SILC data under Eurostat Contract EU-SILC/ 2012/69. The results and conclusions in this chapter are mine and not those of Eurostat, the European Commission or any of the national authorities whose data are used.

For Chapters 3 and 4 I gratefully acknowledge use of the European Labour Force survey under Eurostat Contract 26/2014-LFS. Again, the results and conclusions in these two chapters are mine and not those of Eurostat, the European Commission or any of the national authorities whose data are used.

Contents

Appendices

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Motivation and Contributions

During the past decades a great number of industrialised countries have experienced growing levels of educational attainment as a result of increasing investments in their educational systems by governments. Such investments are expected to foster economic development (Kiersztyn, 2013) if the surge in the supply of more educated workers can be absorbed into the labour market in jobs commensurate with their education. If the labour market fails to do so, then a share of these workers end up in over-education, a situation whereby an individual's level of education exceeds the educational requirements for one's job. If over-education is found to be a temporary phenomenon in workers' lives with mismatched jobs acting as stepping stones to matched ones, as a number of labour market theories postulate, then over-education is self-corrected without the need for policy interventions. On the other hand, if over-education is found to be a large scale long-run phenomenon, as is often found in the empirical literature (e.g. Flisli 2017; Cedefop 2010), with over-educated jobs acting as a dead-end, then a real negative cost exists not only at the individual but also at the society level as they could both be seen to be over-investing in education (Borghans & de Grip, 2000). Policy interventions would then be required to correct it. Therefore, one of the main possible implications of over-education, is its negative effect on overall economic productivity as there is a failure to meet the full potential of the population based on their overall education level (Buchel, 2000).

Given the importance placed by governments on educational policies, i.e. promoting higher levels of education with the expectation of economic development and growth and the consequent budgetary investments in education, as well as the individual investments in education in terms of foregone immediate income and actual expenses related to studying, over-education presents itself as an inefficient outcome. This, together with the finding in the empirical literature that over-education is associated with a number of other negative outcomes at the individual, firm and macroeconomic levels, motivate the thorough investigation of the phenomenon in Chapter 2 of the present thesis.

An important issue in the study of over-education is its dynamic properties, an area that has received markedly less attention in the literature, mainly because such analyses require panel data sets. More specifically, on top of studying the permanent vs temporary nature of this phenomenon, it is important to examine whether over-education entraps people in mismatched jobs, increasing their likelihood of remaining over-educated in subsequent periods, i.e. over-education state dependence. State dependence is formally defined as “the degree to which the effect of any initial endowments (i.e. individual heterogeneity) on an outcome may be attenuated or accentuated by the continued presence of that outcome” (Lancaster 1979; Heckman 1981; 1991 in Mavromaras& McGuiness 2012, 620).

A finding that over-education is self-perpetuating can have important policy implications as it can translate into a scarring effect[1], for over-educated individuals who could develop a long-term labour market disadvantage (Mavromaras et al., 2012). Moreover, the risk of over-education having ‘a dampening effect on the growth potential of the economy’ (Mavromaras et al. 2012, 10) exists if over-education is found to be a genuine labour market imbalance that is also state- dependent. These reasons motivate the examination of over-education dynamics in Chapter 2 rather than constraining the analysis of over-education within a static framework. More specifically, given the panel nature of the EU-SILC dataset, utilised in Chapter 2, and the dynamic econometric modelling employed, it is possible to isolate the causal effect of previous over-education status on present over-education. This is a clear contribution to the limited number of such studies in the over-education literature, especially in the case of Cyprus where the dynamic properties of over-education have not been examined in the past.

The final contribution of Chapter 2 is to examine a set of macroeconomic level determinants of over-education in a dynamic setting. Macro level determinants of over-education have received considerably less attention in the literature of the determinants of over-education and there are calls in the literature for further research. This motivates working towards closing this gap in Chapter 2 in an attempt to bridge the micro and macro level causes of over-education. Apart from the unemployment level at the start of paid employment and the current unemployment level which can been occasionally found in

[1] Mavromaras et al. (2012) explain that: ‘a scarring effect is presented as a disadvantage that is self-perpetuating for the individual and is clearly over and above any positive or negative effect that their individual characteristics may play regarding the presence or absence of this disadvantage’ (Mavromaras et al 2012, 39).

earlier over-education studies, the rest of the supply and demand macro variables chosen in Chapter 2 to reflect overall labour market conditions, have not been examined in the past. Taking macro conditions into consideration also allows accounting for the fact that any observated immobility out of over-education could be a result of the general economic climate.

One of the costs of over-education confirmed in the literature is the wage penalty incurred by over-educated workers with the returns to surplus education falling short of the returns to required education (e.g. Duncan & Hoffman 1981; Hartog 1985; Hersch 1991; Alba-Ramirez 1993). What is more, another strand of the literature examining a number of dimensions of employee attidutes finds over-education to be positively associated with a number of worker behaviours that result in lower productivity. One of these bahaviours is a higher propensity of over-educated workers to engage in on-the-job search compared to their well-matched counterparts (e.g. Wolbers 2003; Wald 2005; Di Pietro and Urwin 2006) and consequently higher turnover. This means that over-education could have real behavioural consequences for workers. The phenomenon of on-the-job search with a special interest in its relation to over-education is the theme of Chapter 3, the second empirical investigation of this thesis. More specifically, the purpose of Chapter 3 is to empirically examine job search behaviour of employed workers and to contribute to the literature on the consequences of over-education by empirically linking it to on-the-job search.

As mentioned earlier, over-educated individuals may view a job for which they are over-educated as a stepping stone to a matched job, by for example gaining work experience, or a temporary income while acquiring additional information on labour market opportunities and adjusting their current position via on-the-job search (Hartog, 2000). Job separations have been argued to be a result of inadequate matches (McGuinness and Wooden, 2009) and on-the-job search can act as a correction mechanism leading to an exit from over-education. In other words, if the outcome of on-the-job search is successful, this leads to voluntary turnover which in turn is a mechanism through which job match imperfections can be restored, leading to a more efficient allocation of human resources (Ponzo, 2012).

In the US, Fallick and Fleischman (2001) find that, among college workers, job-to-job transitions account for 50% of separations while for the UK, Pissarides (1994) shows that job-to-job flows accounted for at least 40% of all separations in the 1980s. Similarly, Bachman (2005) shows that during the period between 1980 and 2000, job-to-job flows in Germany accounted for about 35% of monthly separations. Pissarides and Wadsworth (1994), claim that on-the-job search is an important study area given that job-to-job changes are responsible for the greatest part of labour turnover and note that job search is an indispensable process via which both the pecuniary and non-pecuniary quality of a job match can be improved. It follows that job-search plays a vital role towards an efficient labour market and hence it is important to examine its determinants. DeLoach and Kurt (2018), point out that even if theoretical models of on-the-job search date back to Burdett (1978), the lack of high-quality data related to search activities of employed workers has resulted in little empirical evidence on the phenomenon. The first contribution of Chapter 3 is to empirically analyse the determinants of on-the-job search for employees in Cyprus using a direct measure of on-the-job search from the EU-LFS.

Despite the fact that on-the-job search can act as a correction mechanism for labour market mismatches, it can also be viewed as a withdrawal behaviour from the worker's side, indicating lower commitment to the firm and therefore a productivity level below the full potential of the employee. It follows that on-the-job search could potentially become very expensive not only for the over-educated worker but also for the firm and the economy as a whole, even if it does not lead to turnover. Understanding the relationship between over-education and the propensity to engage in on-the job search is a key topic with implications of vital significance both for firms as well as for policy makers at a country level. This makes the study of the phenomenon and its relationship with over-education imperative, hence motivating the second contribution of Chapter 3 which is to examine the effect of over-education on firm productivity by investigating its relation with on-the-job search (Buchel, 2000).

Apart from the surge of higher education graduates during the past decades, another important labour market phenomenon in developed countries has been the so called polarisation or hollowing-out of the jobs distribution. This phenomenon is characterised by a simultaneous growth in the shares of total employment and labour demand for jobs at the high and low ends of the job spectrum relative to middle-ranked jobs over time. The routinisation hypothesis of Autor, Levy, and Murnane (AML) (2003) has been the most successful and widely accepted explanation of job polarisation in the literature and stresses the role of technology as the main driving force in determining the tasks that people perform in their jobs (Manning, 2004). More specifically, the Task Biased Technological Change (TBTC) theory postulates that as jobs involving more routine tasks are the ones most likely to be performed by machines following the IT revolution, the employment share in the middle of the jobs distribution, which is where such tasks are typically found, is expected to shrink. On the other hand, tasks within jobs at the low and high ends of the jobs distribution are not easily computerised and are complementary to technology, hence leading to an increase in the employment share of such jobs causing a U-shaped job distribution.

The examination of the quantitative evolution of jobs allows drawing a picture of how the jobs distribution and hence labour market demand and supply is evolving and is therefore important for properly anticipating future skill needs, an essential prerequisite for sound economic policy (Kampelmann and Rycx, 2011). Similarly, via the investigation of the process of job change, the changing job opportunity set faced by workers at different ages and education levels can be explored (Autor and Dorn, 2009). It is hence imperative to observe how the jobs distribution evolved over the years and this provides the motivation for studying job polarisation in Chapter 4.

Job polarisation has been extensively documented in the literature over the years with empirical work showing how it has been taking place in the US (Autor and Dorn 2009; Autor et al. 2006, 2008; Smith 2008), the UK (Goos and Manning, 2007), West-Germany (Spitz-Oener 2006; Dustmann et al. 2009) and EU countries (Goos and Manning 2009; Eurostat 2013). The handful of EU-level studies that include Cyprus in their analyses, place it in the group of countries whose job distribution has polarised when jobs are ranked according to wages while it has upgraded when jobs are ranked according to education (e.g. Eurofound, 2013). However, no in-depth country analysis for Cyprus seems to exist. To this end, the first contribution of Chapter 4 is to identify net employment shifts and to analyse the quantitative evolution of jobs in Cyprus using pooled cross sectional data from the EU-LFS.

More specifically, Chapter 4 will plot changes in the jobs distribution so as to observe patterns in job change and to empirically document any potential job polarisation. Jobs in this chapter are defined following the jobs approach which defines jobs as occupations within sectors and provides timely data in relation to levels of employment and job quality in both expanding and shrinking sectors and occupations (Eurofound, 2013). Jobs are then ranked according to two proxies of job quality, namely education and wages. If evidence is found that the middle of the jobs distribution has contracted compared to the extremes, then a central question would be where do displaced workers previously working in mid-level jobs go (Dorn, 2009). Smith (2013) notes that even though the polarisation phenomenon has been well documented in the empirical literature, questions linked to its potential implications such as whether displaced middle-skill workers stay in mid-level jobs, or somehow move to the higher- or lower-level labour markets have not been answered (Smith, 2013). Shedding light on the topic of job mobility due to polarisation is imperative so as to examine the labour market position and outcomes of this group of workers who are the ones directly affected by a possible hollowing-out. The second contribution of Chapter 4 will therefore be to examine job mobility of displaced mid-level workers as a result of the routinisation of job tasks. Conclusions from such analyses can then inform policies so as to safeguard the position of these workers in the labour market.

As a whole, the present thesis contributes to education and labour economics by addressing three important phenomena using data sets and methodologies that are original for the specific topics, in an attempt to fill gaps in the existing literature and provide the first case of country-specific evidence for the chosen country. All three chapters utilise individual level data sets, namely the longitudinal EU-SILC and the cross-sectional EU-LFS for the island of Cyprus. Chapter 3 also employs British and German data from the EU-LFS to serve as a comparison to the Cyprus results. A variety of methodologies are employed throughout the thesis, including the Wooldridge (2005) dynamic probit model with Mundlak corrections and both micro and macro level regressors, pooled binary probit models as well as random effects probit models and Instrumental Variable (IV) regressions as well as pseudo panel IV regressions at the cohort level to deal with the econometric and evaluation issues encountered in each chapter.

The current thesis uses Cyprus as its empirical case study for the reasons mentioned in the next section of this introduction and in doing so contributes to the over-education, on-the-job search and job polarisation literatures but most importantly to the limited country-specific evidence of the studied phenomena.

1.2 The Context

1.2.1 Cyprus Industrial Structure[2]

Cyprus is a small open, free market economy with a strategic location offering a business gateway between Europe, Asia, the Middle East and Africa. It is a member of the EU and the Eurozone and has a highly educated, English speaking population, excellent information and communications technology (ICT) infrastructure and a business-friendly environment.

The island's economy has undertaken a number of transformations over the years, from being an exporter of minerals and agricultural products in 1961-73 to being an exporter of manufactured goods from the late 1970s to the early 1980s, to being an international tourist, business and services hub since the 1980s. The island's economy is nowadays mainly built upon the service sector which accounts for over 80% of both total GDP and employment. The services it specialises in include tourism, financial and business services such as company formation, tax planning, trusts and foreign exchange trading as well as real estate. Industry (mainly manufacturing) and agriculture account for the remainder of the country's GDP. In terms of exports, Cyprus' main domestic export goods are pharmaceutical products, raw and manufactured food products, and scrap products.

In 2013, the Cyprus economy faced a huge setback due to a severe crisis of its banking system and was forced to request a financial assistance package from the European Commission (EC), the European Central Bank (ECB) and the International Monetary Fund (IMF) resulting in a €10-billion bailout deal and the controversial and unprecedented Eurogroup decision to enforce a depositor bail-in, an event that echoed around the world. The country exited the programme in 2016 and has been on the road of quick economic recovery ever since.

[2] This sub section is based on information available on the following website: https://www.cyprusprofile.com/en/economy

Even though Cyprus is small, its strategic location, EU membership and the fact that it trades with a lot of countries both in goods but mostly in services, means that events that happen in Cyprus have a potential impact on other countries and on the EU as a whole.

The Educational System in Cyprus

The educational system in Cyprus is provided by the government for free and is compulsory up to the age of fifteen. Compulsory education consists of pre-primary (5-6 years old), primary (6-12 years old) and lower secondary general education (Gymnasio) between the ages of 12 to 15. Lykeio (upper secondary education) offers a second three-year cycle between the ages of 15-18. Alternatively, in place of attending Lykeio, students can choose to attend secondary technical and vocational education. In addition to the above-mentioned public schools, private secondary education schools (as well as a few private primary schools), often subsidised by the government, also exist where students attend 6 or 7 years of uninterrupted secondary education depending on the school (6 years in the case of primary schools). In terms of tertiary education, three public and six private universities operate in Cyprus at present. Tertiary education in public universities is tuition free and an undergraduate degree is usually completed in four years. Examinations at the end of the final upper secondary grade provide a route into public universities for those students that managed to collect the highest scores for their chosen field of study. On the other hand, private universities accept students with their own criteria and they charge fees. It has to be noted that these private institutions only gained university status in 2007 as they were previously considered as colleges.

A number of Public and Private Higher Education Institutions also exist, none of which have university status. Such (public) institutions offer a number of vocational programmes of study with a duration ranging from one to three academic years as well as academic postgraduate programmes of study in Business and Public Administration. In terms of Private Institutions of Higher Education registered with the Ministry of Education and Culture these offer both academic and vocational programmes of study at the undergraduate and postgraduate levels and their programmes of study are evaluated and accredited by The Cyprus Agency of Quality Assurance and Accreditation in Higher Education[3].

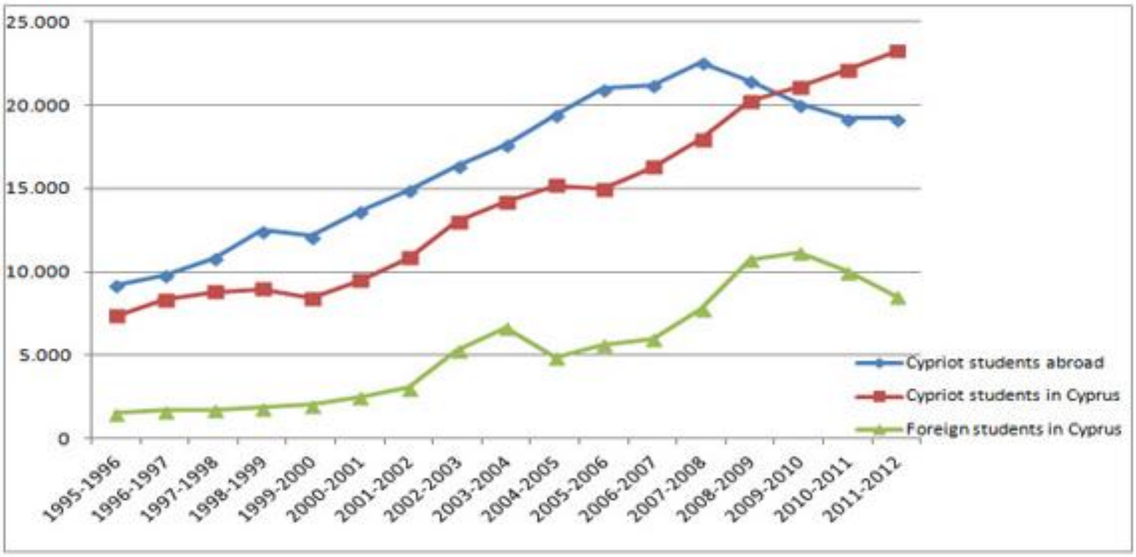
1.2.2 Higher (and Tertiary) Education and the Labour Market in Cyprus

Cyprus is a country characterised by a remarkably strong demand for higher education and like other small economies relies on its human capital as a key factor in production. This means that the quality of its human resources is of crucial importance for economic growth (Bacchus, 2008). More specifically, the percentage of secondary school graduates who chose to pursue further studies during the years of analysis fluctuated between 76-82%[4]. According to the Cyprus department of Higher and Tertiary education[5], one of their main goals over the past years has been to increase the number of people attending higher education in Cyprus, enhance the governance and funding of the Higher Education institutions as well as promote the knowledge triangle (education, research and innovation) and excellence in research, technology and innovation in Higher Education. To this end and with the aim of expanding Higher Education especially at University level, a number of reforms that aim to promote higher education have taken place in Cyprus during the past decades. Examples include the establishment and operation of the University of Cyprus in 1992 and another two state universities as well as private universities and the approval in 2005 of the law which regulates the establishment and operation of private universities in order to further upgrade the private tertiary education.

As a result of these reforms and increased expenditure on education, the number of students in Cyprus has increased rapidly over the last two decades. The following graph shows the number of higher education Cypriot and foreign students in Cyprus from 1995/96 up to 2011/2012, also illustrating the number of Cypriot students studying abroad. As is evident from the below figure, there has been a very rapid increase in the number of students.

[3] Information in this section is mostly based on information from the website of the Cyprus ministry of education and culture. It can be accessed at the following link: http://www.highereducation.ac.cy/en/educational-system.html
[4] Statistics of Education (2010/11) Report, Statistical Service of Cyprus (2011) : Summary table IX: Percentage of graduates of upper secondary level who pursue further studies in Cyprus and abroad, 1991/92-2010/11 %
[5] http://www.highereducation.ac.cy/en/

Figure 1.1: Cypriot and Foreign Students in Cyprus and Cypriot Students Studying Abroad



Source: http://www.highereducation.ac.cy/en/historical-background.html

Figure 1.1 also demonstrates another important trend that is the large number of Cypriot students pursuing their studies abroad, mostly in other EU countries especially Greece and the UK[6]. Given the accession of Cyprus to the EU in 2004 and the subsequent reduction in university fees, there has been a steep increase in students choosing to pursue their studies in countries such as the UK where higher education is not free and hence Cypriot students are now entitled to pay home-student instead of international-student fees. For example, in 2003/04 the number of Cypriot students in the UK amounted to 2806 while it has reached 8420 students in 2010/11. The reforms of the Cyprus higher education sector with the attainment of university status of previously private colleges, also played an important role in the surge in the numbers of students choosing to pursue higher education studies in Cyprus and this is likely to be the reason for the change in the pattern of the Cyprus vs abroad curves in Figure 1.1.

Table 1, below, demonstrates the percentage of people in possession of tertiary qualification across EU countries in order to observe the relative position of Cyprus.

[6]According to the Statistics of Education (2009-2010) report disseminated by the Cyprus Statistical Service, 51.2% of Cypriot students were pursuing higher education studies in Greece and 39.8% in the UK, 1.8% in the U.S. and smaller percentages in other EU countries.

Table 1.1: Percentages of People in Possession of Tertiary Education across Europe

Country	Percentage of people between the ages 25-64 with tertiary education qualifications, 2011	Percentage of people between the ages 30-34 with tertiary education qualifications, 2011	Percentage of students studying in another EU-25, EEA or Candidate country, 2010
Cyprus	37.4	45.8	54.9
Austria	19.3	23.8	4.3
Belgium	34.6	42.6	2.7
Bulgaria	23.4	27.3	8.1
France	29.8	43.4	2.5
Germany	27.6	30.7	3.9
Denmark	33.7	41.2	2.5
Switzerland	35.3	44.0	...
Greece	25.4	28.9	5.4
Estonia	36.8	40.3	5.6
United Kingdom	37.0	45.8	0.7
Ireland	37.7	49.4	13.0
Iceland	33.9	44.6	17.6
Spain	31.6	40.6	1.1
Italy	14.9	20.3	2.4
Latvia	27.7	35.7	4.6
Lithuania	34.0	45.4	5.0
Luxembourg	37.0	48.2	74.6
Malta	15.3	21.1	16.8
Norway	37.6	48.8	10.3
Netherlands	32.1	41.1	5.9
Hungary	21.1	28.1	5.1
Poland	23.7	36.9	5.1
Portugal	17.3	26.1	5.8
Romania	14.9	20.4	4.2
Slovak Republic	18.8	23.4	4.1
Slovenia	25.1	37.9	5.7
Sweden	35.2	47.5	3.6
Czech	18.2	23.8	2.9
Finland	39.3	46.0	2.9

Source: Statistics of Education 2010/2011[7] (Statistical Service of Cyprus).

As can be seen from the above percentages, Cyprus ranks fourth in terms of the highest percentage of people between the ages of 25-64 who have completed tertiary education

[7] Statistics of Education Report (2010/2011); Summary table XVI: Comparison of education indicators between European countries, pp 61-64

behind Scandinavian countries such as Finland and Norway, and Ireland, and seventh for the 30-34 age group. It is also the second country in terms of the percentage of people who pursue their studies abroad.

According to the website of the department of Higher and Tertiary Education of the Cyprus Ministry of Education and Culture[8], globalisation and the rapid technological change which has resulted in the restructuring of economies in the majority of countries around the world make higher education vital for the personal, economic and social development of individuals, groups and states. According to the same source, higher education being directly connected to employment is the most significant education level and plays an important role for development and social coherence. It is furthermore noted that: “the state policy for the substantial increase of expenditure on research and the establishment of Cyprus as a regional educational and research centre, creates new prospects for the development of tertiary education and indicates that there are great possibilities for tertiary education to contribute to the economic development of the country” (http://www.highereducation.ac.cy/en/). The above sentences demonstrate the importance placed on higher education by the Cyprus government. In other words, via investing in education, the government expects an increase in the country’s productivity and economic development and if over-education could potentially impair such a prospect, it is important first to identify it and second to find policy solutions so as to correct it even if this means that it is more optimal to reduce educational spending.

From the student perspective, Menon et al. (2012) quoting earlier work of theirs, explain that the high demand for higher education in Cyprus has been associated with the desire of students to improve their employment prospects in the island's small labour market (Menon 1998; Menon 2008a and Menon 2008b). So there is a high level of investment in higher education by both the government and individuals. At the same time, the country’s small-sized economy, combined with the great number of university graduates limit employment and career opportunities for young graduates (Bacchus, 2008).

In terms of unemployment, and as demonstrated in Figure 1.2 (A) below, the unemployment rate has been following a downward trend from 2006 onwards up until 2009 when it started to increase. Lastly, in terms of the unemployment level by education,

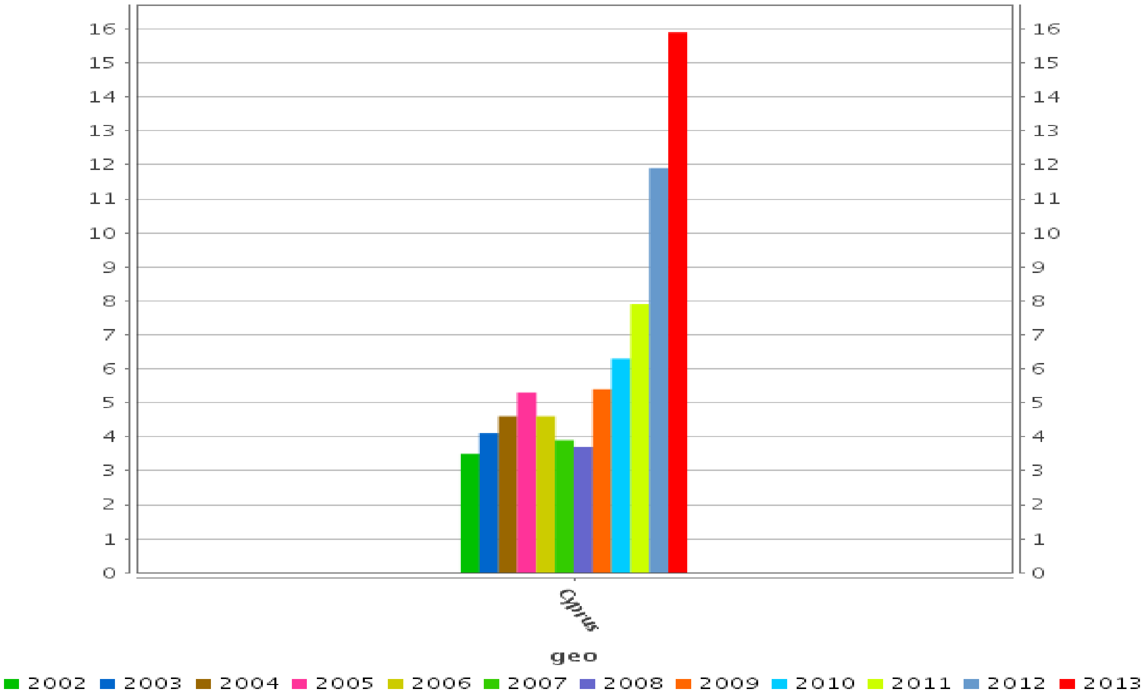
[8] http://www.highereducation.ac.cy/en/

the unemployment rate for tertiary level[9] graduates has also been increasing dramatically in recent years as Cyprus was negatively affected by the global economic crisis in 2009. As can be seen in Figure 1.2 (B), 2011 was the first year in which the unemployment rate for tertiary level graduates surpassed the EU(27) rate with the onset of the financial crisis in Cyprus and it has been increasing in the period illustrated and for some years after that, especially following the collapse of the Cypriot banking industry and its entry into the unenviable group of EU program countries, in March 2013 (Koutsampelas and Polycarpou, 2013). For upper secondary and post-secondary non-tertiary education (Figure 1.2 (C)), the trend has been similar to that for tertiary level graduates, however, for the time period under consideration, the unemployment rates for this group have been lower than the ones at EU-level. This is also the case for people with less than primary, primary and lower secondary education (Figure 1.2 (D)).

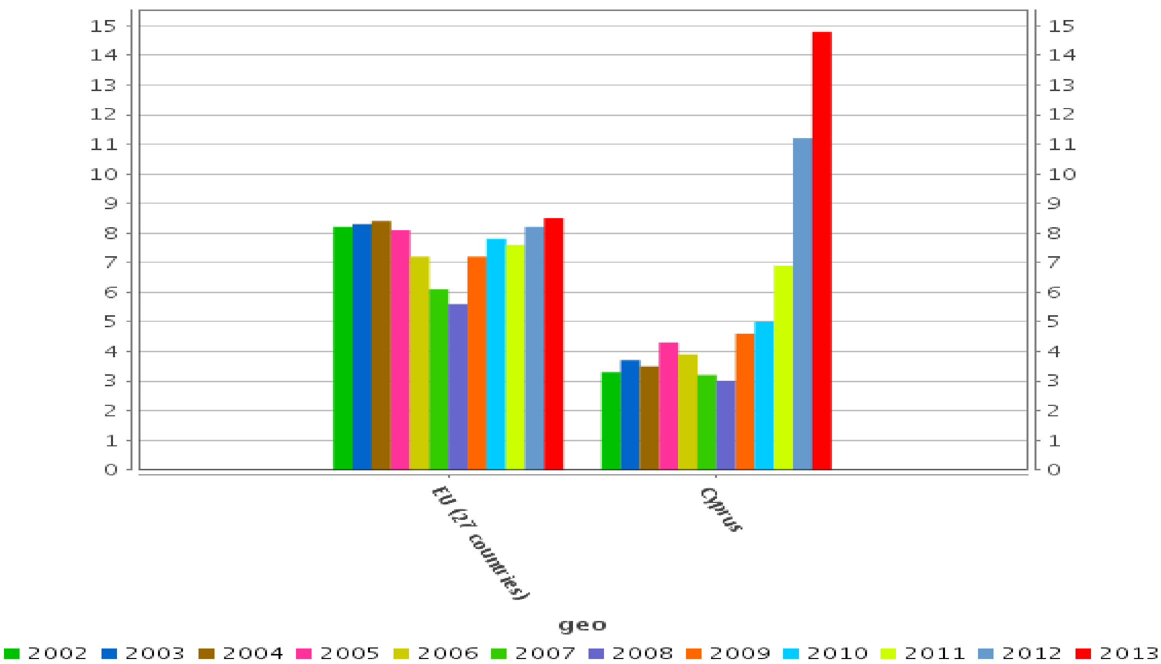
[9] Short-cycle tertiary, bachelor or equivalent, master or equivalent and doctoral or equivalent (levels 5-8)

Figure 1.2: Cyprus Unemployent Rates of the Population aged 25-64 (Annual Average)

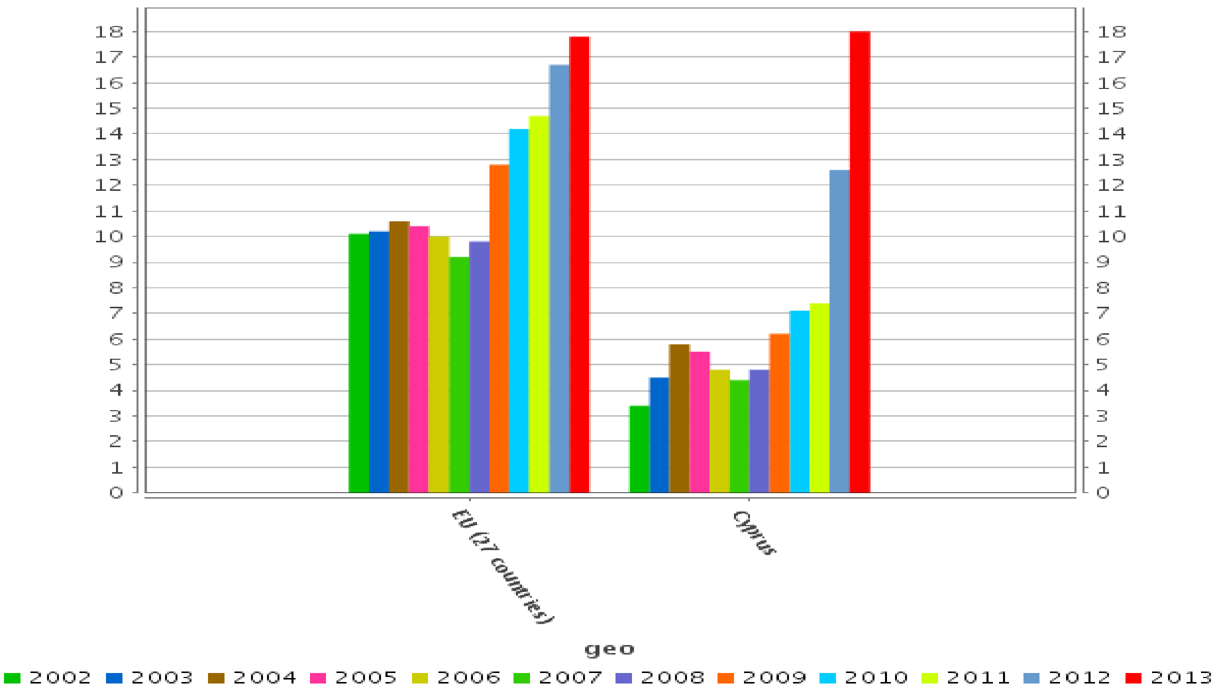
A- Population aged 25-64



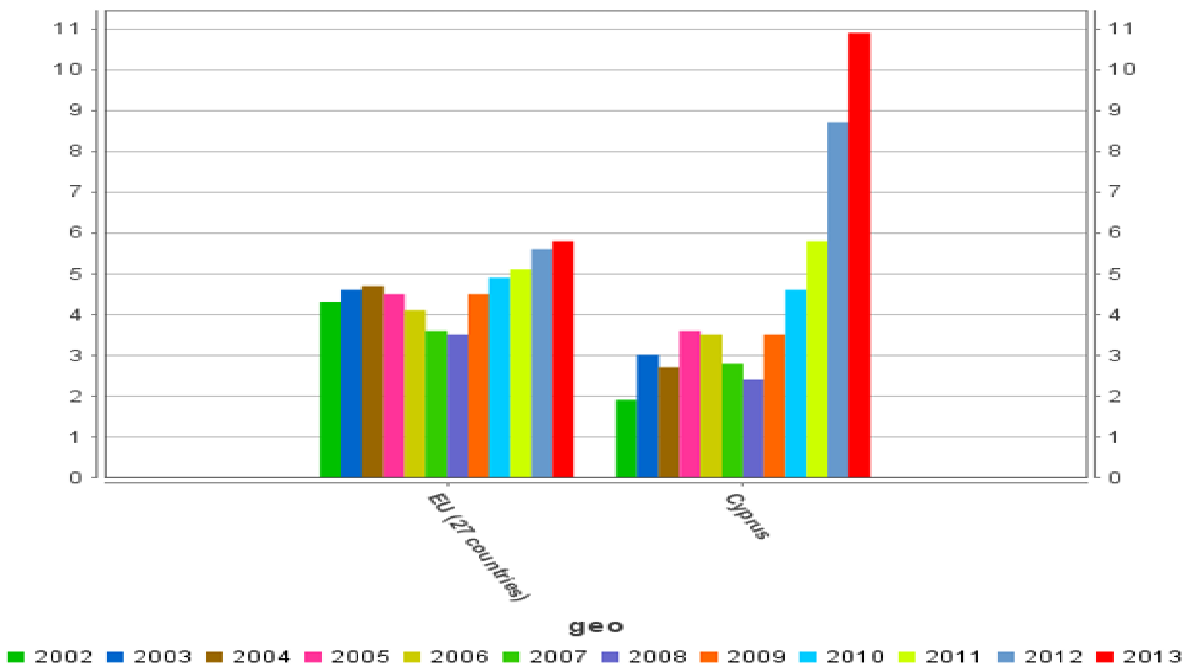
B- Upper secondary and enseignement post-secondary non-tertiary education (levels 3 and 4)



C-Less than primary, primary and lower secondary (levels 0-2)



D- Short-cycle tertiary, bachelor or equivalent, master or equivalent and doctoral or equivalent (levels 5-8)



Source: Eurostat (http://ec.europa.eu/eurostat) Last Update 30.10.2014; date of extraction 31.10.14 [10]

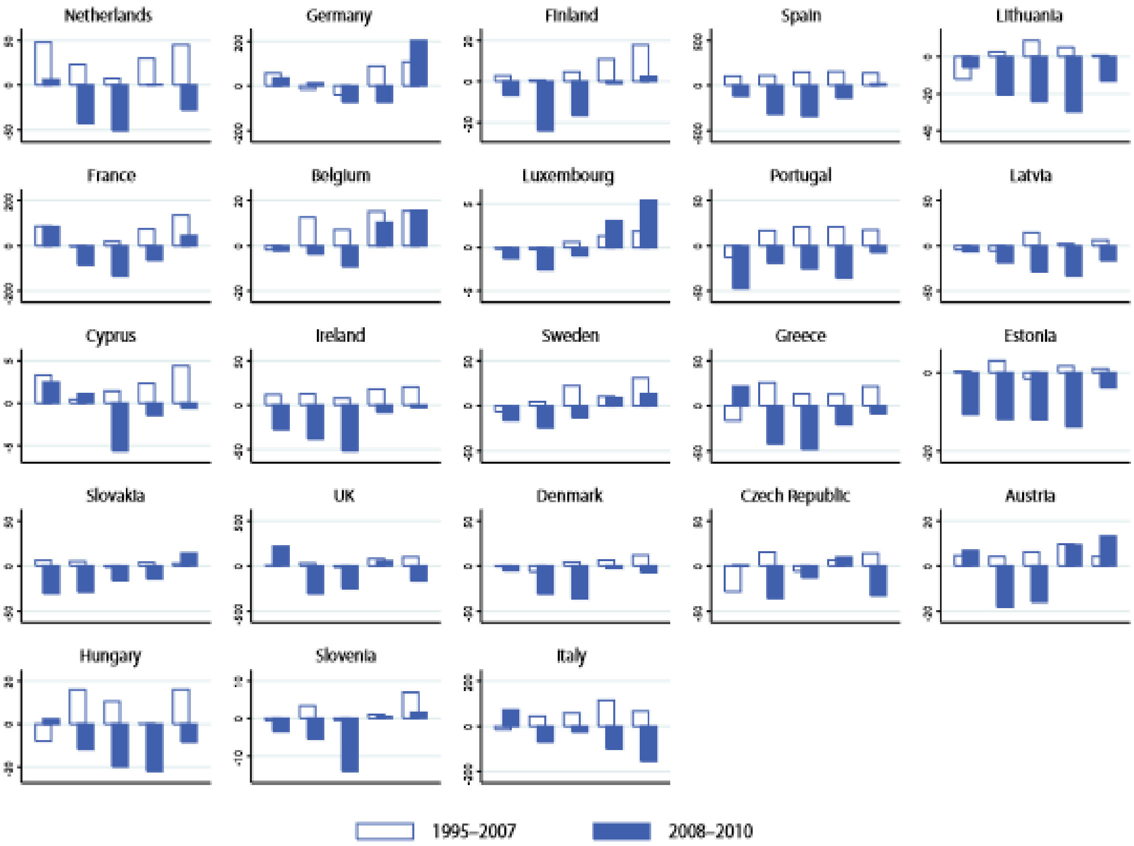
[10] Disclaimer for Figure 1.2: These graphs have been created automatically by Eurostat software according to external user specifications for which Eurostat is not responsible

Combined, the above factors make Cyprus an interesting case study for analysing the phenomenon of occupational mismatch as a failure to correctly match workers to jobs could potentially be very costly for the over-educated individuals, firms and the economy as a whole. Moreover, in the context of the recent recession that has considerably impaired the employment prospects of younger graduates, the risk of a strong persistence of over-education and hence of enduring economic costs, increases as it becomes increasingly difficult to find a good match by escaping over-education.

In terms of trends in changes in the jobs distribution, Figures 1.3 and 1.4 below, demonstrate the annual average change in employment, first when jobs are ranked according to wage and then when they are ranked according to education quantiles as they appear in Eurofound (2013), a comparative EU-level study of job polarisation. More specifically, the two figures below show how the structure of the job distribution has changed over the years when jobs are ranked according to two different proxies of job quality so as to get an overall picture of structural employment change across Europe and to observe whether a polarisation pattern has been also identified for Cyprus.

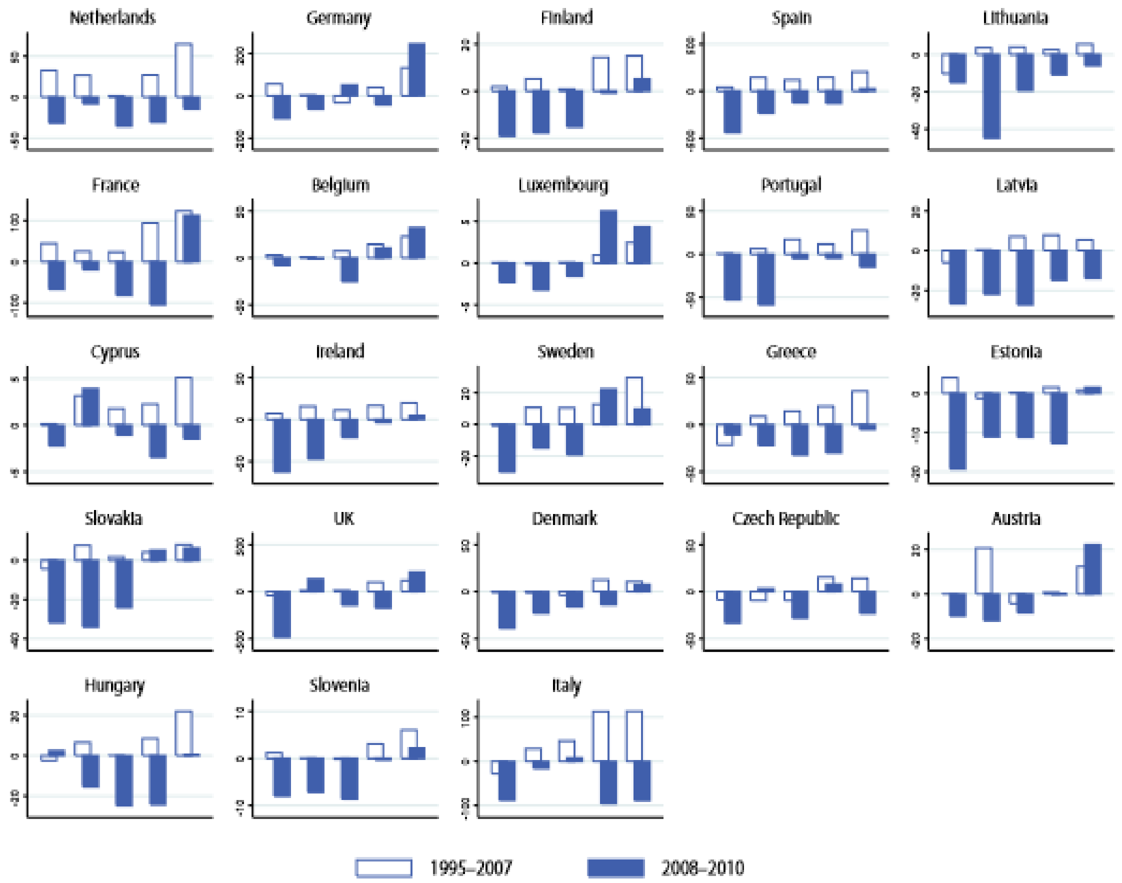
In the two figures below, the white bars demonstrate structural employment change during the expansion period (1995-2007 for most other countries and 1999-2007 for Cyprus) while the blue bars show structural employment change during the subsequent recession (2008-2010).

Figure 1.3: Annual Average Change in Absolute Employment by Wage Quintile and Country, 1995-2010 (Thousands)



Source: Eurofound (2013); Figure 18, pp 30

Figure 1.4: Average Annual Change in Employment according to Educaton Quintiles, by Country, 1995-2010 (Thousands)



Source: Eurofound (2013); Figure 34, pp 58

In the above-mentioned study, which is the only one available (to my knowledge) for Cyprus, Cyprus appears to be within the group of counties whose job distribution hollowed-out when jobs are ranked according to wages while it has upgraded (i.e. more high-level jobs and fewer mid and low-level jobs)when ranked according to education. This provides evidence that jobs in the middle of the jobs distribution when these are ranked according to wages have probably lost their employment share making it imperative to examine the phenomenon further focusing in-depth on Cyprus. Most importantly, the possibility of a hollowing-out of the jobs distribution calls for research in regards to its possible implications for the job mobility of those workers previously working in mid-level jobs who are displaced as a consequence.

In addition to the preceding discussion, Cyprus being a small open economy presents itself as an interesting case study for the investigation of the three phenomena outlined above while at the same time it is a country that is in shortage of empirical investigations relating to such issues

1.3 Structure and Content of Thesis

Each chapter in this thesis is a micro-econometric investigation of related issues in the spectrum of Education and Labour Economics. Brief summaries of each of these chapters are presented below.

1.3.1 Brief Overview of Chapter 2

Firstly, Chapter 2 provides an in-depth study into the phenomenon of over-education using panel data from the EU-SILC for the period 2005-2011 for people up to 40 years old in Cyprus. Apart from contributing to the literature on the determinants of over-education, it aims to disentangle the effect of past over-education experience on the likelihood of current over-education. More specifically, it employs a range of probit models including a Wooldridge dynamic probit model with Mundlak corrections, to estimate the effect of past over-education on the likelihood of current over-education and to investigate the micro and macro determinants of over-education. In order to reveal whether the effect of past over-education and hence state dependence varies according to the years of work experience possessed, separate regressions are also run on sub-samples based on the various career stages. Interaction terms between some of the macro variables and years of work experience are also incorporated in further analysis so as to test whether negative macro conditions have a greater effect on people with less work experience and subsequently whether work experience shelters workers against over-education during adverse macroeconomic situations.

1.3.2 Brief Overview of Chapter 3

Chapter 3 is an empirical investigation into the phenomenon of on-the-job search in Cyprus with a special interest in shedding light into its relationship with over-education. It uses pooled cross sectional data from the EU-LFS for the period 2000-2015 and a direct self-reported measure of search for another job while in employment. In order to eliminate concerns of a possible endogeneity of over-education and apart from binary probit and Ordinary Least Squares (OLS) regressions, an Instrumental Variable (IV) approach is also implemented. The methodology involves instrumenting over-education using one of the macro-level determinants found to have a significant impact on the probability of over-education in the previous chapter, namely the annualised change in the supply of labour by education category and sex. This chapter also runs IV regressions for the determinants of on-the-job search using one of the reasons for looking for another job, namely ‘because of the wish to have better working conditions (e.g. pay, working or travel time, quality of work)’ as a dependent variable to observe possible differences in the results compared to when no distinction as to the reason of on-the-job search is made. Lastly, the analysis is also replicated for the UK and Germany in order to compare results across three countries with different labour market characteristics and a different level of labour market flexibility.

1.3.3 Brief Overview of Chapter 4

The final chapter of this thesis, looks into the phenomenon of job polarisation in Cyprus using EU-LFS data between 1999 and 2014. Defining jobs as occupations within sectors according to the jobs approach, the net employment changes are plotted over time to observe trends in job change. Chapter 4 then moves on to examine job mobility of workers displaced from mid-level jobs. To do so, pseudo cohorts based on age and education level are constructed and followed over four distinct periods of time. In order to be able to make causal inferences as to the direct effect of a change in the proportion of people previously working in mid-level jobs on the cohort proportions in lower and higher level jobs, specifically due to routinisation as postulated by the Task Biased Technological Change theory, IV panel regressions at the cohort level are run. In this case, the IV methodology involves instrumenting the change in the proportion of a cohort in mid-level jobs with the proportion of each cohort in routine occupations in the previous period based on the occupation’s routine task score. It is expected that those cohorts with a larger proportion of people in occupations with a high routine score in the previous observation period will provide exogenous variation in the change in the proportion of a cohort in mid-level jobs from one period to the next. In this way, the job mobility of displaced workers, specifically due to routinisation, towards other job groups is revealed. The possibility that workers previously working in mid-level jobs are forced out of the labour market due to routinisation is also tested.

Chapter 2: Over-education in Cyprus: Micro and Macro Determinants, Persistence and State Dependence. A Dynamic Panel Analysis.

2.1 Introductory Remarks

2.1.1 Introduction

Over-education is commonly defined as the situation whereby an individual has a higher level of attained education than what is required by his/her job. Over-education, or vertical mismatch[11] as it is sometimes termed, can represent an inefficient allocation of resources as it implies an underutilisation of (educational) skills (Linsley, 2005) and can therefore be costly not only for the individual but also for the firm and the economy as a whole.

More specifically, at an individual level, apart from the wage losses that over-educated workers are likely to suffer in terms of diminished returns to their educational investments compared to matched individuals with comparable education (McGuiness, 2006), over-educated workers may also endure lower levels of job (e.g. Tsang et al. 1991; Battu et al.2000) and life (e.g. Piper, 2015) satisfaction, may experience a cognitive decline (De Grip et al., 2007) and are also found to have poorer mental and physical health (e.g. Tsang and Levin, 1985). Lastly, as over-educated workers move into lower level occupations, thus driving the mean educational level within these occupations upwards, some previously well-matched individuals could now be considered under-educated and could hence be 'bumped down' (e.g. see Battu and Sloane, 2000) or forced out of the labour market completely (McGuiness, 2006).

At the macro level, over-education may translate into lower national welfare and wasted tax revenues if individuals are equipped with non-productive education (McGuiness, 2006) while firms may suffer productivity losses, for example via lower employee satisfaction (Allen and van der Velden, 2001) and higher intention to quit[12] (e.g. Tsang

[11] The terms: over-education, vertical mismatch, occupational mismatch and educational mismatch are used interchangeably in this chapter.

[12] On-the-job search behaviour, a precursor to quitting is the theme of the third empirical investigation of the present thesis.

1987; Tsang et al. 1991) or via higher rates of turnover that translate into lost investments in training, screening and recruitment (e.g. Alba-Ramirez 1993; Sloane et al. 1999). Moreover, over-education has been consistently found to have a negative effect on earnings that could translate into a substantial productivity loss on a macro level, even if over-education was only a short-term phenomenon in workers' lives (Kucel et al., 2008).

On the other hand, another strand of the literature views the phenomenon as a statistical artefact resulting from either inadequate measurement techniques or from the absence of adequate controls within the ordinary wage equation context (McGuiness, 2006), or as temporary and not associated with high costs. If over-education is indeed found to be temporary and a path or a stepping stone towards a more productive job matching one's education, then the costs of over-education for all parties involved are expected to be minimal. Nevertheless, and even though some theories support the idea of a temporary or individual-level phenomenon, over-education has been found to be rather long-lasting at both micro and macro levels and to act as a stronger negative signal to employers than unemployment (McCormick, 1990). For example, in the UK, Lindley and McIntosh (2008) find that of those mismatched in 1991, 46% are still mismatched five years later and 34% ten years later, while Rubb (2003) finds that, in the US during the 1990s, 74% of those over-educated are still in the same situation one year later.

Correcting occupational mismatch is argued to be indispensable for improving a country's competitiveness and the welfare of individuals and as a consequence bringing about productivity gains and economic growth (Flisi et al., 2017). Moreover, in a harsh economic climate where public finances are constrained, understanding whether over-education results in labour market scarring is of utmost policy significance given the increasing number of people who pursue higher education studies and the large investments in educational spending incurred by governments.

2.1.2 Chapter Objectives

The purpose of the present chapter is manifold. Firstly, this chapter endeavours to contribute to the literature on the determinants of over-education, focusing on a more homogenous age group of people, i.e. people aged 16-40, in the Cypriot labour market, where the phenomenon has rarely been investigated. By restricting the focus of the chapter onto a specific age group of people that is nevertheless not too narrow, this chapter reduces the risk of capturing the cohort effect of over-education given the rise in the overall qualification level of the population. The micro level determinants of over-education examined are grouped into personal and job characteristics and work history variables.

An important issue in the analysis of over-education that has significant policy implications and has been receiving increased interest, especially due to the rise in the availability of panel datasets that allow such analyses, relates to the dynamic properties of over-education. In addition to investigating the permanent versus temporary nature of over-education, it is also important to examine whether previous over-education has a causal effect on future over-education, a situation known as state dependence. As demonstrated by recent literature, state dependence of over-education could be even more serious than simple over-education persistence whereby the state of over-education continues due to the continued presence of determining characteristics of the individual.

State dependence and heterogeneity being two different phenomena have to be explicitly modelled in estimation in order to provide a proper understanding of time-persistent effects (Mavromaras and McGuiness, 2012). Taking advantage of the panel nature of the EU-SILC dataset, and controlling for individual heterogeneity and initial over-education, the second objective of this chapter will therefore be to look into the dynamic aspects of over-education so as to examine the possibility of state-dependence, i.e. to examine whether lagged over-education is found to be significantly affecting the probability of the following year's over-education even after controlling for background factors that initially caused the over-education (Mavromaras et al., 2012). At the same time, by engaging in a dynamic analysis of the determinants of over-education, the chapter will isolate the main determinants of over-education within a methodological setting that copes with unobserved heterogeneity and state dependence (Boll et al., 2016). As an extension, and in order to examine whether state-dependence in over-education differs based on the career stage one is in, separate regressions will also be run on sub-samples of respondents at different stages of their careers. If evidence of state dependence is found, especially if this persists throughout one's career, then policies should be diverted to preventing entry into over-education in the first place and to discourage people from accepting mismatch jobs as a career strategy.

The third objective of the chapter is to go beyond the mere examination of micro level determinants and by incorporating a number of both demand and supply side macro level variables to attempt to link the micro and macro literature[13] on over-education and to examine how these labour market conditions affect the likelihood of over-education. The effect of overall macroeconomic conditions on the probability of over-education has attracted markedly less attention in the literature. More specifically, questions as to whether the likelihood of over-education increases, for example as result of negative labour market conditions that decrease the availability of jobs or increase labour market competition, have not been extensively targeted. This chapter undertakes to fill this gap and to widen the understanding of the effect of changing macro conditions on the likelihood of over-education by taking into account labour market conditions both at labour market entry and at the time of the survey. The macro variables examined here are the unemployment level at the start of paid employment as well as at the time of the survey, the annualised change in the labour supply by education category and sex and the annualised change in the employment share by occupation and sex.

2.1.3 Over-Education in Cyprus

Table 2.1 below demonstrates the relative position of Cyprus in terms of education mismatch against a number of EU countries as it appears in Flisi et al. (2017), based on information from the OECD Programme for the International Assessment of Adult Competencies (PIAAC). Information such as shown below is very scarce in the over-education literature as most comparative studies do not include Cyprus in their analyses.

[13] See Morano (2014) for a similar approach.

Table 2.1: Percentage of Individuals in Over-Education by Country

Country	Over-education (%)
Austria	13.4
Belgium	13.4
Cyprus	19.8
Czech Republic	6.8
Denmark	18.2
Estonia	17.9
Finland	10.1
France	9.9
Germany	12.8
Ireland	22.2
Italy	16.7
Netherlands	12.5
Poland	7.1
Slovak Republic	6.1
Spain	24.0
Sweden	10.5
United Kingdom	13.5
EU average	13.3

Source: Flisi et al 2017, 34.

As demonstrated by the above table, Cyprus has a 19.8% rate of over-education which is above the EU average of 13.3% and ranks third in terms of the highest over-education level behind Spain and Ireland. Furthermore, Flisi et al. (2017) find that Cyprus belongs to the group of countries (along with countries such as Italy, Spain and Ireland) in which educational as opposed to skill mismatch prevails, a type of mismatch they argue is the one generating the most negative effects (for example, lower productivity and psychological stress). This provides support for the choice of the conceptualisation of occupational mismatch in this chapter which is proxied by educational rather than by skill mismatch (as educational mismatch is more pervasive than skill mismatch). The authors find this result suggestive of an educational system[14] that fails to make appropriate human capital investments, either because it primarily provides general education or because the kind of education provided endows people with inadequate levels of skills that do not match those required by the labour market. In terms of over-education research specifically for Cyprus, one of the only related examinations is Kyrizi's (2011) PhD thesis which uses pooled cross-sectional Cyprus data from the EU-SILC and finds almost 50% of tertiary sector graduates as being over-educated.

All in all and even though over-education has received much attention in the economic literature to date, the evidence on Cyprus is almost non-existent. This chapter aims to fill this gap. To my knowledge, there is no empirical or other study on the issue of over-education using panel data for Cyprus. Moreover, there appears to be no other examination in the literature bridging the micro and macro determinants of over-education and I am not aware of any empirical evidence in regards to state dependence of over-education in general nor at different points in one's career in particular. This may in part be explained by the lack of panel datasets for Cyprus that have only recently become available but are still in great part limited in terms of the sphere of analysis they allow.

The remainder of the present chapter is structured as follows: Section 2.2 presents some of the theories that attempt to explain over-education and reviews the literature in the area, Section 2.3 describes the data and the derivation of the over-education variable and offers some descriptive statistics, Section 2.4 discusses the methodology, Section 2.5 outlines and discusses the results and Section 2.6 offers some concluding remarks.

[14] Further information on the educational system in Cyprus is available in Chapter 1.

2.2 Theories of Over-Education and Review of the Literature

2.2.1 Theories of Over-Education

Even though no single integrated theory of over-education exists, semi-formal economic models have provided a framework within which a number of authors have tried to conceptualise and explain the issue (e.g. Freeman 1976). These theories differ in terms of both the determinants as well as the private and social costs of over-education and the permanency of the over-education phenomenon (Linsley, 2005). In what follows, the main theoretical frameworks within which researchers have tried to explain the over-education phenomenon are briefly summarised.

Human Capital Theory (HCT)

According to Becker (1964), workers will always be paid the value of their marginal product which in turn will be commensurate to their level of human capital, obtained through either formal education or on-the-job training (Becker, 1964). The HCT postulates that, in order to make full use of their workforce skills, firms are prepared to adjust their production processes in response to fluctuations in the relative supply of labour. It follows that over-education, a state in which workers are under-utilised and wage rates paid are below the value of marginal product, can only be explained as a short-run disequilibrium condition that arises during the firms' production adjustment process that aims at fully utilising the individuals' human capital or until workers manage to find a more appropriate match via job search[15] (McGuinness, 2006). It follows that the HCT cannot be supported if over-education proves to be a permanent phenomenon that persists in the long-run.

Another explanation of over-education that is in line with the HCT stems from Mincer's (1974) earnings framework in which earnings regressions are based on years of schooling with less formal measures of human capital ignored. This gives rise to the possibility of an omitted-variables problem in the wage equation as less formal forms of human capital accumulation are not controlled for (Sala, 2011). In other words, over-educated workers could be compensating for their lack of work-related human capital (McGuinness 2006;

[15] The relation between over-education and on-the-job search is closely examined in the following chapter of this thesis.

Sala 2011). Similarly, it might be that there is an unobserved heterogeneity between over-educated workers and their adequately matched counterparts that results in lower productivity, something that is in turn reflected in the wages received. In other words, the lower earnings of over-educated workers compared to their matched colleagues could be reflective of their lower ability levels (McGuinness, 2006).

All in all, HCT can explain over-education if over-education proves to be a short-run transitory phenomenon and/or disappears when controls for work-based human capital investments and/or worker skill heterogeneity are included (McGuinness, 2006). The use of panel data in the present chapter allows controlling for unobserved heterogeneity and together with the inclusion of work experience as a control in the regressions will offer a direct test of the HCT. A potential finding that over-education continues into the long-run or that is state-dependent via the dynamic regressions will provide evidence against the HCT.

Career Mobility Theory (CMT)

Similar to the HCT, the CMT (Sicherman and Galor, 1990) considers education, experience and training as substitutes, all of which are positively related to productivity and earnings and views over-education as a temporary equilibrium labour market outcome. However, unlike the HCT, the CMT postulates that educated individuals entering the labour market for the first time choose positions for which they are apparently over-educated in order to gain on-the-job training and work experience and hence to improve their future labour market prospects. According to Sicherman and Galor (1990), workers may choose a job for which they are over-educated if the chances of promotion are higher and hence this model suggests that, in some occupations the returns to schooling come in the form of a higher probability of advancement to occupations with higher wages rather than directly in the form of higher wages (Sicherman and Galor, 1990). Similarly, according to Sicherman's (1991) career mobility hypothesis, workers temporarily enter jobs for which they are over-educated so as to acquire the work experience and training needed for progression to higher-level jobs. According to this theory, over-education is a typical characteristic of a well-functioning labour market that bares trivial economic costs, given that it corrects itself as those over-educated progress to higher level occupations in which they fully utilise their qualifications. Groot and

Maassen Van den Brink (1997) find evidence in support of the CMT and conclude that younger age cohorts have a higher incidence of over-education but this is dominated by the negative relation that over-education has with years of labour market experience. Support for the CMT is also provided by Sicherman (1991) who shows that, compared to correctly matched workers, those over-educated for their jobs have a higher probability of job change and are more likely to move up the occupation hierarchy.

In this chapter, the CMT is tested via the inclusion of work experience (and age) as a control in all the regressions as well as by the sub-regressions examining the probability of over-education and state dependence of over-education at different career stages. If over-education is found to affect only those at the early stages of their careers then the CMT can be used to explain over-education. On the other hand, a potential finding of persistent over-education across career stages and/or of over-education state-dependence will provide evidence incompatible with the CMT as it will be suggestive of the fact that people accepting over-educated jobs fail to progress to matched ones and hence over-education does not correct itself.

Therefore, the HCT and CMT perceive over-education as a temporary phenomenon and a normal feature of a well-functioning labour market and hence unproblematic from a policy perspective. On the other hand, the theories presented below suggest that over-education could also exist in the long run.

Job Competition Theory (JCT)

The ease with which firms can rapidly change their production methods to facilitate changing factor input prices, suggested by the HCT, is questioned by numerous economists who argue that this might not be possible due to the structure of work practices but also due to institutional arrangements like national pay agreements (Duncan and Hoffman 1981; Hartog and Oosterbeek 1988).

The JCT, based on Thurow (1975), provides a demand side explanation for the existence of over-education contrary to the supply side explanation of the Human Capital model. It suggests that the marginal product (and earnings) resides in the job rather than the individual's characteristics with workers investing in human capital for its signalling power rather than for its content (McGuinness, 2006). According to the JCT, people compete for job opportunities by acquiring more education, which they view as a way to signal their productivity to prospective employers who are bound by imperfect information. Given that the JCT stresses the importance of a person's relative position in the labour market, it follows that as the number of educated people in the economy increases, the level of competition for jobs also increases and as a result over-education arises as people acquire more and more education in their effort to get to the front of the job queue. In this chapter, a proxy for the level of labour market competition, i.e. the annualised change in the labour supply by educational level and sex, is entered into the regression to examine whether an increase in the number of individuals at the various education levels affects the likelihood of over-education. If evidence of a significant and positive relationship between this variable and over-education is found, then JCT could offer a possible framework of explaining over-education.

In many ways the Job Competition model is very similar to the signalling framework proposed by Spence (1973) where, as a response to imperfect information held by employers in relation to the productivity of workers, individuals invest in education to signal their quality (McGuinness 2006; Linsley 2005). Over-education takes place when there is a signalling equilibrium under which it is optimal for individuals to overinvest in education (Spence 1973), due to either low education investment costs or due to high expectations of individuals or firms about education (Linsley, 2005).

Assignment Theory (AT)

The AT formulated by Sattinger (1993), attempts to bring together the HC and JC theories by employing matching theory to unite both demand and supply side factors into the analysis of over-education. It accepts that skills obtained in education contribute positively to productivity but claims that skill utilisation is determined by the education-job match as job constraints may allow only a limited use of these skills. According to the AT, productivity and wages are also contingent on this match (Levels et al. 2014; McGuinness 2006). It follows that AT views over-education as a function of both the individual and the job with over-educated workers unable to fully utilise their education/skills as they are constrained by the nature of their jobs. Therefore, over-education is a form of allocative inefficiency in which skills are underutilised, that continues until a more efficient match is accomplished via improved matching methods or government policies to lessen inefficiencies (Linsley, 2005).

2.2.2 Review of the Empirical Literature

Richard Freeman's study of the US graduate labour market in 1976 (Freeman, 1976) was the first to draw attention to the phenomenon of over-education. Since then, the over-education literature has mainly focused on its adverse wage effects with the common finding being that over-educated workers receive a positive rate of return to over-education that is nevertheless lower than the rate they would have received had they been appropriately matched (e.g. Duncan & Hoffman 1981; Hartog 2000; Korpi & Tåhlin 2009). Even though this general finding in the literature on the wage effects of over-education has also been found to be consistent over time and across countries, according to McGuinness (2006), there is no agreement as to the main causes of over-education and its temporary vs permanent nature. In what follows, some of the empirical literature on over-education is reviewed[16].

2.2.2.1 Determinants of Over-Education at the Individual Level

The literature examining the determinants of over-education has generally found age to be negatively related to the probability of being over-educated either due to a better quality of job match as one's career advances (McGuiness and Wooden, 2009) or because an over-educated individual's excess education compensates for their lack of work experience (Sicherman and Galor, 1990). Alternatively, the negative relation between age and over-education could be the result of younger workers accepting over-educated jobs as part of a career strategy at the start of their career (Leuven and Oosterbeek, 2011). Similarly Groot (1993, 1996) and Sicherman (1991) find that over-educated workers have less experience, tenure and on-the-job training than correctly allocated workers (Groot

[16] Along with the review of the over-education literature, some of the related literature on over-skilling (e.g. Allen and van der Velden 2001; Green and McIntosh 2007;Mavromaras and McGuiness 2012; Mavromaras et al. 2012) is also cited where over-skilling is defined as a situation whereby an individual has more skills and knowledge (rather than just education) than those utilised in the job. Skills are usually measured via employee self-assessments as to whether their jobs require a lower skill level than the one they possess or whether they do not fully utilise their skills at their current job.

and van den Brink, 2003). Alba-Ramirez (1993) also finds that over-educated workers have less work experience.

Being a woman (Dolton and Silles, 2001) and belonging to a minority in terms of ethnic origin (Green et al. 2007; Battu and Sloane 2004) have also been associated with a higher probability of over-education. In terms of health limitation, Blazquez and Malo (2005), using Spanish data do not find a significant relationship between disability and educational mismatch, a surprising result, as they note, given that there are good reasons to anticipate the problem of over-education to be more severe for those with disabilities. On the other hand, Jones and Sloane (2010) find evidence that the disabled have a higher likelihood of both over and under-skilling than the non-disabled.

Being of a lower individual ability which is compensated by excess educational attainment has also been found to increase the likelihood of over-education (e.g. Green and McIntosh, 2007; Chevalier and Lindley, 2009 and Allen and Van der Velden, 2001).

More recently, Baert and Verhaest (2014) estimate the stigma effect of unemployment and over-education within one framework and examine the "scarring effects" of over-education due to negative signalling. They argue that, in the absence of information about worker quality, a candidate's labour market history is often used by employers as a signal of future productivity. It follows that past unemployment as well as past over-education could be serving as a negative signal to employers. Their results show that, unemployment has a more negative impact on subsequent employment than over-education and suggest that accepting jobs such as short-term and part-time jobs that are of a lower-level yet have low risks of habituation and reduced job search, should not be problematic.

2.2.2.2 Labour Market Conditions and their Effect on the Probability of Over-Education

In terms of the empirical literature examining the effects of macro labour market conditions on over-education, this has mostly focused on how labour market conditions such as unemployment at the time of graduation (labour market entry) and at present affect the probability of over-education and future career paths. For example, Kahn (2010) using the National Longitudinal Survey of Youth (NLSY79) analyses wages, labour supply, occupation, and educational attainment as a function of economic conditions in the year of college graduation and finds large, negative and persistent labour market consequences of graduating in a bad economy. She moreover finds that underemployment and job mismatch are more likely to affect those workers who graduate in times of economic hardship compared to those employed individuals who graduate in good economies, as the number of jobs to choose from are less.

Oreopoulos et al. (2012), using Canadian data, analyse the long-term effects on earnings, job mobility and employment characteristics of graduating in a recession and find that graduating college in a recession causes very persistent but not permanent earning losses while the effects of recession shocks are more severe for young workers, compared to workers with more work experience.

Liu et al. (2012) using panel data from Norway show that cyclical skill mismatch is a significant driver behind persistent career loss as a result of graduating in recessions and find a strong countercyclical pattern of skill mismatch among college graduates, with initial labour market conditions having a decreasing yet persistent effect on the likelihood of early career mismatch. More specifically, they find that a 1% increase in the unemployment rate increases the probability of over-education by 3.4 percentage points.

On the other hand, Dolton and Silles (2001) examine unemployment as a possible cause of over-education among college graduates testing whether it increases the probability of accepting a job for which one is over-educated, but do not find a statistically significant result.

Lastly, Morano (2014) looks at whether labour market conditions have an effect on the likelihood of over-education, hypothesising that as an increasing unemployment rate reduces the number of opportunities faced by a worker, it increases the probability of accepting a job below ones' level of attained education. They find that the rate of unemployment has a small, negative and usually not statistically significant impact on over-education yet when accounting for the effect of unemployment on over-education across age groups, they find that the coefficients of the interaction terms between unemployment and the 20-24 and 25-29 age groups are positive and statistically significant but they find the latter to be considerably smaller than the former. They conclude that, labour market conditions are more likely to affect new labour market entrants to a greater degree than those with longer job tenures.

2.2.2.3 Duration and Dynamics of Over-Education: Persistence and State Dependence

As mentioned above, over-education has been generally found to decrease with age and work experience, something that could be used as evidence to suggest that over-education is, for a lot of workers, a temporary phenomenon (Groot and van den Brink, 2003). Lindley and McIntosh (2008), for example, use UK panel data to examine the determinants of over-education and of the transitions out of over-education and find a negative relationship between over-education and job tenure and a significant movement out of over-education. They take this as evidence to suggest that over-education is a temporary phenomenon at the start of an individual's career. Nevertheless, they accept that over-education is a reasonably permanent state for a minority of individuals.

On the other hand, a number of studies that connect current to previous over-education have found evidence suggesting that over-education is a permanent phenomenon and a trap for many workers. For example, Dolton and Vignoles (2000) use a survey of UK graduates and find that a substantial percentage of graduates (30%) remain over-qualified six years after graduation while Lindley and McIntosh (2010) using British panel data (1991-2005), show that over-education in the first job can potentially leave a scarring effect on workers' wages later on in their careers. At an EU-level, Verhaest and Van der Velden (2010) find that between 30-58% of those over-educated in their initial employment, were still over-educated five years post-graduation whereas of those who were not over-educated in their first job only 3-6% were found to be over-educated five years later. Some of the relevant studies linking past and current over-education are summarised below.

Firstly, Dolton and Silles (2001) model over-education in first and present employment making use of both probit and bivariate probit estimation. They show that being over-educated for one's first job has a strong effect on the match quality later on and that, when compared with graduates who are in matched jobs at the start of their careers, graduates who are initially over-educated find it more difficult to get graduate jobs later on. They explain that this could be either because of the fact that over-education could cause a deskilling or an obsolescence of graduates' skills that are not used, or because a bad start is difficult to recover from.

Scherer (2004) in her cross country comparison of West Germany, the UK and Italy also finds that mismatch at labour market entry is related to less prosperous career chances and higher unemployment risks later on.

Over-education has also been shown to result in habituation in terms of decreasing negative association between over-education and job satisfaction (Verhaest and Omey, 2010) which can translate into greater chances of long-run or permanent over-education. Baert et al. (2013) using Flemish data also find that it takes longer for young graduates who accept a job below their level of qualifications to get a matched job than what it would have taken them had they continued to be unemployed. They argue that even if job search models suggest that it may be optimal to take on a lower level position and to engage in on-the-job search as suggested in Dolado et al. (2003), it is uncertain whether the same search intensity can be maintained and conclude that it does not seem to be the case that these jobs act as stepping stones to a matched job.

Moreover, Mavromaras et al. (2012) argue that over-skilling scarring is equivalent to unemployment and under-employment scarring stemming from lost product through human capital under-utilisation, i.e. using too few skills[17] and abilities and that the end result of both forms of under-utilisation is a lower product and lower pay. They use Australian panel data to estimate the lagged effect of persistence and find that there is a high degree of state-dependence in over-skilling, a finding that provides evidence that over-skilling is not a temporary phenomenon. Using a random effects dynamic probit model they furthermore estimate the effect of over-skilling dynamics on wages and find that over-skilling mismatch is common and that those who have been over-skilled in the past are more likely to be over-skilled at present. They also find that there is an inverse relation between over-skilling persistence and the level of education.

Similarly, Mavromaras and McGuiness (2012) use panel data from the Australian HILDA survey (2001-2006) to examine the possible presence of state dependence in over-skilling, and whether it differs by education pathway. They use a dynamic random effects probit model which includes Mundlak corrections and model the initial conditions following Heckman's method and find widespread over-skilling state dependence particularly for highly educated individuals while workers with vocational education show no state dependence. According to the authors, the fact that over-skilling is found to be state

[17] This can easily extend to over-education i.e. not using the level of education attained.

dependent has significant policy implications in that the cost of labour market mismatches for individuals is dependent not only on the size but also on the duration of the wage penalty.

Frenette (2004), using a representative survey of Canadian graduates also shows that over-qualification is highly state-dependent, with those who are over-qualified two years after graduation being significantly more likely to still be over-qualified five years later.

In a different context, Cuesta and Budría (2011) investigate the impact of personality traits on over-education dynamics using German panel data and employ first-order Markov models controlling for the endogeneity of initial over-education status (non-random selection of people into over-education the first time they are observed in the data) and non-random attrition. They show that structural state dependence in over-education is moderate and that irrespective of the considerable differences in over-education transitions associated with individual heterogeneity, a non-trivial state dependence in over-education exists. More specifically, they find that the probability of remaining over-educated is 89.2% and that 17.6% of the observed persistence in over-education is due to the fact of having been over-educated in the previous year.

The work of Cuesta and Budria (2011) is nevertheless criticised by Piper (2015) in his article in relation to over-education and happiness. Piper (2015) argues that given that the average age of the respondents in Cuesta and Budria (2011) is 41.5, it is unlikely that years of schooling will be changing for many of these individuals, something he argues means that what is really being measured is transitions into and out of occupational categories. Moreover, Piper (2015) argues that the remarkably large persistence rate of over-education found by Cuesta and Budria (2011) may have little to do with over-education as such and more to do with people changing jobs and entering different occupations. Piper (2015) goes on to note that, given that not many people appear to change jobs in Cuesta and Budria (2011) and given the increase in participation in higher education in Western Nations like Germany, it is likely that their over-education dummy is mostly capturing younger people. To overcome this limitation and in order not to capture the cohort change of increasing qualifications amongst the young, Piper (2015) focuses on the twenties age group, an arguably more homogenous group. The study investigates the relative over-education and life satisfaction using British longitudinal data and employs a dynamic panel analysis, to account for the existence of serial correlation, testing the hypothesis of a negative relationship between being over-educated and life satisfaction. Evidence is found that the relatively over-educated are relatively less happy hence confirming the negative relationship between over-education and life satisfaction.

Kiersztyn (2013) using Polish panel data between 1988 and 2008 examines changes in the persistence of over-education in the long-run paying close attention to the relationship between the rate of these changes and the general economic conditions. She finds a high over-education persistence with 50% of those over-educated remaining over-educated five years later and around one in ten respondents being persistently over-educated. She also finds that the incidence of over-education during her studied period increased quicker during the recession and that those aged between 26 and 35 in 2008 ran the greatest risk of persistent over-education compared to other cohorts. In order to disentangle state dependence, she employs a dynamic random effects logistic regression correcting initial conditions using the Wooldridge (2005) approach and controls for unobserved individual heterogeneity using Mundlak (1978) corrections and finds that workers who are over-educated in the previous period are around four times more likely to be over-educated at the next panel wave. She notes that this figure seems reasonably high, particularly given that five years should be enough to allow many workers to achieve a better job match and that her findings are in support of the job competition and assignment theories.

More recently, Boll et al. (2016) relying on German data for 1984-2011, employ probit models in order to estimate the likelihood of entering over-education as well as dynamic mixed multinomial logit models with random effects to address state dependence and unobserved heterogeneity. Their model estimates the covariates' effect on the probability of newly entering a specific type of over-education and find that over-education is mainly state dependent and that employment experience lowers the probability of entering over-education with persons in a later stage of their career having a lower likelihood of entering over-education than persons at the very beginning, a result they take as supporting search theory. Moreover, they find that having been unemployed in the past mostly increases the risk of entering subjective over-education as well as the transition to twofold over-education (according to the authors these are the individuals classed as being over-educated by both objective-realised matches- and subjective-self assessment-methods) [18] for both genders and regions and stress the importance of continuous employment. Lastly, they find that changing one's employer is not an appropriate exit strategy and being in possession of a dual qualification does not insure a graduate against job mismatch with overall effects mainly relying on the operationalisation of over-education (Boll et al., 2016).

2.3. Data and Descriptive Statistics

2.3.1 Data

2.3.1.1 The EU-SILC

The data used in this chapter come from the European Statistics on Income and Living Conditions (EU-SILC) that is coordinated and released by the statistical office of the European Union (Eurostat). The EU-SILC is a multi-dimensional instrument used to undertake analyses on poverty, inequality and deprivation. It focuses on income but also covers housing, material deprivation, labour, health, demography and education so as to allow studying the multidimensional approach of social exclusion. There are two components to the EU-SILC, a household questionnaire answered by the household reference person and an individual questionnaire for each household member over 16 years of age. The EU-SILC provides two types of annual data: cross-sectional annual data with variables on income, poverty, social exclusion and other living conditions, and longitudinal data pertaining to individual-level changes over time, observed over a four year period[19].

For the purposes of this chapter, only longitudinal data files for Cyprus were used. Cyprus joined the EU-SILC in 2005 and the survey is administered by the Statistical Service of Cyprus with a harmonised version then supplied to Eurostat. The full information set is obtained through surveys among households and interviews with household members. EU-SILC is a four-year rotational panel dataset meaning that participants can be followed

[18] A discussion of the various methods of measuring over-education found in the literature can be found in the next section of the thesis
[19]More information on the EU-SILC can be accessed at the following link:
http://epp.eurostat.ec.europa.eu/portal/page/portal/income_social_inclusion_living_conditions/introduction#

for a maximum of four consecutive years. The SILC is the only individual level longitudinal survey currently available for Cyprus.

The longitudinal component of the EU-SILC is to some extent limited in terms of variables that could be of interest and that are available in other cross sectional surveys explicitly focused on the labour market, such as for example the Labour Force Survey (LFS) but also in the cross sectional component of the EU-SILC. For example, variables of potential interest such as the field of study of respondents, sector of employment and information on the educational background and economic situation of family members that are available in surveys such as the LFS are absent from both the cross sectional and longitudinal files of the EU-SILC. Similarly, variables such as citizenship and the industry in which one works that could both be argued to affect the probability of over-education are only available in the cross sectional files of the EU-SILC which are not linkable to the panel ones.

Nevertheless, the fact that individuals are followed for four years offers an important advantage when compared to cross sectional data sets specifically designed for labour market analyses. More specifically, panel data are indispensable in order to examine transitions in and out of over-education and to test whether over-education is a temporary or permanent phenomenon. More importantly, longitudinal data allows the introduction of dynamic regression analysis enabling the control of individual heterogeneity and the examination of the possibility of state dependency of the dependent variable.

The CY-SILC sample is drawn from the 2001 Census of Population sampling frame and the sample design was one-stage stratification. The sampling units are private households which were selected with simple random sampling within each stratum. Geographical stratification criteria were used for the sample selection. The minimum effective sample size for Cyprus according to European Regulation is 3250 households for the cross sectional component and 7500 persons aged 16 or over, while for the longitudinal component it is 2500 households and 5500 individuals aged 16 or over.

In this chapter, respondents who have successfully completed a personal interview in all four years of their respective panel are kept in the data set so as to create balanced sub-panels with the maximum number of observations per respondent, i.e. to be able to observe each respondent for the maximum four-year duration of the survey. In summary, the four four-year panels: 2005-2008, 2006-2009, 2007-2010 and 2008-2011 available for Cyprus are used[20] and people who entered in the first year of each panel and who completed an individual interview in all four rotations of the respective panel are kept. Details of the data cleaning and preparation of the final sample are provided in Appendix 2A. Following the elimination of duplicate observations, the files of the various years[21] were then appended together into one unbalanced panel of seven years duration[22]. This means that the resulting dataset is a seven-year unbalanced panel with four sub-panels where each individual is observed for four years (i.e. either between 2005-2008, 2006-2009, 2007-2010 or 2008-2011).

As mentioned above, at the time of writing only seven years of data are available which is a relatively short time span. A longer panel duration both in terms of more years for the same individual but also more sub-panels would allow a more in depth examination of the state dependence and of the dynamics of over-education and would also allow an assessment of the longer-term effects of over-education.

Table 2.2 below, demonstrates the level of attrition between the first and the fourth year of data for each of the four original sub-panels (before being merged into one) by showing the mean value of the backgound characteristics of the individuals in the sample. The initial and final sample sizes of each panel are also demonstrated. As can be seen in the table, the proportions of the various background characteristics are relatively stable throughout the years of each panel and hence any concerns about non-representativeness of the final sample due to choosing to keep just the people who successfully reached the fourth round of their respective panel are not warranted[23].

[20] The first User Database (UDB) file available for Cyprus is the 2006 UDB that contains data collected in 2005 and 2006. Given that Cyprus entered the EU-SILC in 2005, the first complete 4-year panel for Cyprus is the 2008 UDB.
[21] I.e. the 2005-2008 panel (the 2008 longitudinal UDB file), the 2006-2009 panel (2009 UDB file), the 2007-2010 (2010 UDB file) and the 2008-2011 (2011 UDB file) with respondents who have entered in the first year of each panel.
[22] A number of variables of interest have changed categories (and names) from 2009 onwards. Details of how this was dealt with are provided in Appendix 2B.
[23] The attrition statistics presented here, could also be used as evidence against a potential argument that out migration could be acting as a safety valve for the over-educated as they show that the average characteristics of the people interviewed on the first and last round of the survey are very similar. Therefore, it does not appear to be the case that for example migration could be a safety valve for specific age groups or people with higher education or the over-educated even if relevant migration statistics are not available.

Table 2.2: Attrition Check-Descriptive statistics in the First and Fourth Round of Each Sub-Panel in the Original Dataset

Panel:	2005-2008		2006-2009		2007-2010		2008-2011	
	Round 1=2005 n=2259	Round 4=2008 n=1780	Round 1=2006 n=2258	Round 4= 2009 n=1753	Round 1= 2007 n=2195	Round 4=2010 n=1672	Round 1=2008 n=1974	Round 4= 2011 n=1488
Age	**44.58** (17.9)	**47.48** (17.6)	**45.1** (17.9)	**49.1** (17.9)	**45.09** (18.23)	**48.28** (18.07)	**46.40** (18.08)	**49.75** (17.84)
Education:								
Pre-Primary	**0.08** (0.27)	**0.09** (0.28)	**0.08** (0.27)	**0.09** (0.28)	**0.07** (0.26)	**0.08** (0.27)	**0.07** (0.25)	**0.07** (0.26)
Primary	**0.20** (0.40)	**0.22** (0.41)	**0.19** (0.39)	**0.20** (0.40)	**0.20** (0.40)	**0.21** (0.41)	**0.22** (0.41)	**0.23** (0.42)
Lower secondary	**0.12** (0.33)	**0.10** (0.27)	**0.11** (0.31)	**0.08** (0.27)	**0.12** (0.33)	**0.09** (0.28)	**0.12** (0.32)	**0.08** (0.26)
Upper Secondary	**0.38** (0.49)	**0.37** (0.48)	**0.38** (0.48)	**0.36** (0.48)	**0.38** (0.48)	**0.37** (0.48)	**0.35** (0.48)	**0.36** (0.48)
Post-secondary non-tertiary	**0.02** (0.13)	**0.02** (0.15)	**0.02** (0.15)	**0.02** (0.14)	**0.02** (0.12)	**0.01** (0.12)	**0.02** (0.13)	**0.01** (0.12)
1st and 2nd stage tertiary	**0.20** (0.40)	**0.22** (0.41)	**0.22** (0.42)	**0.25** (0.44)	**0.21** (0.41)	**0.24** (0.43)	**0.23** (0.42)	**0.26** (0.44)
Work Experience	**22.62** (15.16)	**23.54** (15.49)	**22.62** (15.25)	**24.42** (15.43)	**23.24** (15.58)	**24.53** (15.54)	**24.14** (15.31)	**25.4** (15.2)
Married	**0.66** (0.47)	**0.69** (0.46)	**0.65** (0.48)	**0.69** (0.46)	**0.65** (0.48)	**0.66** (0.47)	**0.65** (0.48)	**0.67** (0.47)
Female	**0.51** (0.5)	**0.52** (0.5)	**0.53** (0.5)	**0.52** (0.5)	**0.53** (0.5)	**0.53** (0.5)	**0.51** (0.5)	**0.52** (0.5)
Health limitation	**0.25** (0.43)	**0.21** (0.41)	**0.20** (0.40)	**0.23** (0.42)	**0.24** (0.42)	**0.23** (0.42)	**0.23** (0.42)	**0.28** (0.45)

Note: Standard deviation in parenthesis

2.3.1.2 Data Sources for Aggregate-Level Variables

For the derivation of the macro determinants of over-education, a number of additional datasets were used. Two of those aggregate statistics: the unemployment level by age and sex during each survey year (2005-2011) and the number of people employed by occupation and sex (from which the annualised change in the employment share by occupation and sex was then calculated), were taken from the Cyprus Labour Force Survey Reports prepared and released by the Statistical Service of Cyprus, based on the results of the EU-harmonised Labour Force Survey (years 2005-2011).

Similarly, the unemployment rates at the start of regular/paid employment by sex, were taken from the Cyprus Labour Statistics Report that is published annually by the Statistical Service of Cyprus and incorporates data on employment, unemployment, vacancies, placements, government labour force, foreign workers, social insurance statistics, wages, salaries etc[24].

The annualised change in labour supply by level of education and sex was calculated using information on the highest level of educational attainment for persons 20 years of age and over taken from the Cyprus Statistics of Education report (2004-2011) that presents the results of the Annual Survey on Education that is prepared by the Statistical Service of Cyprus[25].

2.3.2 Final Sample

As discussed above, this chapter uses Cyprus data from waves 2005-2011 of the EU-SILC survey. Given that the focus of the present chapter is people over-educated for their jobs, observations of respondents who are: (i) students, (ii) soldiers[26] (ii) retired, unemployed or inactive/disabled, and/or (iv) self-employed or family workers[27] for the full four years in which they are interviewed, are excluded from the sample. However, observations of individuals who worked for at least some of the panel years are not excluded. People working in the armed forces are also dropped before calculating over-education as the calculation of the norm (required) educational level in the wide range of professions encompassed by the term 'armed forces' is not straightforward. Lastly, as the focus of this chapter is youth, people over the age of 40 in their first round of the survey were also dropped. This means that the oldest respondent in the data is 44 years old in their last survey rotation. Following the calculation of the over-education variable, those

[24] The data on unemployment, used here are collected by the Cyprus Ministry of Labour and Social Insurance.
[25] All the reports mentioned in this section are accessible on the website of the Cyprus statistical service at: http://www.cystat.gov.cy
[26] Serving their 18 month compulsory military service
[27] The self-employed and family workers are excluded as their level of mismatch cannot be easily and reliably identified

with an attained education below secondary level (i.e. pre-primary and primary education) were also dropped from the sample[28].

Following the above data manipulations, the final sample size retained is approximately 1062-1490 observations per cohort with the total sample size being 5333 observations across 1617 respondents. For the panel regressions, anyone with fewer than two observations is dropped from the sample. This is 186 observations, bringing the sample size down to 5147 observations and 1431 individuals. The cross-sectional regressions are run on all available observations as well as on the reduced sample when only respondents with at least two years of data are kept. The results from the cross sectional regressions are almost identical in the two instances hence the change in the sample size is not driving the difference in results between the cross-sectional and panel regressions. For reasons of comparability, the cross sectional regression results shown in the following sections are the ones based on the reduced sample i.e. the one that is also used for the panel specifications.

2.3.3 Measurement of Over-education[29]

Due to the fact that empirical research depends on the availability of appropriate data, no uniform measurement of over-education exists (Verhaest and Omey, 2006). On the contrary, a number of both subjective and objective measures have been used in the over-education literature over the years (e.g. see Groot et al., 2000). More specifically, in subjective measures, individuals self assess the skills/education required to do their job whereas in objective measures job requirements are either inspected and set out, or an individual's education is set against that of a reference group (frequently constructed based on broad occupational categories) (Piper, 2015). Verhaest and Omey (2006) present two subjective and two objective methods. More specifically, the subjective methods are direct self-assessment whereby individuals are asked whether they are over-educated for

[28] Respondents with pre-primary and primary education were only used when estimating the modal education level by occupation and the over-education variable but were then dropped from the sample as over-education exclusively affects people with a higher educational attainment. However, results are robust to first dropping those with pre-primary and primary education and then calculating over-education and to keeping them in the sample. In the latter case, the overall over-education percentage in the sample stands at 15.3% instead of the 16% when they are dropped. Given that this is a very small difference (only 94 individuals or 342 observations have an education below secondary level), not including them in the sample does not exaggerate the extent of over-education, or the estimated results.

[29] Further details of the derivation of the over-education variable together with detailed descriptions of all the variables used in this chapter as well as the secondary variables used to construct some of the explanatory variables used in the regressions can be found in Appendix 2D.

their job and indirect self-assessment whereby the responses of individuals as to what is the education level required to do their job is compared to their actual education to deem whether they are over-educated. In terms of the objective measures, in Job Analysis (JA) the required level of education is created by job analysts for each occupational classification whereas in the Realised Matches (RM) approach the individual's education level is compared to the mean or modal education level of workers in each occupation. None of the methods of measuring over-education mentioned above are without limitation[30]. For example, the self assesment methods are likely to be problematic for quantitative analysis given the subjective nature of the worker's responses as to what education and/or skills are required for his/her job. In terms of the two objective methods, one of their main downsides is that the possible heterogeneity of jobs within occupations cannot be taken into account. This is discussed in detail further on.

The choice of over-education measure is likely to be important and Verhaest and Omey (2010) in their analysis of over-education measurement recommend measuring over-education in various ways so as to make reliable conclusions. Unfortunately, the EU-SILC does not contain a question about the level of education required in order to perform one's job or to be employed in the respondent's occupation and for this reason the self-assessment method of measuring over-education, cannot be used in this chapter. Moreover, it was not possible to find any reliable databases with detailed analysis of educational requirements by occupation for Cyprus, and hence Job Analysis could not be employed in this case either. Therefore, only one operationalisation of over-education is possible in this chapter. This does not allow any robustness checks of the sensitivity of the results to the measure of over-education chosen.

Over-education in this chapter is measured via the RM approach. More specifically, the modal educational level for each occupation group is assigned to the pooled set of observations irrespective of year[31] so as to serve as a proxy or an indication of the required level of education within each occupation[32]. Even though the choice of the mode rather

[30] See Groot et al. (2000) ; Rubb (2003); Verhaest and Omey (2006) for more details.
[31] Here, the assumption that occupations have a fixed level of required education is made. Given that the total time period in this chapter is only seven years, it is safe to assume that jobs have fixed requirements during this period as jobs are not expected to change that quickly, so pooling the data to calculate modal education is the right approach.
[32] Observations of respondents with a self-defined current economic status other than "Working Full time" or "Working Part time" were dropped before generating the educational requirement by occupation. Similarly, given that the educational requirements for people who are self-employed are likely to differ and their over-education status not being easily defined and hence not easily comparable to that of employees, observations of respondents with a status in employment other than employees (i.e. observations belonging to any of the following categories: Self-employed with employees; Self-employed without employees; Family Worker) were also dropped before calculating the modal

than the mean here is due to the structure of the education level information in the data (categorical rather than continuous), the mode is usually preferred, as a reference point. This is because it is less sensitive to outliers and technological change (Frisli et al., 2017) and does not entail the arbitrary use of one standard deviation above the mean when identifying who is over-educated (Mendes de Oliveira et al., 2000)[33]. In order to maintain a satisfactory number of observations within each occupational cell and to provide more precise modal values, the occupation variable, originally classified according to the 2 digit International Standard Classification of Occupations (ISCO-88), was reclassified resulting in 20 occupational groups[34] [35].

The over-education binary (dummy) variable is then equal to 1 if the highest level of education attained by the respondent at the time of the survey is above the modal level of education or the norm education level within their occupation (i.e. a respondent is over-educated) and equal to 0 if the highest level of education attained by the responded is less than or equal to the modal level of education within their occupational group (i.e. a respondent is not over-educated)[36].

In general, it has been argued that, when set against JA and worker self-assessment, the RM approach yields lower incidences of education-job mismatch than the other measures (McGuinness, 2006), thus being a more demanding criterion for assessing over-education (Piper,2015). On the other hand, being determined by actual qualifications held rather than required education, the RM approach as a whole has been criticised in the literature as potentially failing to capture the true incidence of over-education and instead capturing

---

educational level by occupation. Therefore, the modal level of education within each occupational group is derived solely based on the observations of people that are in full time or part time work and working as employees and aged up to 40 in the first instance they are surveyed. Observations belonging to respondents in any of the following categories are also dropped before calculating the modal level of education by occupation: Unemployed; Pupil, student, further training, unpaid work experience; In retirement or in early retirement or has given up business; Permanently disabled or unfit to work; In compulsory military community or service; Fulfilling domestic tasks and care responsibilities; Other inactive person.

[33] Lindley and McIntosh (2010) give an example to demonstrate this point whereby in an occupation where 75% of the workforce are graduates, the modal method, as desired, would classify anyone with a degree as adequately educated, irrespective of the qualification level of the remaining 25% of workers. On the other hand, using the mean to classify those who are adequately matched could mean that, if a proportion of workers are employed in the same occupation with lower-level education because they are in possession of longer work experience for example, this could drive the mean level of education within that occupation down to less than a degree, and depending on the size of the standard deviation, the possibility exists that all the graduates in this occupation could appear as being over-educated (pp 38).

[34] Details provided in Appendix 2C

[35] The use of more aggregate job titles for the derivation of the realised matches norm is not unusual in the literature and guarantees a satisfactory number of observations within each occupation. According to Verhaest and Omey (2010), although at the cost of more required heterogeneity within job titles, using more aggregate job titles mitigates biases that are related to the more detailed occupations but not to the aggregate (Verhaest and Omey, 2010).

[36] Given that the focus of this chapter is over-education, respondents with an educational level equal or below their occupation's modal level of education i.e. correctly matched for their job or under-educated, are classed as 'not over-educated'.

only the non-structural part of over-education within occupations (Verhaest and Omey 2005; Verhaest and Omey 2010). More specifically, even if job requirements are unchanged, an increase in the qualifications held for example due to a general rise in the qualification level of the population will cause the average qualifications of workers hired in all occupations to rise therefore increasing the measured required education (Lindley and McIntosh, 2010).

Similarly, Verhaest and Omey (2010) note that it is often the case in the literature that the derivation of the RM measure is based on data for the full labour force which means that the mean or modal level of education within occupations is not reflective of the mean or modal human capital level because of the additional experience or training of older workers (Kiker et al., 1997). In other words, due to the rise in participation in higher education and the fact that the average education level among younger people entering the labour market is higher than the existing work force, an over-education dummy that is estimated based on the whole population irrespective of age is likely to be reflecting 'the cohort change of increasing qualifications amongst the young' (Piper 2012, 15) or the qualifications of people who were hired at different times (Frisli et al., 2017). It is for this reason that simply comparing this overall educational norm by occupation with the individual level of education can then lead to misrepresentative inferences about the mismatched situation.

For the above reasons, and for the fact that the RM approach tends to be rather sensitive to cohort effects, it was decided that the calculation of the required education by occupation and hence the over-education variable was best to be based on the respondents who were aged up to 40 years old at their first survey year rather than on the whole sample regardless of age[37]. On the other hand, it could be argued that by excluding respondents older than 40 years of age, the norm educational level by occupation and therefore the over-education variable will only be representative of younger cohorts and hence will not reflect the true over-education level in the labour market. In other words, the true extent of over-education and hence the competition faced by the youth in the labour market may not be truly reflected as the mode and over-education variables will refer solely to the younger workers rather than everyone within the same occupation. Moreover, by dropping the older cohorts, the fact that employers, faced with an increased supply of

[37] Results are robust to small changes in this cut-off age point.

educated labour, might have upgraded jobs and employ highly educated workers in jobs that do not require such high qualifications might not be (easily) captured. All in all, the argument of avoiding to capture the cohort effect by calculating the norm educational level by education and hence basing over-education on only the younger cohorts seems more appropriate in this case.

As mentioned earlier, one of the main shortcomings of measuring over-education using objective measures, is that it does not take heterogeneity of jobs within occupation codes into account (Sloane et al., 1999). More specifically, a person who works in one occupation category could be doing a different job or have different roles than another worker in a different firm/sector whose job falls under the same occupational category. It follows that, one of the two employees could be in possession of a higher level of education than what is required for the specific occupation category (based on the modal education of other employees in that occupational category for example), while the other worker whose role involves more complex tasks and has an equivalent education level as the other employee is adequately educated for that job. The RM measure would neverthelees class both workers as over-educated. Similarly, certain occupations accept entrants with a wider range of education (Piper, 2015) oferring different career progression opportunities however, the available datasets, as is the case of the EU-SILC, utilised in this chapter,do not allow controlling for such occupations.

In terms of the dynamics of over-education, the above weakness of the RM measure means that all the action (over-education transitions) is restricted to between occupation changes, meaning that this measure is not going to pick up changes within occupation codes. In other words, the present chapter can only take into consideration changes in the over-education status as a result of occupation changes[38] whereas another route out of over-education could also take place via promotion within ones'occupation (Piper, 2015). For example, in the case of occupations in which individuals enter occupations that require a lower level of education but have higher prospects of progression wihin that occupation, the exit route from over-education is likely to be promotion rather than ocupational change and this stays uncaptured by objective measures possibly driving the level of state dependence upwards. In other words, if a worker's occupational category

[38] In the present sample, 87.58% of the respondents stay in the same occupation for the whole period they are observed whereas 87.27% stay in the same job (i.e. with the same employer).

does not change following a promotion, they will still be considered as over-educated even if they have actually escaped over-education meaning that the observed permanency and state dependency of over-education could be over-estimated. The key issue is whether the occupation codes used are narrow enough to pick up promotions – even if, ideally, a promotion should result in a new occupation code, because a new job is being performed. This is a downside of most of the studies of over-education in the literature especially the ones using objective over-education measures. In this chapter this is made worse due to the relatively small number (20) of occupation codes on which the over-education variable is based. For example, even if a cashier being promoted to a store manager would surely change occupation code, even within the 20 categories used in this chapter, a cashier being promoted to a section manager, or a nurse being promoted to a ward sister may end up being recorded as not changing occupation code and therefore not changing their over-education status, when actually they have. Piper (2015) suggests that such a bias can only be removed if for example future studies utilise very detailed occupational data. For the moment, such data especially for the case of Cyprus are not available. Nevertheless, as long as there is not too much variation in the required education level of different jobs within these occupation codes, even disaggregating to a lower level occupation classification would still give everyone in the higher level occupation the same required education level.

Lastly, another limitation of the RM method is that the quality of education is difficult to take into account (Sloane et al., 1999). However, the fact that the education variable used in this chapter refers to the highest level of education attained rather than simply years of education as is often found in the over-education literature (e.g. Piper, 2015), mitigates the concern about variations in terms of the quality of education.

2.3.4 Over-Education Summary Statistics

The proportion of over-educated workers in the final employee sample[39] amounts to 16.32%[40]. Grouping age into age groups in order to have more comparable groups in terms of cell numbers, it appears that the 25-29 age group has the largest percentage of over-educated employees with 29.76% followed by the 35-39 year olds with 25.24% and the 30-34 age group with 20.48%, while the two smallest groups, the youngest (20-24) and oldest (40-44) age groups, have the smallest percentages of over-educated workers, 15.24% and 9.29% respectively. In general, the percentage of people who are over-educated for their jobs does not show dramatic variation in terms of age and this refutes possible criticisms that the over-education dummy could be capturing the cohort change of increasing qualifications amongst the young (see e.g. Piper, 2015).

The table below shows the evolution of the incidence of over-education in the present sample, broken down according to the year of survey.

Table 2.3: The Incidence of Over-Education by Year

Year	Percent Overeducated
2005	15.67
2006	15.38
2007	14.92
2008	16.40
2009	16.42
2010	17.17
2011	22.35
Overall	16.32

[39] The employee/working sample here refers to the population in work and working as employees (excluding the self-employed) and who are up to 40 years of age in the first instance they are observed, not working in the armed forces or skilled agriculture and having at least a secondary education and who have at least 2 usable observations in the data (i..e when they were working as employees)
[40] It is reminded here that the modal over-education measure used to obtain this statistic produces lower estimates of over-education, on average, compared to other measures as discussed on page 45-46.

As demonstrated in the above table, with the exception of 2006 and 2007 where the rate of over-education dropped slightly, the overall trend from then onwards has been a steady increase in the incidence of over-education in Cyprus. Such a pattern of growing over-education incidence over time is consistent with increasing participation in higher education of young people as demonstrated in Chapter 1 as well as with the financial crisis. The onset of the financial crisis in Cyprus could also potentially explain the jump in the incidence of over-education in 2011 as it is likely to have caused a reduction in job openings for young people in professional jobs.

The following table presents the over-education transition probabilities matrix, i.e. the change in one categorical variable, in this case over-education, over time.

Table 2.4: Transition Probabilities Matrix

	Over-education at time t+1	
Over-education at time t	**0**	**1**
0	**98.87%**	**1.13%**
1	**3.70%**	**96.30%**

The above table of transition probabilities shows that 96.3% of those over-educated in one year are still over-educated the next year whereas only 3.7% of those who are over-educated in one year find a matched job in the next year. This is strong evidence suggesting that over-education is a permanent state/long-run phenomenon for the great majority of respondents who have an over-education experience and could also be a signal of a state dependency in over-education, where being over-educated in year one causes over-education in year two for example.

Table 2.5, below, draws a picture of the longitudinal patterns of over-education present in the data. Zeros denote observations of respondents who are not over-educated at the time of the survey whereas ones represent observations of respondents who are over-educated at each survey round with zeros and ones placed in chronological order. For example, the first row shows the number and percentage of respondents with two observations in the data set who were not over-educated in either of the two occasions that they were observed. Similarly, row three shows respondents who are observed three times within the data set and who were not over-educated in any survey round while for example, row four presents the frequency and percentage of respondents who had four observations in the data and were not over-educated in the first three survey round but became over-educated in the fourth survey round that they were observed.

Table 2.5: Patterns of Over-education History

Over-education History Patterns	Frequency	%
00	302	5.87
000	465	9.03
0000	3436	66.76
0001	20	0.39
001	12	0.23
0011	40	0.78
01	4	0.08
0100	12	0.23
011	15	0.29
0111	32	0.62
10	2	0.04
100	18	0.35
1000	20	0.39
101	3	0.06
11	54	1.05
110	3	0.06
1100	12	0.23
111	129	2.51
1110	4	0.08
1111	564	10.96

The above table shows that 14.52 % of the sample stay over-educated during all survey rounds, while 3.83% who are not over-educated for the full period in which they are observed spent at least one year in over-education. Moreover, of those over-educated the first time they are observed, 92.4% remain over-educated in every subsequent survey year, making over-education a long-run and persistent phenomenon. Table 2.5 is consistent with the transition matrix in Table 2.4, in that it shows few transitions while it also draws a picture about the consistency in the over-education status over a longer period than just t and t+1.

2.3.5 Descriptive Statistics

Table 2.6 below provides summary statistics for all the explanatory variables used throughout the econometric analysis.

Table 2.6: Descriptive and Summary Statistics

	Mean	Standard Deviation
<u>**Variables Used for the Derivation of the Over-education Variable:**</u>		
Education (Lower Secondary)		
Upper Secondary	0.442	0.50
Post-secondary Non-tertiary	0.032	0.18
1st and 2nd stage Tertiary	0.434	0.496
Occupation (Legislators, senior officials / managers, corporate managers etc.		
Physical, mathematical and engineering professionals	0.028	0.166
Life science and health professionals	0.027	0.162
Teaching professionals	0.075	0.264
Other professionals	0.053	0.225
Physical and engineering science associate professionals	0.045	0.208
Life science and health associate professionals	0.016	0.125
Other associate professionals plus &Teaching associate professionals	0.149	0.357
Office clerks	0.162	0.368
Customer services clerks	0.053	0.224
Personal and protective services workers	0.089	0.285
Models, salespersons and demonstrators	0.053	0.223
Extraction and building trades workers	0.057	0.232
Metal, machinery and related trades workers	0.038	0.190
Precision, handicraft, craft printing and related trades workers	0.008	0.089
Other craft and related trades workers	0.008	0.087
Machine operators and assemblers plus small cell of stationary-plant and related operators	0.015	0.121
Drivers and mobile plant operators	0.028	0.164
Sales and services elementary occupations	0.054	0.225
Labourers in mining, construction, manufacturing and transport& Agricultural, fishery and related labourers	0.031	0.174
<u>**Dependent Variable:**</u>		
Over-education	0.163	0.37
<u>**Independent Variables used in Regression Analysis:**</u>		
Initial Over-education	0.227	0.419
Age	32.05	6.23
Years of Work Experience	10.65	6.44
Married(Single)	0.61	0.49
Female (Male)	0.53	0.50
Temporary contracts (Permanent-unlimited duration contract)	0.12	0.32
Part time work (Full time work)	0.06	0.24
Limited in Activity because of health problems (no limitation)	0.06	0.23
Recent Entry into the labour market:		
No recent entry	0.87	0.33
From Unemployment	0.09	0.28
From other inactivity	0.03	0.18
Reason of Job Change		
No job change	0.87	0.33
Self-induced job change	0.10	0.30
Involuntary Job change	0.03	0.16
Proportion of past year spent in unemployment	0.04	0.13
<u>**Macro Level Variables**</u>:		
Initial Unemployment upon start of paid employment by sex	3.64	1.37
Unemployment by sex and age group	5.41	3.20
Annualised change in the labour supply by educational category and sex	0.02	0 .03
Annualised change in the employment share by occupation and sex	0.0012	0.073

Note: Omitted Reference categories in brackets

In terms of the variables based on which the over-education variable is calculated and as demonstrated by the above table, the majority of the sample have an upper secondary education (44.2%), closely followed by 1st and 2nd stage tertiary education (43.4%), while 9.3%[41] of the respondents have a lower secondary education and only 3.14% a post-secondary non-tertiary education. Following the reclassification of the occupation variable so as to ensure satisfactory cell sizes, there are twenty different occupations in the sample. The largest occupation group is office clerks who occupy 16.18% of the overall sample whereas the smallest occupation group is the other craft and related trades workers who only make up 0.76% of the sample.

As mentioned earlier, in terms of the micro level determinants of over-education examined, these are grouped into personal and job characteristics and work history variables. The personal characteristics examined are gender, age, years of work experience, limitation in activity because of health problems and marital status. Firstly, the likelihood of over-education is expected to be higher for women if they act as ‘tied stayers’ or ‘tied movers’ (Mincer, 1978) and to decrease as age increases and as workers acquire more work experience. Furthermore, activity limitation because of health problems may increase the likelihood of over-education as disabled individuals could face difficulties in finding work and making a good job match. On the other hand and as pointed out by Dolton and Silles (2001), disabled individuals may work in matched jobs as a result of government legislation or because of their personal determination to find a good match. Lastly, being married, even though it has been found to increase the chances of over-education due to considerations such as relocation because of a partner’s job (Dolton and Silles, 2001), is expected to have a less clear effect in the case of Cyprus, due to the small country size and minimal travelling distances.

The job characteristics controlled for in the over-education equation are: part-time and temporary work arrangements which are both expected to increase the likelihood of over-education. According to Dolton and Silles (2011), people working in part-time jobs may not be able to fully use all forms of human capital including qualifications attained while those working under a temporary contract may see such jobs as a ‘quick fix’ rather than

[41] A weakness of the present chapter as a result of the fact that there are only 20 broad occupation categories is that, because there are only a few low educated people in the sample, they do not form the modal category in one of these large occupations. As a result, no over-educated people are found either in lower or upper secondary education-all of them are gathered in post-secondary/non-tertiary (38.95%) but mostly in tertiary education (61.05%).

a permanent life-long career, increasing the chances of accepting jobs for which they are over-educated.

In terms of work history, the proportion of the past year spent in unemployment, recent entry into the labour market and self-induced versus involuntary job change are investigated in this chapter. It is expected that time spent in unemployment serves as a negative signal for prospective employers as human capital may depreciate during this time as well as due to the fact that previous unemployment spells may be perceived as caused by the job seeker's lower productivity. This was conceptualised theoretically by Vishwanath's (1989) stigma effect model according to which a candidate's unemployment duration could act as a signal of otherwise unobservable components of her/his productivity, with longer unemployment periods suggestive of the possibility that employers learned that the worker was unproductive. These negative signals could in turn increase the chances of over-education as employers may compensate the perceived lower quality of these job seekers by offering lower level jobs that do not match their formal education level. At the same time, the likelihood of over-education might be expected to decrease as the proportion of the past year spent in unemployment increases if workers deliberately choose to wait in unemployment rather than to accept jobs for which they are over-educated. The same reasoning applies for the effect of recent entry in the labour market from unemployment and from other inactivity. Lastly, workers who voluntarily (i.e. self-induced reasons) change job in the past year are expected to have a lower likelihood of over-education as they are more likely to have found themselves a good match before quitting their old job while involuntary job change is expected to have the opposite effect as it can send out a negative signal to new employers.

As mentioned earlier, apart from the micro-level independent variables, a number of macro-level variables are also included in the regressions. The first macro variable examined in this chapter is the unemployment rate at start of first paid employment. This ranges from 1.3 % of the economically active (in 1992) to 8.1% (in 2011) with an average of 3.67% (2.68% for men and 4.55% for women). It is expected that a higher unemployment level upon start of paid employment increases the likelihood of over-education at present if workers initially accept a job for which they are over-educated due to limited opportunities for matched work and then fail to find a good match after that. For example, Dolton and Silles (2001) argue that graduates entering the labour market during a recession and who accept non-graduate jobs may find it difficult to recover from this bad start to their careers, as it can send out a negative signal to prospective employers who might then be reluctant to offer a matched job. Apart from macro conditions at the time of labour market entry, current labour market conditions could also be said to affect the likelihood of over-education and most importantly its persistence as they may affect the chances of over-educated individuals escaping over-education. For example, a high level of unemployment can force people to stay in jobs for which they are over-educated, as opportunities for matched work are limited. As demonstrated in Table 2.6, the unemployment rate at the time of the survey by age group[42] is on average 5.41% of the economically active population. However, this figure masks important differences among the various age groups. For example, the average unemployment rate for ages 16-19[43] stands as high as 15.92% and at 11.53% for ages 20-24, while the rates for the other groups are substantially lower and declining as age increases (6.19% for ages 25-29; 3.91% for ages 30-34; 3.76% for ages 35-39 and 3.67% for ages 40-44). Hence, as expected, the unemployment rate is consistently higher for younger cohorts.

Moreover, an increase in the number of people with an equivalent level of qualifications i.e. an increase in the labour supply by education level and sex, can be argued to increase the competition faced by individuals in the labour market and hence reduce the opportunities of job seekers or of already over-educated individuals to find a good match. The annualised change in the supply of graduates by educational category and sex during the years of the study is on average 2%. More specifically, males with lower and upper secondary education and post-secondary (non-tertiary) education have an annualised increase of 0.8% and males with 1st and 2nd stage of tertiary education an annualised increase of 3.06% demonstrating a substantially larger rise in the number of tertiary level graduates and hence candidate/employee competition compared to secondary level graduates. The corresponding figures for females are -0.77% and 5.06% demonstrating a decline in the number of secondary level graduates and a substantial growth in the number of tertiary level female graduates in the period covered by this study.

On the other hand, an increase in the employment share by occupation, a proxy of labour demand within one's occupation, is expected to increase the opportunities of finding a good match and/or escaping over-education both within and outside the firm. In terms of the percentage change in the employment share by occupation and sex from one survey

[42] The age groups used here are the following: 16-19; 20-24; 25-29; 30-34; 35-39 and 40-44.
[43] This is significantly small group that makes up only 0.8% of the overall sample

year to the next i.e. the annualised change, this stands at an average percentage change of 0.12%. For example, Legislators, senior officials and Managers experience the largest increase with a mean annual change in employment share of 12.26% followed by Professionals with a percentage increase in their employment share of 2.8%, Technicians who experience an average percentage increase of 1.01% and Elementary workers with a 0.6% average percentage increase during the years of the survey. Clerks, Service and Sales workers, Craft and related workers and Plant and machine operators all experience an average fall in their employment share with the respective figures being, -1.73%, -0.06%, -3.1% and -1.88%.

Interaction terms between some of the macro variables and years of work experience are also incorporated in further analysis so as to test whether negative macro conditions have a greater effect on people with less work experience and subsequently whether work experience shelters workers against over-education during adverse macroeconomic situations.

2.4 Methodology and Model Specifications

This section overviews the various estimation methods used in the analysis of the determinants of over-education and its persistence. These non-linear (probit) models include both static and dynamic specifications and the pros and cons of each of the models are discussed. The probit models used are the static and dynamic pooled probit, random effects probit and the random effects probit model with Mundlak correction. The Wooldridge (2005) dynamic random effects probit model with Mundlak corrections is also used.

2.4.1 Binary Choice Probit Models (Static)

In order to examine the determinants and character of over-education, a series of binary response probit models are employed. Probit models are commonly motivated as latent variable models where the (latent) variable/outcome of interest, in this case the propensity to be over-educated, cannot be observed or measured.

In all the model specifications described below the determinants of over-education are split into micro and macro level. In turn, and as discussed in detail in the previous section, the micro variables are split into personal and job characteristics (marital status, sex, age, limitation in activity because of health problems, contract type i.e. permanent vs temporary and part time vs full-time work arrangements); work history variables (number of years of work experience, proportion of the past year spent in unemployment, self-induced or voluntary job change vs employer-induced or involuntary job change and other reasons for job change, over-education in the previous year and recent entry into the labour market during the past year). The macro variables are the unemployment rate at the time of the survey by sex and age group, the unemployment rate at the start of paid employment by sex, the annualised change in the number of graduates by educational level and sex and the annualised change in the employment share by occupation and sex. Year dummies are also included.

2.4.1.1 The Pooled Probit Model

Even though not the correct specification given the panel nature of the data set, the pooled probit model will be used as a benchmark for the analysis in this chapter as it can provide an overview of the relationships of interest in terms of the cross sectional differences in the sample and can therefore be largely informative in a descriptive sense (Mavromaras et al., 2010).

The unobserved or unmeasured (latent) variable $\mathrm{OE^*}_\mathrm{i}$, in this case over-education, has the following regression:

$$\mathbf{OE^*}_\mathbf{i} = \mathbf{x_i}\boldsymbol{\beta} + \mathbf{u_i} \qquad \textbf{(2.1)}$$

where $\mathrm{u_i}$ is assumed to have mean 0 and a standard normal distribution with (known) variance 1.

What is observed is only a dichotomous indicator of the latent variable defined by the below expression:

$$OE_i = \begin{cases} 1 \text{ if } OE^*_i > 0 \\ 0 \text{ if } OE^*_i \leq 0 \end{cases} \quad \textbf{(2.2)}$$

In this case, this dichotomous variable is the indicator of whether the highest level of education attained by the respondent is greater or not than the modal educational level within one's occupation.

The statistical specification of this binary choice model is then given by the following expression:

$$\mathbf{Pr(OE_i = 1) = F(x_i\beta) = \Phi(x_i\beta)} \quad \textbf{(2.3)}$$

Specifically, it is assumed that the model takes the form where Pr denotes probability, and Φ is the Cumulative Distribution Function (CDF) of the standard normal distribution. And the parameters β are estimated by maximum likelihood. In treating the data as cross-sectional, the pooled probit model ignores the fact that even though independence across individuals is assumed this is not the case for the correlation over time for a given individual. In other words, even though observations of different people are thought to be independent, observations of the same person are likely to be clustered together or correlated, biasing the standard errors and leading to wrong inferences. According to Gibbons and Hedeke (1996), treating these data as if they were independent (i.e. from separate individuals) would provide overly optimistic or too small estimates of precision (i.e. standard errors) (Gibbons and Hedeke, 1996). The fact that heterogeneity is ignored completely means that pooled regression estimates are subject to biases and for this reason, the very restrictive assumption of no unobserved systematic individual effects, will be relaxed in subsequent model specifications[44].

2.4.1.2 Random Effects Probit Models

Since even modest changes in standard errors can have large effects on statistical inference (Miller et al., 2009), it is important to correct for the fact that observations of the same individual in a panel and hence the standard errors, cluster (or are correlated). The random effects probit model allows for serial correlation in the unobserved factors determining OE_{it}, i.e. in ($c_i + u_{it}$) in equation (2.4) below[45]. According to Mavromaras et al. (2010), the main advantage of introducing the random effects model is that it allows controlling for unobserved effects that do not change over time and hence permits accounting for some of the unobserved differences between mismatched and matched workers within the model.

Therefore, to take account of the panel nature of the data and to allow for individual effects (unobserved heterogeneity) using the latent variable framework, the binary choice model can be re-written as:

$$\mathbf{OE^*_{it} = x_{it}'\boldsymbol{\beta} + c_i + u_{it} \qquad u_{it} \sim N(0,1) \qquad (2.4)}$$

where x_{it}' is a vector of explanatory variables and time/year dummies associated with observation i at time t (not all variables within x vary both across individuals and over time), c_i is the individual-specific random component capturing the effect of time-invariant individual unobserved heterogeneity and u_{it} is an idiosyncratic error term associated with each observation i at time t and follows a normal distribution $N(0, \sigma^2_u)$.

The observable dependent variable in the model OE_{it} is defined as follows:

44 The assumption of no unobserved individual effects allows the use of a cross sectional model to estimate the parameters of interest even with panel data.
45 In Stata, the xtprobit command is used with the vce(cluster clustvar) option where the clustvar is the personal identification variable(pid). This option ensures a consistent VCE estimator when the disturbances are not identically distributed over the panels or there is serial correlation in u_{it}.

$$\mathbf{OE_{it} = 1 \text{ if } OE^*{}_{it} > 0} \qquad \textbf{(if over-educated)}$$

$$\textbf{(2.5)}$$

$$\mathbf{OE_{it} = 0} \qquad \textbf{(if not over-educated)}$$

The statistical specification of the binary choice model in the presence of heterogeneity is then given by:

$$\mathbf{Pr(OE_{it} = 1|x_{it}, ci) = F(x_{it}{}'\beta, u_{it})}$$

A key assumption underlying this estimator is the unrealistic and restrictive assumption of no correlation between individual heterogeneity, c_i and the explanatory variables $x_{it.}$

2.4.1.3 Random Effects Probit Model with Mundlak (1978) Correction

As mentioned above, a drawback of the random effects model is its unrealistic assumption pertaining to the independence between the covariates and the error term (Mavromaras et al., 2010). According to Mavromaras et al. (2010), this is not a viable assumption as it is not supported by empirical evidence most of the time. The Mundlak (1978) correction provides a solution to this problem from within the random effects framework and in doing so helps towards accounting for potential correlations between the explanatory variables and the individual specific component of the error term therefore allowing for the control of individual heterogeneity (Mavromaras et al., 2010). In other words, the Mundlak (1978) correction is correcting for the effects of any unobserved characteristics that do not vary over time, with ability likely to be one of the most important.

The Mundlak correction assumes that the relationship between c_i and the means of the time-varying x-variables can be witten as $c_i=\bar{x}'_i\delta+\varepsilon_i$, where $\varepsilon_{i\sim}$iid follows a normal distribution and is independent of x_{it} and u_{it} for all i and t and δ is the coefficient on the individual-specific variable mean $\bar{x}'_{i.}$

In practice, the Mundlak correction is applied by including the individual means of each of the time-varying variables that are assumed to be correlated with the unobserved heterogeneity (Mundlak, 1978) on the right hand side of the regression equation hence permitting an interpretation of the point estimates as being pure within-person effects (Boll et al., 2016). In the case of the determinants of over-education, the individual means over age, years of work experience and marital status are included on the right hand side of the random effects probit regression. Nevertheless, it has to be noted that even though individual means of all time variant covariates that are suggested to be potentially correlated with individual unobservable heterogeneity are included in the regression, the possibility that time invariant variables are also correlated with individual unobservable traits cannot be ruled out (Boll et al., 2016).

The random effects probit specification with Mundlak correction is expressed as follows:

$$\mathbf{OE^*}_{\mathbf{it}} = \mathbf{x_{it}}'\boldsymbol{\beta} + \bar{\mathbf{x}}'_{\mathbf{i}}\boldsymbol{\delta} + \boldsymbol{\varepsilon}_{\mathbf{i}} + \mathbf{u_{it}} \qquad \mathbf{(2.7)}$$

2.4.2 Dynamic Probit Models

Dynamic probit models have been motivated in different ways in the literature. In this chapter, the primary reason for using dynamic specifications is to examine whether over-education in the previous period increases the likelihood of over-education in this period (and hence the coefficient of the lagged dependent variable is of primary interest). Nevertheless, the use of dynamic estimation in the present chapter is furthermore necessitated by the high over-education persistence observed in the aggregate over-education (as shown by the transition probabilities in Table 2.4), as well as the serial correlation present in the idiosyncratic error term of the static models[46], even in the case of the random effects with Mundlak corrections model, which all call for a respecification of the static models. According to Bond (2002), allowing for dynamics in estimation can be crucial for obtaining consistent estimates of the parameters even when the lagged dependent variables coefficients are not of major interest (Bond, 2002). Piper (2012), adds that the presence of first order serial correlation in the idiosyncratic error term points

[46]Tested using Wooldridge's (2002) test for serial correlation, implemented in Stata by the user-written xtserial command (Drukker, 2003), which rejects the null of no first order autocorrelation.

to the fact that there are omitted dynamics in the FE estimates which means that static panel analysis estimates are no longer efficient and possibly misspecified (Piper, 2012)[47]. Piper (2013) discussing King and Roberts' (2015) study of robust standard errors further adds that serial correlation in the idiosyncratic error term, should be seen as an opportunity to take advantage of this information and respecify the static model by for example employing dynamic panel methods rather than be treated as a problem to be fixed by adjusting the standard errors (Piper, 2013).

According to Greene (2008), adding dynamics to a model alters the interpretation of the equation. In the absence of the lagged variable, the independent variables reflect the full set of information behind the observed outcome while in its presence, the entire history of the right-hand-side variables is incorporated in the equation. This means that any measured effect is conditional on this history; in this case, any impact (and therefore the interpretation) of the independent variables is representative of the effect of new information, whereas the lagged dependent variable reveals the effect of the past (Greene, 2008).

Similar to the static models and as described above, the dynamic pooled probit specification ignores heterogeneity completely and so is likely to overestimate the coefficient on lagged over-education (Soderbom, 2009). It is nevertheless used here as a benchmark for comparison to the random effects specifications. The dynamic random effects probit model and the dynamic random effects probit model with Mundlak corrections are also run and are the same as the static versions described above with the addition of the lag of over-education in the regressors.

For example, the Dynamic Random Effects Probit model is demonstrated by the latent equation:

$$\mathbf{OE^*_{it} = \gamma_i\, OE_{i\,t-1} + x'_{it}\,\beta + c_i + u_{it}} \tag{2.8}$$

Where i=1,….,N denotes individuals observed over t=1,…T periods. OE^*_{it} is the latent dependent variable for being over-educated with the observable outcome as it appears in

[47] Piper (2012) also claims that dynamic modelling can offer a robustness check for static regressions.

(2.5). $OE_{i\,t-1}$ represents the lag of the dependent variable, i.e. the over-education status of the individual at t=-1 with γ_i being the coefficient of interest to be estimated. x_{it}, c_i and u_{it} are as in the static specifications described in Section 2.4.1.

2.4.3 Wooldridge (2005) Dynamic Probit

A potential problem stemming from the above dynamic model specification is that, when modeling a dummy variable y_{it} (in this case OE_{it}) that is a function of the lagged dependent variable, $y_{i\,t-1}$ (in this case $OE_{i\,t-1}$), the lagged dependent variable may be correlated with the error terms. More specifically, given that a person's employment history in the data is not observed from the very beginning, there is a risk that the initial value arising from a person's first observation in the sample is conditional on observed or unobserved variables in the unknown past of that person. This means that the initial value of a respondent's over-education might be affected by his or her previously held over-education status (Boll et al., 2016). In other words, in a dynamic equation, any unobserved heterogeneity could be picked up by the lagged dependent variable, e.g. whatever made people over-educated in the first place could still be making them over-educated at present. For example, a person of a lower ability might have an increased probability of initially becoming over-educated, which in turn overestimates his likelihood of remaining over-educated as this unobserved ability is correlated with the error term. This could lead to a high persistence and spurious state dependence in over-education. This is known as the initial conditions problem (Heckman 1981; Blundell and Bond 1998; Arellano and Carrasco 2003). Not taking initial conditions that are correlated with the individual-specific error term u_i into account results in an overestimation of state dependence which means that the estimated coefficient γ_i of $OE_{i\,t-1}$ in Equation (2.8) will be larger than the true value of state dependence.

Three methods have been suggested in order to correct for the fact that in a dynamic setup y_{i0} is likely to be correlated with unobserved heterogeneity c_i affecting y_{it}. The initial conditions problem was first examined by Heckman (1981), followed by less computation-intense estimators by Orme (1997), Arulampalam and Stewart (2009) and Wooldridge (2005). Given that the three methods' performance in the context of dynamic probit models is equivalent (Arulampalam and Stewart, 2009) and the fact that Woodridge's approach is simpler to implement in practice, similar to what is often done in the literature, this is the preferred method applied in this chapter. Wooldridge (2005), suggests including the individual's over-education outcome in year $t=1$ as an additional covariate that captures part of the unobserved heterogeneity between persons.

Another probem arising from equation (2.8) above, and like in the static random effects probit model case, is the unrealistic assumption of independence between the covariates and the unobervered heterogeneity term which is likely to bias the results. As explained in the previous section, this is resolved through the application of the Mundlak (1978) correction which in this case is combined with the Wooldridge initial conditions correction and expressed by:

$$\mathbf{OE^{*}_{it}} = \boldsymbol{\gamma_i}\, \mathbf{OE_{i\,t-1}} + \mathbf{x'}_{\boldsymbol{it}}\, \boldsymbol{\beta} + \bar{\mathbf{x}}'_{\boldsymbol{i}}\boldsymbol{\delta} + \boldsymbol{\theta}\mathbf{OE_{i\,1}} + \boldsymbol{\varepsilon}_{\mathbf{i}} + \mathbf{u_{it}} \qquad \mathbf{(2.9)}$$

Equation (2.9) above, is expected to reveal the true extent of over-education state dependence.

2.5 Estimation Results and Discussion

2.5.1 Micro-level Determinants and Over-Education State Dependence

Table 2.7 below reports the marginal effects of the different probit models. The dependent variable is over-education (equal to 1 if over-educated). The various specifications presented below are defined as follows: Specification 1: Static Pooled Probit Model; Specification 2: Dynamic Pooled Probit Model; Specification 3: Dynamic Random Effects Probit Model; Specification 4: Dynamic Random effects Probit with Mundlak Correction and Specification 5: Wooldridge Dynamic Probit with Mundlak Correction[48]**.**

[48] Given that Mundlak correction coefficients incorporate 'both the steady state relationship and any correlation between the specific variable and the error term and are hence difficult to interpret' (Mavromaras and McGuiness 2012,622) they are omitted from the table of results here.

Table 2.7: Micro-level Determinants of Over-Education Regression Results

	Spec. 1	Spec. 2	Spec. 3	Spec. 4	Spec.5
Age	**0.003****	**-0.000**	**-0.00**	**0.003**	**0.005**
	(0.001)	(0.001)	(0.001)	(0.005)	(0.005)
Work Experience	**-0.007*****	**0.000**	**0.000**	**-0.002**	**-0.003**
	(0.002)	(0.001)	(0.001)	(0.005)	(0.005)
Married	**-0.011**	**-0.020*****	**-0.020***	**-0.063****	**-0.06****
	(0.012)	(0.005)	(0.011)	(0.032)	(0.03)
Female	**0.143*****	**0.011****	**0.011***	**0.010***	**0.10****
	(0.010)	(0.004)	(0.006)	(0.006)	(0.005)
Health limitation	**0.007**	**-0.005**	**-0.005**	**-0.005**	**-0.003**
	(0.022)	(0.007)	(0.008)	(0.008)	(0.008)
Temporary Contract	**0.018** (0.016)	**0.009** (0.008)	**0.009** (0.008)	**0.009** (0.008)	**0.01*** (0.007)
Part time	**-0.034***	**0.013***	**0.014**	**0.013**	**0.012**
	(0.019)	(0.009)	(0.012)	(0.010)	(0.009)
Past Year unemployment	**0.182**** (0.056)	**-0.038** (0.043)	**-0.038** (0.048)	**-0.04** (0.046)	**-0.04** (0.039)
Over-education at t-1	-	**0.145*****	**0.145****	**0.145****	**0.91*****
		(0.013)	(0.072)	(0.06)	(0.02)
Recently employed					
a) from unemployment	**-0.062***** (0.022)	**-0.001** (0.016)	**-0.001** (0.017)	**-0.001** (0.017)	**-0.001** (0.015)
b)from other inactivity	**-0.025** (0.028)	**-0.019*** (0.009)	**-0.019*** (0.011)	**-0.019*** (0.010)	**-0.017**** (0.009)
Job Change					
a) Self-induced	**0.022**	**-0.002**	**-0.002**	**-0.002**	**-0.002**
	(0.017)	(0.008)	(0.009)	(0.008)	(0.008)
b)Employer and other reasons	**0.033** (0.033)	**0.027** (0.019)	**0.027** (0.017)	**0.027*** (0.017)	**0.03*** (0.015)
Year Dummies					
2006	**-0.001**	-	-	-	-
	(0.024)				
2007	**-0.007**	**-0.013**	**-0.013**	**-0.014**	**-0.016**
	(0.022)	(0.011)	(0.014)	(0.014)	(0.014)
2008	**0.006**	**-0.011**	**-0.011**	**-0.012**	**-0.014**
	(0.022)	(0.011)	(0.012)	(0.013)	(0.013)
2009	**0.008**	**-0.015**	**-0.015**	**-0.015**	**-0.017**
	(0.023)	(0.011)	(0.013)	(0.013)	(0.013)
2010	**0.015**	**-0.004**	**-0.004**	**-0.006**	**-0.008**
	(0.025)	(0.011)	(0.011)	(0.013)	(0.014)
2011	**0.071****	**-0.009**	**-0.009**	**-0.01**	**-0.013**
	(0.032)	(0.011)	(0.012)	(0.014)	(0.015)
Initial Over-education	-	-	-	-	**0.02****
					(0.008)
N	**5333**	**3646**	**3646**	**3646**	**3646**

Notes: Standard errors clustered around personal ID in parentheses; For the cross-sectional probits margins, dydx(*)are reported while for the panel probit models margins, predict(pu0) dydx(*) is used instead; Significance denoted by: *** p<0.01, ** p<0.05, * p<0.1; In Specifications 4 and 5, individual means over age, years of work experience and marital status as per the Mundlak correction are included.

As can be seen from the regression results in Table 2.7, and similar to the over-education literature, being a female increases one's chances of being over-educated in all probit specifications while being limited in activity because of health reasons is not found to significantly affect the probability of over-education in any specification. Age[49] and work experience are only found to significantly affect the probability of over-education in the benchmark model i.e. pooled static probit model. More specifically, keeping everything else constant, contrary to what was expected and to previous studies of over-education, age appears to be increasing the probability of over-education. This might be explained by the fact that, keeping work experience (and other variables) constant, an older age might send out a negative signal that may cost workers a good match or they may be considered as less productive and with fewer prospects of advancement than their younger counterparts. The effect of age, nevertheless disappears in the dynamic specifications demonstrating that age could have been picking up the effect of state dependence of over-education where older workers might find it more difficult to escape over-education or in other words their current over-education status might be causing their next year's over-education. Years of work experience are found, as expected, to slightly reduce (by 0.7 percentage points per year) one's chances of over-education but this variable is only found to significantly affect the probability of over-education in the static pooled probit specification. All in all, age and work experience do not seem to affect the probability of over-education in any of the dynamic models and controlling for previous over-education status, age and experience are largely irrelevant, meaning that it is history of over-education that is important. This is because, as mentioned earlier, any measured influence of independent variables in dynamic models is conditional on the over-education history and hence any impact (and hence the interpretation) of the independent variables represents the effect of new information, whereas the lagged dependent variable reflects the influence of the past. Squared terms for age and years of work experience were also included in all specifications but they were not found to attract significant coefficients, meaning that there does not appear to be a non-linear effect of age/work experience on over-education.

Being married on the other hand is found to significantly reduce the probability of over-education in all specifications except in the static one, where its sign is again negative though not significant. The negative relation between being married and over-education

[49] Results are robust to dropping employees less than 20 years old as they are a small group in terms of number.

can be explained by the possibility that married individuals can afford to wait longer until they find a matched job due to the fact that they can rely on their partner's income until they do so[50].

Job characteristics such as having a temporary contract is only found to slightly increase the chances of over-education yet only in Specification 5 while part time work has a mixed effect on the probability of over-education. More specifically, part time work is found to reduce the probability of over-education in the static pooled specification (Spec. 1) while it is found to slightly increase it in the dynamic pooled specification (Spec.2). The effect of working on a part time basis becomes insignificant once random effects are introduced and once unobserved heterogeneity is controlled for, pointing to the fact that any effect reflected in the pooled specifications might be due to unobserved characteristics of individuals who work on a part-time basis.

In terms of the work history variables, the proportion of the past year spent in unemployment is found to be increasing one's chances of over-education and recent entry into the labour market from unemployment to be significantly decreasing one's chances of over-education, yet only in the pooled static specification, possibly because they are picking up some of the missing dynamics or state dependence of past over-education. On the other hand, recent entry into the labour market from other inactivity is found to be decreasing the probability of being over-educated in all the dynamic specifications yet not in the static pooled probit. The negative relation between recent entry from other inactivity and the probability of over-education might be explained by the fact that people making the choice to leave inactivity and enter the labour market have waited to enter the labour market until they find a matched job rather than accept a job for which they are over-educated. Lastly, self-induced job change is not found to be significantly affecting the probability of over-education in any specification while involuntary job change is found, as expected, to significantly increase the probability of over-education once fixed effects are introduced via the Mundlak (1978) correction.

Overall, the fact that the dynamic pooled and random effects results are almost identical for all of the variables suggests that unobserved heterogeneity is not very important in this case.

[50] Information on parental status that could be argued to have the potential to affect/alter this result is not available in my dataset.

Last but not least, focusing on the dynamic specifications, what is clear from Table 2.7 is the highly significant, positive effect of past years' over-education on this year's over-education which demonstrates that over-education is a self-perpetuating state. This is the case, even after the application of econometric corrections to eliminate spurious state dependence (for example, in Specification 5). As can be seen in the table above, the marginal effect of the lagged over-education is robust to the different specifications and hence is not a spurious finding due to statistical modelling error. According to Mavromaras and McGuinness (2012), the intuition behind this finding is that the negative impact of those characteristics that were responsible for becoming over-educated in the first place, will be heightened via the continued presence of over-education, thus reinforcing the labour market disadvantages associated with over-education.

More specifically, the lag variable of over-education is found to be increasing the probability of over-education this year by 14.5 percentage points in the pooled, random effects and random effects with Mundlak correction models pointing to a strong state dependence of over-education, meaning that being over-educated in the previous year increases one's chances of over-education this year. The fact that the coefficient of the over-education in the past period is the same in the cross sectional (Spec. 2 , i.e. dynamic pooled) and dynamic panel specifications points towards the fact that individual heterogeneity does not have much effect. When initial conditions are controlled for, the coefficient on the lagged variable increases dramatically with workers who were over-educated in the previous period being 91 percentage points more likely to be over-educated in the next period compared with people who were not over-educated in the previous period. This result suggests that there is little movement into or out of over-education, which is consistent with the transition matrix in Table 2.4 and the over-education history patterns in Table 2.5 above. This is a large figure, however previous studies in the literature also find significant over-education state dependence. For example, Kiersztyn (2013) reports odds ratios for state dependence in the range of 4.3 to 7.45 meaning that being over-educated at t-1 causally increases the chances of an individual being over-educated at time t by 4 to 7 times. Mavromaras and McGuiness (2012) also find considerable state dependence of over-skilling in Australia.

As mentioned earlier, in order to guarantee sufficient sample sizes per occupation so as to have a reliable modal education level and hence over-education variable, only 20 broad occupation groups are defined. This means that, whereas with more narrowly defined occupations, a promotion would probably be expected to take a worker into a new occupation category, with only 20 occupations someone could stay within the same occupation code even if doing a different level of their job. This means that it could be the case that one of the main reasons for the persistence of over-education found in the results is that people do not seem to move jobs much. Nevertheless, the fact that the regressions presented in Table 2.7 do control for job change means that, even holding job change constant (e.g. amongst those who change jobs), those who were over-educated last year are still more likely to be over-educated this year. This is an important finding.

Lastly, the coefficient of the initial conditions variable is positive and significant meaning that being over-educated in the first instance a respondent is observed increases his/her chances of being over-educated at present by 2 percentage points[51]. Hence, not only over-education in the last period causes over-education this period but also over-education in the first instance a person is observed in the survey also carries over or causes over-education at present.

2.5.2 Over-Education State Dependence by Career Stage

Having concluded from Table 2.7 that state dependence exists in the data and in order to examine whether this self-perpetuating nature of over-education differs depending on the career stage one is in, separate regressions were run restricting the sample to each of the four career stages, listed below and using the random effects probit model with Mundlak and initial conditions correction. The aim here was to use a direct measure of the career stage one is in, rather than use age as a proxy, given the diverse educational and work profiles of people of the same age in the sample. For this reason, and since a direct measure of work experience is available in the data, the four career stages were defined based on this variable and are the following: 0-3 years of work experience-Career Stage 1 or Early career; 4-9 years of work experience- Early to Mid-Career; 10-20 years of work experience- Mid careers and more than 20 years of work experience- Late careers. However, given that only people up to the age of 40 in their first survey rotation were left in the final sample, it was not possible to correctly run the regression for this last group,

[51] As a robustness test regressions were also run after dropping those with only 2 observations in the sample (so just leaving people with 3 and 4 observations) so as to eliminate the possibility that initial over-education and the lag of over-education are measuring over-education in the same period. Results are almost identical when this is done.

due to its small sample size (6.39% of the whole sample) and hence the results for this group are omitted from Table 2.8. Table 2.8 below, shows the marginal effects for the lag of the over-education variable and the initial over-education coefficient for career stages 1 to 3. Full regression results for each of the 3 career stages are omitted here but can be found in Appendix 2E.

Table 2.8: State Dependence of Over-Education by Career Stage Regression Results[52]

	Career Stage 1 (0-3 years of work experience)	Career Stage 2 (4-9 years of work experience)	Career Stage 3 (10-20 years of work experience)
Over-education at t-1	0.88*** (0.04)	0.90*** (0.05)	0.97** (0.04)
Initial Over-education	0.01 (0.02)	0.023* (0.02)	0.004 (0.01)

Notes: Standard errors in parentheses; Significance denoted by: *** p<0.01, ** p<0.05, * p<0.1

Table 2.8 confirms the fact that over-education is highly state-dependent in all career stages while initial over-education is only statistically significant for those in early-mid careers. This means that, for those who have between 4 and 9 years of work experience, over-education carries over for a longer period of time than those who are in stages 1 or 3. In terms of the lag of over-education, it appears that being over-educated in the past year causes this year's over-education at an increasing rate as a respondent moves up the career ladder. More specifically, the coefficient on the lag of over-education rises with the career stage with those who have between 10 and 20 years of work experience facing the largest state dependence. This can be taken to mean that the self-perpetuating nature of over-education is present in all career stages and hence affects all respondents irrespective of whether they just entered the labour market or have 20 years of work

[52] Marginal effects from random effects probit model with Mundlak correction and initial condition correction and clustering around personal ID. Recent entry into the labour market variable omitted in order for all the group regressions to converge. It was not possible to get any results for Career Stage 4 (more than 20 years of work experience) due to the small sample size of this group. Results are robust to the choice of cut-offs between the work experience groups.

experience, pointing to the fact that over-education is a permanent phenomenon and that once in it, it is not easy to escape. In other words, being over-educated in one period will cause next years' over-education irrespective of how many years of work experience one accumulates. Moreover, the increasing coefficient on the lag variable means that being over-educated in one period will increase one's chances of being over-educated in the next period to a larger extent the more years of work experience a respondent has[53]. This could be potentially explained by the fact that being over-educated at a later stage in one's career might send a more negative signal to prospective employers than being over-educated at an earlier carer stage impeding the chances of over-educated employees to escape over-education and find a good match. Another channel via which this finding may be explained is that people with more work experience might have been over-educated for longer than people who are just entering the labour market or employees with only 2-3 years of work experience and might have knowingly accepted a mismatched job while they gain on-the-job experience. Unlike these early career-stage workers, employees in their early-mid or mid careers might become habituated to their jobs, lowering their on-the job search[54] for a good match and hence having a higher probability of being over-educated in the next year.

2.5.3 Macro Determinants of Over-Education

Table 2.9 below demonstrates the relations between over-education and the various macro conditions discussed earlier. Marginal effects from the dynamic random effects probit model with Mundlak correction and initial conditions correction are presented.

[53] However, the difference between career stage group 1 and 2 is in not statistically significant (t statistic= 0.313) and there is only a weak evidence that over-education state dependence significantly increases between stages 2 and 3.

[54] Chapter 3 of this thesis looks into the relationship between over-education and on-the-job search

Table 2.9: Macro-level Determinants of Over-Education Regression Results

Independent Variables	Marginal Effect	Standard Error
Age	0.004	(0.005)
Work Experience	-0.002	(0.005)
Married	-0.058**	(0.029)
Female	0.006	(0.007)
Health limitation	-0.002	(0.008)
Temporary Contract	0.013	(0.008)
Part time	0.015	(0.010)
Past Year unemployment	-0.027	(0.036)
Over-education at t-1	0.89***	(0.030)
Recently employed:		
a) from unemployment	-0.002	(0.013)
b)from other inactivity	-0.015 *	(0.010)
Job Change		
a) Self-induced	-0.003	(0.008)
b)Employer and other reasons	0.023*	(0.014)
Initial unemployment by year of labour market entry and sex	0.001	(0.002)
Unemployment rate by age group[55] and sex	0.001	(0.001)
%change in the employment share by occupation	-0.10 ***	(0.04)
%change in the supply of graduates by educational category and sex	0.35***	(0.106)
Initial Over-education	0.015**	(0.007)
N	**3646**	

Notes: Robust Standard errors in parentheses[56]; Significance denoted by: *** p<0.01, ** p<0.05, * p<0.1; Year dummies and individual means over age, years of work experience and marital status (Mundlak correction) are included but coefficients omitted from the table of results here).

[55] Given that the 16-19 age group was small and unemployment for that age group was significantly higher than for other groups a robustness check when employees below 20 years are not included in the sample was also run. Results from this regression are robust.

[56] The presence of macro-level independent variables means that standard errors are likely to be correlated based on the groups in which they fall in these macro variables. Even if not controlling for multiway clustering, taking the example of unemployment by age group and sex, it can be argued that, if it is assumed that the effect of sex or age themselves are common, e.g. there is common female effect for all women, then due to the fact that age and sex are included in the regression as explanatory variables, concerns about error correlation can be eliminated via the use of normal robust errors. This can be extended to all other macro variables used in this chapter.

As can be seen in the above table, initial unemployment at the start of paid employment by sex as well as the unemployment rate at the time of the survey by age group and sex do not appear to have a significant effect on the probability of over-education. On the other hand, the annualised change in the labour supply by educational category and sex, used to serve as an indication of the level of worker competition in the labour market, and the annualised change in the employment share by occupation and sex serving as a proxy of the labour market demand are both strongly significant and have the expected sign. More specifically, the probability of over-education increases by 0.35 percentage points as the labour supply increases by 1 percentage point from one year to the next while a 1 percentage point increase in the employment share by occupation appears to reduce the chances of over-education by 0.10 percentage points[57].

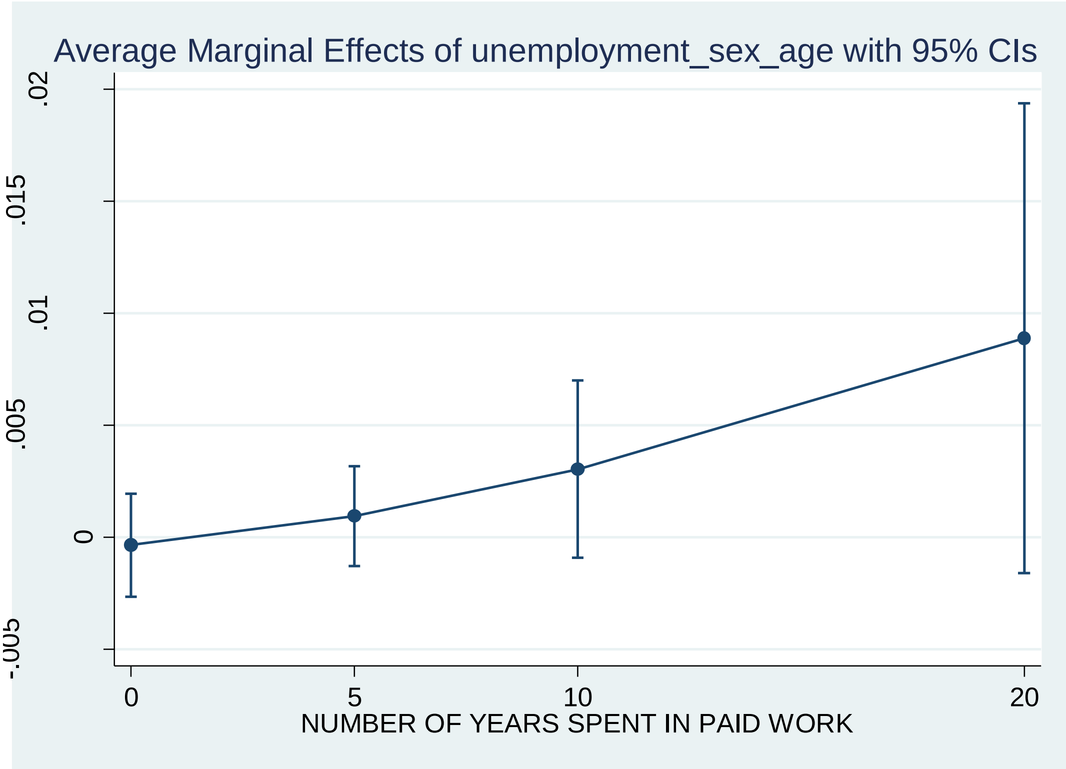
2.5.4 Micro-Macro Level Interactions

In order to examine whether adverse macro level conditions fall harder on new labour market entrants or in other words to observe whether work experience shelters workers in times of economic hardship, a number of micro-macro level interaction variables were entered in the pooled dynamic probit regressions[58] and are plotted below so as to offer an overall picture of the relations of interest.

The first interaction examined and presented in Figure 2.1 below is the interaction between the unemployment rate at the time of the survey by age group and sex and years of work experience[59].

[57] The results are almost identical when one macro variable at a time is entered in the regression.
[58] The preferred model specification from the earlier sections, i.e. the Wooldridge dynamic probit model with Mundlak corrections, could not be used in this section given that the Mundlak specification mirrors the FE specification and using interactions with variables that don't vary over time for individuals would not work in such a setting. For this reason the pooled probit model was the most appropriate specification to use in this case.
[59] One interaction at a time was entered in the regressions in this section, however, when both interactions are entered at once the results don’t change.

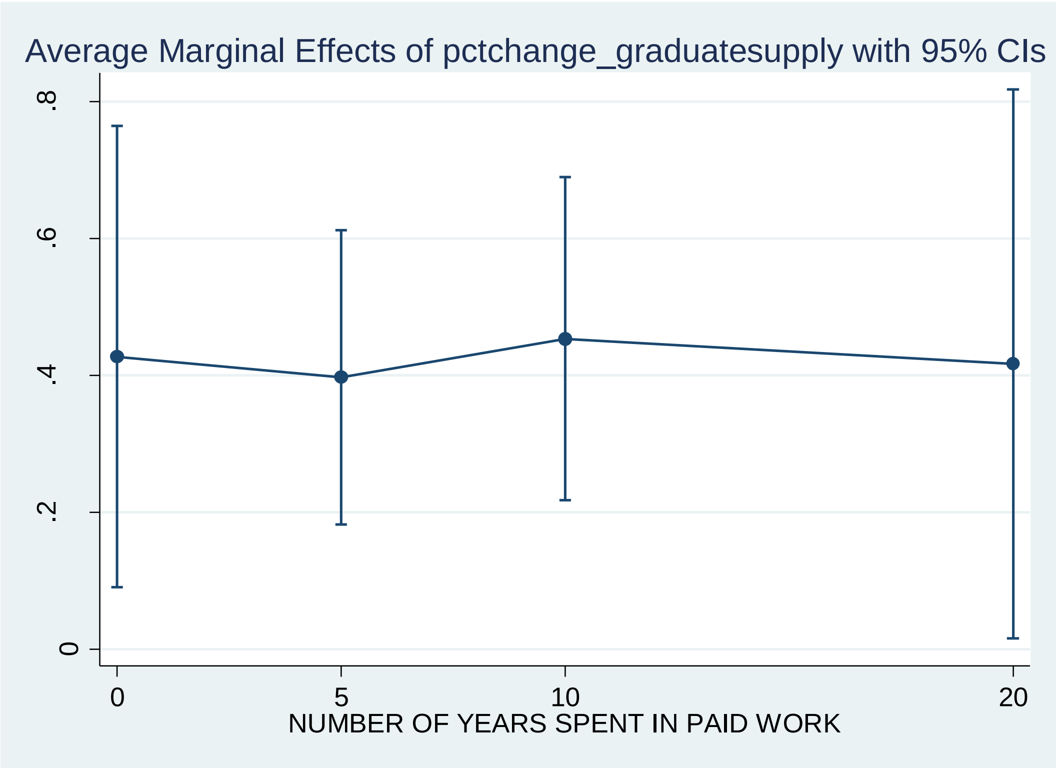
Figure 2.1: Interaction of the Unemployment Rate by Age Group and Sex with Work Experience



Even though the effect of the unemployment rate at the time of the survey is not statistically significant, its marginal effects at representative values of work experience demonstrate that as a person gains more years of work experience, the probability of over-education increases. In other words the negative relation between over-education and unemployment exacerbates as a person acquires more years of work experience and hence people with more years of work experience are more negatively affected by the unemployment rate than people with less. Even though strange at first, this finding might be explained by the fact that the more time that people spend in the labour market (and the older they become) the harder it is for them to escape over-education or to find a good match in times of high unemployment. In other words, at times of high unemployment for a specific age group and sex, the harder it becomes for a person to find a matched job. People with longer tenures might have a higher likelihood of being over-educated given that they may become trapped in mismatched jobs or they may become habituated in their current jobs and may not want to risk losing a safe mismatched job to compete for a matched job. This finding might have been different if the work experience variable here referred to job specific tenure rather than years of work experience in general. In that case, accumulation of job specific work experience might have been found to be decreasing the chances of a person being over-educated irrespective of the unemployment rate in the economy if within firm job progression served as a way to a matched job.

Lastly, Figure 2.2 below, demonstrates the marginal effects of the percentage change in labour supply by educational category at representative values of work experience.

Figure 2.2: Interaction between the Annualised Change in the Supply of Graduates by Educational Category and Sex with Work Experience



As can be seen from the confidence intervals of the margins plot in Figure 2.2 above, the effect of a percentage change in labour supply on the probability of over-education is the same for all work experience groups. Therefore, and contrary to what was expected, work experience does not provide protection against employee competition in the labour market and therefore the probability of over-education for people with 0, 5, 10 or 20 years of labour market experience is equally affected by a change in the labour supply.

2.6 Summary and Concluding Remarks

The present chapter employs panel data from the EU-SILC for the period 2005-2011 and examines the factors that affect the probability of being over-educated for one's job, i.e. of being in possession of a higher attained educational level than the modal level of education within ones' occupation. This is done using a range of static as well as dynamic pooled and random effects probit models as well as a dynamic econometric setting that controls for initial conditions and unobserved heterogeneity. It finds that factors such as being a female, married, working on a temporary contract basis, having recently entered the labour market from inactivity and having had a job change due to employer or other reasons during the previous year have a significant effect on the probability of over-education. In this way Chapter 2, contributes to the literature of the determinants of over-education in general and to the very limited over-education literature specific to the country of Cyprus in particular.

Furthermore, this chapter has answered calls for further examination of the dynamic properties of over-education and state dependence in over-education. It has done so by employing estimation methods that allowed the isolation of the main determinants of over-education within a methodological setting that deals with unobserved heterogeneity and state dependence, namely the Wooldridge (2005) dynamic probit model with Mundlak (1978) corrections. It has provided strong evidence suggesting that over-education is highly self-persistent. This means that over-education in one period causes over-education in the next period and hence workers who pass from this state run the risk of developing a long-run labour market disadvantage. In other words, over-education is likely to act as a trap rather as a stepping stone to matched employment. Moreover, results have also demonstrated that this is true for workers at all career stages meaning that being over-educated in one period will increase one's chances of being over-educated in the next period even as a respondent moves up the career ladder.

As an extension, the Wooldridge (2005) dynamic probit model with Mundlak (1978) corrections was re-run also incorporating a number of macro-level factors, answering calls in the literature for more research into the relation between overall macroeconomic conditions and the likelihood of over-education. The regression results demonstrated that apart from micro level determinants, country level indicators of both labour demand and labour supply characteristics that have not been examined in the past, also affect the chances of over-education. More specifically, the annualised change in the labour supply by education category and sex, entered in the regression as an indication of the level of competition faced by workers with the same education level, is found to significantly increase the chances of over-education as expected. In other words, as the number of people with an equivalent level of education increases, i.e. the labour supply at that level of education rises, then a person becomes more likely to be (and remain) over-educated as the labour market may not be ready to accommodate this surge in equivalently educated labour in jobs commensurate with their education. Moreover, on the demand side, the annualised change in the employment share by occupation and sex, serving as a proxy of the labour market demand, is found again as expected, to have a significant and negative relationship with over-education. For example, an increase in the employment share in a given occupation means that there are more employment opportunities within that occupation and hence more chances of finding work equivalent to ones' education. This translates into a lower probability of over-education. At the same time, other macro level determinants that can also be scarcely found in past over-education literature, namely the unemployment rate at the time of the survey and the unemployment level upon start of paid employment are not found to significantly affect the probability of over-education in this chapter. This means that macro variables directly linked to labour market demand and supply seem to be more predictive of the probability of over-education than measures of labour market slack such as the rate of unemployment in explaining over-education.

Lastly, an interaction term between the unemployment rate at the time of the survey by age group and sex and years of work experience has demonstrated that, as a person's work experience increases, his/her chances of over-education are more negatively affected by the unemployment rate, i.e. a person with more work experience will have more chances of being over-educated in times of higher unemployment than a person with less work experience. Moreover, an interaction between the annualised change in the supply of labour by educational category and sex and work experience has demonstrated that work experience does not seem to be able to shelter workers from over-education when the labour supply of equivalently educated workers increases. This means that accumulating more years of work experience does not seem to be able to protect workers from the effect of adverse macro conditions like an increase in the labour market competition when the supply of similarly educated workers increases on their likelihood of being over-educated.

What is more, a higher level of work experience seems to be exacerbating the negative effect of current unemployment on the likelihood of over-education.

The implications of the findings in the present chapter are manifold and these are discussed in detail in Chapter 5 of this thesis.

Appendix 2A: Data Cleaning and Preparation of Final Sample

The original micro data files provided by Eurostat include harmonised data from all participating countries. For this reason, all countries other than Cyprus were first dropped from the dataset. The EU-SILC is moreover released in four separate files: the household and the personal registers and the household and personal data files. All four files contain different variables but also a few common basic variables, for example sex, age, year of survey etc[60].

Step 1: For this reason, the different files for the same year/panel were first merged together into one file (the two household files together and the two personal files together and then the household and personal files together). Four different files each containing observations for four years with respondents that were either in their first, second, third or fourth years were left.

The EU-SILC longitudinal survey engages a rotational panel methodology which means that each year's files released by Eurostat contain information for respondents who are at different stages of the survey. For example, the 2008 file contains data collected from respondents who entered the survey in 2005, 2006, 2007 or 2008, therefore, depending on the year of entry of each respondent, there can be information corresponding to 1, 2, 3 or 4 survey years in the same file.

Step 2: For the analysis in this chapter and in order to be able to make use of the panel aspect of the longitudinal EU-SILC data and to observe transitions into or out of over-education and/or employment-unemployment, it was decided that respondents that have not yet completed the full 4-year duration of the panel are to be dropped. For example, in the case of the 2008 longitudinal file (2005-2008 panel), only those respondents who entered the sample in 2005 and were successfully followed up to 2008 are kept rather than keeping participants who are in their first, second, or third year of the survey. Similarly, in the 2009 User Database (UDB) file, that contains data on individuals who entered in any year between 2006 and 2009, only those respondents who entered in 2006 are kept and so on. In other words only those who entered the panel during the first year of each panel/UDB file were kept.

[60] The main difference between the register and data files is the source of the data collection-the register files mostly contain information in relation to the respondents collected from registers whereas the data files contain information directly collected from in-person interviews.

Appendix 2B: Adjustment of Variables that Change

a) **Self-defined Economic status**

In order to be able to append the different sub-panel files together, the new self-defined economic status variable (PL031) available from the last rotation of the 2009 file onwards, was recoded back to the original, PL030 variable, as per the below. In the table below, arrows demonstrate the categories of the PL031 variable that correspond to each of the original PL030 categories. For example, category 3 of the newer, PL031 variable, was recoded back to category 1 of the original, PL030 variable, as demonstrated by the first arrow in the below table

Self-defined Economic status (PL030): available for years up to 2008	**Self-defined Economic status (PL031): available from the last rotation of 2009 file onwards**
1 Working full time	1 Employee working full-time
2 Working part time	2 Employee working part time
3 Unemployed	3 Self-employed working full time (including family worker)
4 Pupil, student, further training, unpaid work experience	4 Self-employed working part-time (including family worker),
5 In retirement or in early retirement or has given up business	5 Unemployed
6 Permanently disabled or/and unfit to work	6 Pupil, student, further training, unpaid work experience
7 In compulsory military community or service	7 In retirement or in early retirement or has given up business
8 Fulfilling domestic tasks and care responsibilities	8 Permanently disabled or unfit to work
9 Other inactive person	9 In compulsory military community or service
	10 Fulfilling domestic tasks and care responsibilities
	11 Other inactive person

b) **Main activity each month of the year**

Similar to the above variable, the variables reporting an individual's main activity status each year of the survey change in 2009 and the arrows in the below table demonstrate the categories that were recoded back to the original variables.

Main Activity on January (PL210A)- Main activity on December (PL210L) up to 2008	**Main activity on January (PL211A) - Main activity on December (PL211L) from 2009 onwards.**
1 Employee(full time)	1 Employee working full time
2 Employee(part-time)	2 Employee working part time
3 Self-employed (full-time)	3 Self-employed working full time (including family worker)
4 Self Employed(part-time)	4 Self-employed working part time (including family worker)
5 Unemployed	5 Unemployed
6 Retired	6 Pupil, student, further training, unpaid work experience
7 Student	7 In retirement or in early retirement or has given up business
8 Other inactive	8 Permanently disabled or/and unfit to work
9 Compulsory Military Service.	9 In compulsory military or community or service
	10 Fulfilling domestic tasks and care responsibilities
	11 Other inactive person.

Appendix 2C: Small Cell Correction for the 2-digit ISCO-88 Classification of Occupations (PL050) Variable

The original occupation classification variable (PL050) found in the EU-SILC, reports the respondents' occupation, classified according to the 2 digit International Standard Classification of Occupations (ISCO-88). However, in order to ensure that all occupations have a satisfactory number of cases, a number of changes were made to the original PL050 variable and the new (reclassified) occupation classification variable was used for the derivation of the over-education dummy. The problematic occupational categories in terms of small cell size are presented in bold font in the left column of the table below and the reclassified variable categories appear in the right hand side column:

PL050:OCCUPATION (ISCO-88 (COM))	**Reclassified occupation variable used for the derivation of over-education**
1) 11 Legislators, senior officials and managers	1)Legislators, senior officials and managers, corporate managers, managers of small enterprises (**Regrouping of 11-13 of 2-digit ISCO-88 levels**)
2) 12 corporate managers	
3) 13 managers of small enterprises	
4) 21 physical, mathematical and engineering professionals	2) 21 Physical, mathematical and engineering professionals
5) 22 life science and health professionals	3) 22 Life science and health professionals
6) 23 teaching professionals	4) 23 Teaching professionals
7) 24 other professionals	5) 24 Other professionals
8) 31 physical and engineering science associate professionals	6) 31 Physical and engineering science associate professionals
9) 32 life science and health associate professionals	7) 32 Life science and health associate professionals
10) 33 Teaching associate professionals	
11) 34 other associate professionals	8) 34 Other associate professionals **plus the small cell of ISCO-88: 33: Teaching associate professionals**
12) 41 office clerks	9) 41 Office clerks
13) 42 customer services clerks	10) 42 Customer services clerks
14) 51 personal and protective services worker	11) 51 Personal and protective services workers
15) 52 models, salespersons and demonstrators	12) 52 Models, salespersons and demonstrators
16) 61 skilled agricultural and fishery worker	**Dropped:** this group was very small and could not be merged with any other category

	hence it was decided to drop it altogether especially given that it was referring to less educated individuals for whom over-education was unlikely to be as relevant
17) 71 extraction and building trades workers	13) 71 extraction and building trades workers
18) 72 metal, machinery and related trades workers	14) 72 Metal, machinery and related trades workers
19) 73 precision, handicraft, craft printing a	15) 73 Precision, handicraft, craft printing and related trades workers
20) 74 other craft and related trades workers	16) 74 Other craft and related trades workers
21) 81 stationary-plant and related operators	
22) 82 machine operators and assemblers	17) 82 Machine operators and assemblers **plus small cell ISCO-88:81: stationary-plant and related operators**
23) 83 drivers and mobile plant operators	18) 83 Drivers and mobile plant operators
24) 91 sales and services elementary occupations	19) 91 Sales and services elementary occupations
25) 92 agricultural, fishery and related labourers	
26) 93 labourers in mining, construction, manufacturing and transport	20) 93 Labourers in mining, construction, manufacturing and transport **plus small cell ISCO-88:92: agricultural, fishery and related labourers**

Appendix 2D: Description of Variables used in Chapter 2

1) Dependent Variable: Over-education dummy

For the derivation of the over-education variable the following variables from the EU-SILC were used:

a) Highest ISCED level attained (Variable PE040): this variable refers to the highest level of an educational program the person has successfully completed. The educational classification used is the International Standard Classification of Education (ISCED 1997) coded according to six ISCED-97 categories presented below:

   0 Pre-primary education

   1 Primary education

   2 Lower secondary education

   3 Upper secondary education

   4 Post-secondary non-tertiary education

   5 First stage of tertiary education (not leading to an advanced research qualification)

b) As per Appendix 2C, the new (reclassified) occupation classification variable was used for the derivation of the over-education dummy. This consists of the following categories:

1. ISCO-88: 1 Legislators, senior officials and managers, corporate managers, managers of small enterprises(regrouping of ISCO-88: 11-13 groups)
2. ISCO-88: 21 Physical, mathematical and engineering professionals
3. ISCO-88:22 Life science and health professionals
4. ISCO-88:23 Teaching professionals
5. ISCO-88:24 Other professionals
6. ISCO-88:31 Physical and engineering science associate professionals
7. ISCO-88:32 Life science and health associate professionals
8. ISCO-88:34 Other associate professionals plus the small cell of ISCO-88: 33: Teaching associate professionals
9. ISCO-88:41 Office clerks
10. ISCO-88:42 Customer services clerks

11. ISCO-88:51 Personal and protective services workers
12. ISCO-88:52 Models, salespersons and demonstrators
13. ISCO-88:71 Extraction and building trades workers
14. ISCO-88:72 Metal, machinery and related trades workers
15. ISCO-88:73 Precision, handicraft, craft printing and related trades workers
16. ISCO-88:74 Other craft and related trades workers
17. ISCO-88:82 Machine operators and assemblers plus small cell ISCO-88:81: Stationary-plant and related operators
18. ISCO-88:83 Drivers and mobile plant operators
19. ISCO-88:91 Sales and services elementary occupations
20. ISCO-88:93 Labourers in mining, construction, manufacturing and transport plus small cell ISCO-88:92 Agricultural, fishery and related labourers

2) Individual Level Independent Variables from the EU-SILC:

a) <u>Personal and job characteristics:</u>

Marital Status (PB190): The original marital status variable was divided into the following 4 categories: never married; married; separated; widowed and divorced. For the purposes of this chapter, it was reclassified into a binary dummy variable that takes the value of 1 if a respondent is married and 0 otherwise (single-not married), with 0 (single) becoming the reference category in the Chapter's regressions.

Sex (PB150): Dummy variable with 0 corresponding to men and 1 to women. This variable is introduced as a control variable in the probit models, where the omitted category is category 0 (male).

Age: The age of the respondents was calculated by deducting the Year of Birth (PB140) variable from the Year of Survey (PB010) variable.

Contract type (PL140): This is a binary variable with category 0 referring to permanent job/work contract of unlimited duration and category 1 referring to temporary job/work contract of limited duration. It is introduced in the probit regressions as a factor variable, where category 0 (permanent contract) serves as a reference category and is hence omitted.

Part-time/Full time: A dichotomous variable equal to 0 if someone works full time and equal to 1 if part-time.

Limitation in activities because of health problems (PH030): this variable, which is also referred to as Health limitation in the present chapter, is originally divided into 3 categories: 1 yes, strongly limited; 2 yes, limited and 3 no, not limited. For the purposes of the present chapter, it was reclassified into a binary variable with category 1 regrouping the first two categories of the original variable and zero corresponding to category 3 of the original PH030 variable.

Status in employment (PL040) (also referred to as **employment status** in this chapter): This variable refers to the current main job for people at work or the last main job for people who do not have a job and is divided into the following categories: 1 self-employed with employees; 2 self-employed without employees; 3 employees and 4 family workers. It was used to clean the data from people who are not employees since the over-education status of the self-employed and family workers cannot be as easily observed or reliably calculated/derived as in the case of employees.

Self-defined economic status (PL031) and (PL030): As discussed earlier, this variable asks respondents to define their current economic status and is used to derive the final sample by keeping only categories 1 (working full time) and 2 (working part time) for the purposes of examining the determinants of over-education. As mentioned above, this variable was replaced in 2009 but then recoded back to the original variable. The categories used in this chapter to define economic status are: 1 working full time, 2 working part time, 3 unemployed, 4 pupil, student, further training, unpaid work experience; 5 in retirement or in early retirement or has given up business; 6 permanently disabled or/and unfit to work; 7 In compulsory military community or service; 8 fulfilling domestic tasks and care responsibilities and 9 other inactive person.

Year of the survey (PB010): Apart from using this variable in the calculation of other variables, it was also introduced as a factor variable in the probit regressions as a year dummy controlling for anything not elsewhere controlled for that may be different for everyone from one year to the next.

b) <u>Employment Biography or work history variables:</u>

Number of years spent in paid work (PL200) renamed to **Years of Work Experience** in this chapter: This variable is used as a direct measure of the actual years of work experience of respondents. According to the Eurostat EU-SILC manual, this indicator provides a summary measure of the labour force experience of the individual. It reports the number of years, since starting the first regular job, that the person has spent at work, whether as an employee or self-employed.

Proportion of the past year spent in unemployment: In order to create the proportion of the past year spent in unemployment, the variables Main Activity on January to Main activity in December that have as a reference period the income reference year i.e. the previous year were used. The categories of these variables are: 1 Employee(full time); 2 Employee(part-time); 3 Self-employed (full-time); 4 Self Employed(part-time); 5 Unemployed; 6 Retired; 7 Student; 8 Other inactive and 9 Compulsory Military Service.

Over-education in period t-1: This variable is the lagged variable of the over-education dummy.

Initial Over-education: A new variable measuring whether an individual was over-educated or not in the first instance they were observed in the survey. Similar to the over-education dummy and its lag, the initial over-education variable is a dummy variable that takes the value of 1 if the respondent was over-educated in the first time period they were observed and the value of zero if they were not overeducated in t=1. Since the aim of the chapter is to understand over-education, what causes it and whether it is persistent or self-persistent, it was decided that observations of respondents who were unemployed, students, serving their military service or in other inactivity or were working as self-employed were assigned a value of zero in the initial over-education variable and were hence classified as not-overeducated.

Job Change: This variable brings together the **Change of Job since last year (PL160)** and the **Reason for Job Change (PL170)** variables. Its aim is to examine whether changing job and the reason for doing so differentially affect the probabilities of ending up in over-education compared to those who do not change job since the last interview (or in the past 12 months). This new variable has 3 categories: (1) no job change (for those who responded “No”-category 2- to PL160); (2) Self-induced job change (in order to take up or seek a better job) and (3) changed job because of employer and other reasons[61].

Below, the 2 variables mentioned above, based on which the Job Change variable was created are described:

Change of Job since last year (PL160): This is a Yes/No variable with the reference period being ‘’since last year’’ (or since last interview if applicable). According to the EU-SILC documentation, the aim of this variable is to report whether the individual left a job or changed from one job to another since the last interview (or last 12 months for the first year of data collection). For employees, a change of job means a change of employer, not moving from one set of duties to another with the same employer. Nevertheless, a change of contract with the same employer is considered as a change of job.

Reason for Change (PL170): This is a categorical variable that is asking those respondents who responded Yes (category 1) to the Change of Job since last year variable (PL160) above to choose a reason for doing so. The categories for the PL170 variable are: 1: to take up or seek better job (e.g. better wage, better work conditions, less commuting, etc.); 2: end of temporary contract; 3: Obliged to stop by employer (business closure, redundancy, early retirement, dismissal etc.); 4: Sale or closure of own /family business; 5: Child care or care for other dependent; 6: Partner’s job required us to move to another area or marriage; 7: Other reasons.

Recent Entry (into the labour market): This newly created variable is used to identify those respondents who have recently entered the labour market after having spent time in unemployment or in other inactivity. It was derived using the variable: “**Most recent change in the individual’s activity status (PL180)”** which reports whether there is a change in the individual’s activity status since the last interview (or last 12 months for the first year of data collection). **PL180** has the following categories: 1: employed-unemployed; 2: employed-retired; 3: employed-other inactive; 4: unemployed-employed; 5: unemployed-retired; 6: unemployed-other inactive; 7: retired-employed; 8: retired

[61] i.e. because of any of the following: 2: end of temporary contract; 3: Obliged to stop by employer (business closure, redundancy, early retirement, dismissal etc.); 4: Sale or closure of own /family business; 5: Child care and care for other dependent; 6: Partner’s job required us to move to another area or marriage; 7: Other reasons)

unemployed; 9: retired-other inactive; 10: other inactive-employed; 11: other inactive-unemployed; 12: other inactive-retired.

The new recent entry variable uses two of the categories of PL180 variable described above, namely categories 4 and 10, i.e. those categories that involve movement into rather than out of the labour market. This follows from the fact that over-education can only be observed for those who are currently in employment as all other observations are dropped before calculating over-education.

For the purposes of this variable anyone who has not moved into employment (since last interview) is assigned a value of zero, i.e. no recent entry. Moreover, any respondent with a missing value in the most recent change in activity status (PL180) variable is also assigned a value of zero. According to the flag variable of the PL180, missing values refer to those who report no change in their activity status since last year. Therefore, category 1(the reference category-no recent entry) of the recent entry newly created variable contains the observations of those who: a) had a missing value in the most recent change in activity status variable (i.e. no activity status change); b) those who changed from being employed to unemployed; c) employed to retired; d) employed to other inactive; e)unemployed to retired; f) unemployed to other inactive; g) retired to unemployed; h) retired to other inactive; i) other inactive to unemployed and j) other inactive to retired. Moreover, and even though respondents who moved from being retired to being employed are entering the labour market[62], they were not placed in category 3 of the recent entry variable but instead were added into category 1 (i.e. no recent entry) as they do not constitute a group of interest given their age as people up to 40 years old are kept in Chapter 2. Category 2 of this Recent entry variable refers to category 4 of the PL180 variable i.e. recent entry into the labour market from unemployment and category 3 refers to category 10 of the PL180 variable, i.e. recent entry into the labour market from other inactivity.

3) Macro-level independent variables:

Unemployment rate by Sex and Age Group: This variable, taken from the Labour Force Survey Report (years 2005-2011), refers to the rate of unemployment by sex and

[62] This was only a very small group (40 observations) that was going to create a problem if it was assigned a category on its own in terms of large standard errors and statistical significance. Also, more than half of these observations corresponded to people over the age of 65 who were going to be dropped from the analysis anyway.

by age group with age groups divided into the following categories: 1) 16-19; 2) 20-24 ; 3) 25-29 ; 4) 30-34; 5) 35-39 and 6)40-44 years of age.

Unemployment rate at labour market entry by Sex: This variable refers to the unemployment rate at the start of paid employment. For years 1980-2011, this rate is divided by sex, however for years 1976-1979 the statistic was not available by gender and hence the overall unemployment rate is instead used for that period. The year of labour market entry on which the unemployment rate was based is the following:

When begun first regular job (PL190): This variable reports the age at which the respondent first began working in a regular job. In order to find the year in which a respondent started work, this variable was added to the year of survey variable.

Annualised Change in Labour Supply by Educational Level and Sex: This variable is used as a proxy for the competition faced by employees in the labour market depending on their level of education. It is derived using the highest level of education attained for people over 20 years of age that is provided as a percentage of the population in the Statistics of Education Report of the Cyprus Office of Statistics. The annualised change in the supply of graduates in each survey year is calculated as the percentage change from the previous year according to 4 educational categories: 1) never attended school, 2) primary education, 3) secondary education and 4) tertiary education and according to sex.

Annualised Change in the Employment Share by Occupation and Sex: The aim of this variable is to serve as an indication of the number of people employed and hence demanded in each occupation and to examine whether changes in the employment share by occupation (and sex) affect the chances of a person in that occupation to become and/or to remain over-educated. In order to create this variable, the employment level (in thousands) of the employees by sex and occupation (defined by 1-digit ISCO classifications) was first entered and used to get the shares of employment by sex and occupation as a proportion of the total employment (irrespective of occupation) in the economy and then multiplied by 100 to get the percentage. In this way, the share of the total number of employees in the economy employed in each occupation by sex is derived[63].

[63] The final variable i.e. "annualised change in the employment share by occupation and sex" refers to the % change from one year to the next as opposed to the percentage point change. For example, if the employment share in occupation A increased from 60% to 63% from Year 1 to Year 2, then the % change reported in this variable would be the ratio i.e. 1.05, and 1.05 - 1 = 0.05, which (when multiplied by 100) is the percent increase in the value across time.

Appendix 2E: Determinants and State Dependence of Over- Education[64] by Career Stage [65]

	Career Stage 1 (0-3 years of work experience)	Career Stage 2 (4-9 years of work experience)	Career Stage 3 (10-20 years of work experience)
Age	-0.02* (0.013)	0.02 (0.02)	0.002 (0.00)
Work Experience	0.03* (0.018)	-0.01(0.01)	-0.001 (0.001)
Married	-0.07* (0.04)	-0.08*(0.04)	0(0)
Female	0.03**(0.01)	0.01(0.01)	0(0)
Health limitation	-0.03(0.05)	0.00(0.01)	0(0)
TemporaryContract	0.009(0.02)	0.01(0.01)	0.001(0.003)
Part time	0.057(0.04)	-0.00(0.02)	0.01(0.02)
Past Year unemployment	-0.17**(0.09)	0.002(0.04)	0.001(0.002)
Overeducation at t-1	0.88***(0.04)	0.90*** (0.05)	0.97**(0.04)
Job Change (No job change)			
a) Self-induced	0.00 (0.02)	-0.01 (0.02)	0.00(0.00)
b)Employer and other reasons	-0.035 (0.03)	0.05* (0.03)	0.01(0.02)
Year Dummies(2005 and 2006)			
2007	0.06**(0.03)	-0.07 (0.06)	-0.004(0.008)
2008	0.04 (0.03)	-0.05 (0.05)	-0.004(0.008)
2009	0.03 (0.03)	-0.06 (0.06)	-0.005(0.01)
2010	0.05* (0.03)	-0.035 (0.05)	-.005(0.01)
2011	0.06* (0.04)	-0.06 (0.06)	-0.01(0.01)
Initial Over-education	0.01(0.02)	0.02* (0.02)	0.01(0.01)

Notes: Standard errors in parentheses; Significance denoted by: *** p<0.01, ** p<0.05, * p<0.1; Standard errors clustered around personal ID; Individual means over age, years of work experience and marital status as per the Mundlak correction are included.

[64] Results from random effects probit model with Mundlak correction and initial condition correction and clustering around personal ID. Recent entry into the labour market variable omitted in order for all the group regressions to converge. It was not possible to get any results for Career Stage 4 (more than 20 years of work experience) due to the small sample size of this group.
[65] Results are robust to the choice of cut-offs between the work experience groups

Chapter 3: Over-Education and Employee Withdrawal Behaviours: Examining the Relation between Over-Education and On-The-Job Search.

3.1 Introductory Remarks

3.1.1 Introduction

Chapter 2 has demonstrated that over-education is a permanent phenomenon in workers' lives that is also state dependent in the sense that this year's over-education causes or makes next year's over-education more likely irrespective of the career stage one is in. As argued in Chapter 2, individuals may view a job for which they are over-educated as a stepping stone to a matched job, by for example gaining work experience or a temporary income while engaging in on-the-job search. It follows that those who are over-educated are expected to search more on-the-job compared to their well matched counterparts. According to Bjelland et al. (2011), job-to-job movements account for between one-third and one-half of all labour market movements in the U.S and on-the-job search can act as a prediction of such voluntary turnover. This means that on-the-job search can act as a correction mechanism leading to an exit from over-education and can therefore have a vital role to play towards an efficient labour market making the study of the relation between the two imperative.

On the other hand, job search behaviour absorbs time and energy that might be put to other uses and is hence costly irrespective of whether it results in turnover or not (March and Simon, 1958). Moreover, according to Locke (1976), job search can provoke psychological processes that encourage withdrawal behaviour and decrease job and organisational commitment. Therefore, on-the-job search signals empoyee commitment and it is therefore important as a process on its own rather than just due to its connection with job mobility (Wald, 2005). It follows that it can be very expensive both for the firm and the economy as a whole and if not successful, for the worker and hence it is important to examine its determinants.

3.1.2 Chapter Objectives

The results from the previous chapter have shown that using a job for which one is over-educated as a stepping stone to a matched job could result in them being trapped in over-education and hence developing a long-term labour market disadvantage. The literature on the relationship between mismatch and on-the-job search is limited and the connection between the two has received little empirical attention. To my knowledge, the existing research exploring the connection between mismatch and job search relies only on Australian (Mavromaras et al. 2013; McGuinness and Wooden 2007), Canadian (Wald, 2005) and European (Allen and van der Velden 2001; Wolbers 2003; Congregado et al. 2016) data sets, however Cyprus is not included in the existing European studies. The present chapter aims to enrich this literature and to provide evidence for Cyprus, where the phenomenon has never been examined.

More specifically, this chapter will examine the determinants of on-the-job search while paying closer attention to whether those over-educated for their jobs are prone to heightened job search implying lower organisational commitment and lower productivity. Moreover, via the investigation of the relation between on-the-job search and over-education, light will also be shed on the question of whether over-education is voluntary, which has often been debated, but there is little direct evidence to answer it either way. If the over-educated are no more likely to search for jobs than the non-overeducated, this would suggest that they have accepted the lower level job for other reasons such as family or mobility reasons or that they do not really perceive themselves as over-educated. Such a finding could also explain the results in Chapter 2 as it could be one reason why people do not leave over-education very often, if they are not actively searching to do so. On the other hand, if those over-educated are found to be searching more on-the-job but over-education is persistent, as found in Chapter 2, this could be because the labour market lacks the flexibility of allowing over-educated workers to find matched jobs therefore creating an over-education ‘trap’, meaning that policies should be directed towards allowing more flexibility in escaping over-education or preventing over-education from happening in the first place.

As also pointed out by Mavromaras et al. (2013), even though there is a large body of evidence connecting labour market mismatch to adverse labour market outcomes, there is a lack of empirical evidence on whether there is a causal association between the two.

The analysis in Chapter 2 of the determinants of over-education has demonstrated that over-education is not random and is affected by other things such as personal and job characteristics. Hence over-education could be argued to be endogenous, if these same characteristics also influence the likelihood of job search. In order to eliminate the possibility that these factors affecting the probability of over-education also affect the probability of engaging in on-the-job search, causing the over-education coefficients to be biased, one of the macro variables found to significantly affect the probability of over-education in Chapter 2 will be used as an Instrumental Variable (IV) for over-education. The probit/Ordinary Least Squares (OLS) results will then be compared to the IV regression results in order to see whether treating over-education as endogenous makes a difference to the results. Apart from examining the determinants of on-the-job search irrespective of the reason for searching on-the-job, this chapter will also replicate the analysis using one of the reasons of looking for another job, namely 'because of the wish to have better working conditions (e.g. pay, working or travel time, quality of work)' as a dependent variable to see if the results change.

The analysis performed for Cyprus will then be replicated for the UK and Germany. The UK is considered a flexible labour market while Germany has a more rigid labour market and this offers the opportunity to compare results across three countries with different labour market characteristics that may in turn explain differences (or similarities) in the relationship between mismatch and on-the-job search.

The present chapter will therefore make a twofold contribution to the literature. The first contribution will be to add to the literature on the determinants of on-the-job search, especially in the case of Cyprus where the phenomenon has not been examined[66]. The second contribution of this chapter will be to closely examine the connection between over-education and on-the-job search controlling for the potential endogeneity of over-education using IV regression, a methodology that, to my knowledge, has not been employed in the past to shed light on the causal relation between the two.

The structure of the chapter is as follows: Section 3.2 lays out the theories connecting over-education and on-the-job search and summarises the empirical literature, Section 3.3 describes the data and characteristics of the final sample and provides summary statistics

[66] On-the-job search in the present chapter is treated as a uniform activity, and no attempt is made to identify different types of job search, as they are not measured in the data set used.

for the variables used in the regressions, Section 3.4 sets out the methodology used, Section 3.5 presents and discusses the results and Section 3.6 summarises and concludes.

3.2 Literature Review

3.2.1 Job Search Theory

Theoretical models of job search postulate that workers engage in job search if the marginal return of searching is above its marginal cost and hence the search decisions are the outcome of an optimal search strategy. In other words, the decision to search for another job depends on the optimisation decision between the utility received from the current employment and the expected utility in another job over any search costs incurred (direct search costs and opportunity costs of searching related for example to leisure).

Burdett (1978) was the first to incorporate the possibility of job search while already employed into a theoretical model of job search and quit rates. More specifically, Burdett (1978) creates a simple partial equilibrium reservation wage model in an attempt to explain the negative relation that age and tenure have with quit rates. Burdett and Mortensen (1998) then moved this on-the-job theory model into a general equilibrium framework.

Moreover, Pissarides (1994) incorporates on-the-job search into the matching model of unemployment and concludes that the model can explain on-the-job search if the assumption that workers acquire job-specific skills is made. The implication of the aforementioned assumption is that on-the-job search, and hence resigning, happens at short job tenures because the accumulation of job-specific capital makes a worker more productive in his/her current job over time and hence increases his/her wage, causing him/her to give up further search.

3.2.2 Theories Connecting Over-Education and Job Search

Matching Theory

In matching theories of job search (Jovanovic, 1979), over-education is a sign of a poor job match, and over-educated workers are expected to engage in repeated job search activity so as to seek and achieve better matches over time. In other words, school leavers engage in job search until they find an optimal job match (Jovanovic 1979; Tuma 1985). This translates into higher levels of job search for those workers who are mismatched in their jobs compared to their well-matched counterparts. Hence, according to this theory mismatches are temporary phenomena stemming from imperfect information and job search costs (Jovanovic, 1979) and are gradually eliminated via the provision of increased labour market information and heightened job search. It is therefore often found that, as implied by this theory, job search decreases as job tenure increases as the worker-employer match is increasingly likely to improve and uncertainty about the quality of the match is resolved (Jovanovic, 1979).

Career Mobility Theory[67]

As discussed in the previous chapter, in theories of career mobility (Rosen 1972; Sicherman and Galor 1990), workers may intentionally enter their preferred profession at a level below their qualifications so as to obtain the required skills that will allow them to achieve faster career progression in the future, via on-the-job training and learning. Sicherman and Galor (1990) argue that when the expectations of mismatched workers for promotion are not met, they are more likely to quit. In terms of the prediction of career mobility for on-the job search, it is expected that there will be a negative relation between over-education and on-the job search if workers decide to stay in a job for which they are mismatched so as to gain work experience and skills that will offer them a greater career advancement within the firm. On the other hand, if promotions within the firm are not perceived as imminent, on-the-job search is expected to increase and hence the relationship between over-education and on-the-job search is then expected to be positive.

Similarly, in the theoretical literature, Dolado et al. (2002), consider a matching model of heterogenous workers and jobs which includes on the-job-search with high educated workers temporarily accepting unskilled jobs if they are paid a wage above their outside option value while continuing to search for skilled jobs. This results in over-education as workers' attained education exceeds the skill requirements of jobs while at the same time,

[67] This theory is also discussed in detail in Chapter 2 of this thesis

they engage in on-the-job search to achieve a better job match between their education and job skill requirements.

Frank's Theory of Differential Over-Qualification

Frank's theory of differential over-qualification postulates that married women are more likely to be over-educated for their jobs due to geographical constraints resulting from their husbands' career/job choices (Frank, 1978). Frank's theory therefore suggests that marriage constrains a woman's job search as they have to 'compromise and accept something less than the best offer' (Frank 1978, 361). It follows that married women who are over-educated search less on-the-job than similarly over-educated single women.

According to Wald (2005), the nature of the geographic constraints such as labour market size, commuting distance and time, which are important causes of the binding nature of geographical constraints are difficult to be precisely captured as they seldom appear in labour force data sets noting that due to such data issues this theory has only attracted narrow empirical testing and mixed support (McGoldrick and Robst 1996; Buchel and Battu 2003). Due to data constraints, it is also difficult to provide an accurate test of this theory in the present chapter. Moreover, given the small travelling distances between cities in Cyprus, Frank's theory is not expected to play a key role in explaining the relation between over-education and on-the-job search. Lastly, the line between the primary and secondary income earner in the household is increasingly becoming less clear cut.

Summing up the three theories, while job mismatches are viewed as sub-optimal outcomes from the worker's perspective according to matching theory, mismatches are considered to be in accordance with optimising behaviour in the case of Sicherman's (1991) career mobility hypothesis and Frank's (1978) theory of differential over-qualification (Wald, 2005). Furthermore, unlike job matching theory, the last two theories described above, predict that workers can be hired for positions in which they are apparently over-qualified even in the presence of perfect information (Hersch, 1995).

3.2.3 Review of the Empirical Literature

Pissarides and Wadsworth (1994) use the Labour Force Survey (LFS) to provide empirical evidence on on-the-job search for Britain and find that job characteristics are the ones that mostly affect the probability of on-the-job search. More specifically, they find that temporary or part-time work have a positive effect on on-the-job search only for men and that job tenure is a significant determinant of job match quality and hence another factor affecting the likelihood of engaging in on-the-job search. They also find that age is inversly related to job search and that skilled workers search more than the unskilled. Inter-industry wage relativities are another determinant of on-the-job search confirmed by their empirical investigation.

In an examination of the relationship between over-qualification and job search behaviour, Wald (2005) uses maximum likelihood probit estimation on a sample of employed Canadians aged 18 and over who were surveyed in 2000. More specifically, Wald (2005) empirically examines the claim that over-qualified workers are active job searchers meaning that they are less committed and evaluates the three theories explaining over-qualification in which job search plays a key role namely matching theory, the theory of differential overqualification, and the career mobility hypothesis. He finds that over-qualified workers are indeed more likely to engage in job search, a finding that provides support for the matching theory which views over-qualification as sub-optimal from the worker's perspective. According to Wald (2005), his findings point to the fact that the best strategy for employers so as to safeguard high organisational commitment and possibly avoid costly turnover is to turn down over-qualified job applicants (Wald, 2005). He also finds that a high level of employee satisfaction, a feeling that one's job is interesting and that he/she is being treated fairly by employers and good prospects of internal advancement reduce job search while job search increases when employees feel a low workplace morale.

Moreover, Di Pietro and Urwin (2006), using Italian graduate cross sectional data, find that over-educated workers are more likely to be engaged in on-the-job search than well-matched workers and that this is also true for under-educated workers though to a lesser extent. They also find that education rather than skill mismatch seems to have a larger impact on the probability of job search and provides an improved fit to their binomial logistic regression compared to when skill mismatch is entered in the equation instead.

At an EU level, Wolbers (2003) uses a logistic regression to analyse the probability of looking for another job in a group of European countries using LFS cross sectional data. He defines mismatch with regard to field of education rather than the level of education and finds that school-leavers with a non-matching job look for another job more frequently than those with a matching job. He moreover finds that the effect of job mismatch on the likelihood of looking for another job is smaller in countries with a high share of school-based vocational education compared with countries where this share is low.

More recently, DeLoach and Kurt (2018) examining factors that affect the decision to enage in on-the-job search using pooled cross sectional data for the US between 2003 and 2016, find strong evidence that the probability of search is negatively related to worker productivity in general. They moreover find that mismatched workers seem more likely to engage in search compared to non-mismatched workers and that high-productivity mismatched workers are more likely to search than other mismatched workers. The authors in this paper note that a downside of studying on-the-job search using cross sectional data sets is the fact that only a small fraction of workers report on-the-job search at a specific point in time, i.e. the time of the survey. They argue that this means that the analysis of the determinants of on-the-job search is made based on a small sample of workers whereas this number is much greater in the real economy (DeLoach and Kurt, 2018). In other words, even if a small proportion of the sample undertakes on-the-job search at the time of the survey, the extent of on-the-job search in the economy is expected to be much greater given that workers that for example are not looking now have searched on-the-job at some point in their careers. In order to overcome the above issue, panel data is required. DeLoach and Kurt (2018) also point out that they are unable to control for unobserved heterogeneity that is likely to bias the effect for mismatched workers downwards. They state that, unlike as assumed in theory, mismatched workers are usually not exogenously mismatched with unobserved worker heterogeneity likely to be simultaneously correlated with search and the likelihood of being mismatched. They give the examples of lack of ambition, mobility and risk aversion and explain that for example, due to being risk-averse a worker may not search for a new job while at the same time he/she is more likely to become mismatched causing the coefficient on mismatch to be downwardly biased. On the other hand, they argue that a more ambitious or a more mobile worker is more likely to engage in on-the-job search while at the same time being less likely to be mismatched. This issue is tacked in the present chapter via IV regressions which eliminate such potential endogenity concerns even if the data set used is not longitudinal in nature hence contributing to the closing of this gap in the literature..

In terms of the empirical literature that uses panel rather than cross sectional data, Allen and van der Velden (2001) use a data-set with a longitudinal component to examine a cohort of Dutch graduates from 1990-91 in their first job after graduation and five years later to test the assignment theory of over-education by examining the relation between education and skill mismatches. Skill mismatch in this paper is based on the responses of workers to questions about their perceptions as to the level of utilisation of their skills and knowledge in their current jobs. In order to examine whether people whose skills are not matched to their job are motivated to quit their job in favour of another, Allen and van der Velden (2001) use a logistic regression to examine whether mismatched workers looked for other work in the past four weeks. They find that on-the-job search is subject to the variety of work tasks and the possibility to introduce own ideas, and that educational mismathes do not increase the chances of on-the-job search of workers who are also skilled mismatched. More specifically, they find that over-educated graduates are no more likely to search on-the-job compared to their well-matched counterparts and that over-skilled workers are the ones more likely to engage in on-the-job search compared to those who are well matched in terms of skills. They moreover find that underutilisation of skills in a job increases on-the-job search behaviour. All in all, their results establish that skill rather than educational mismatches are the ones that have behavioural consequences as only skill mismatches show a significant effect on on-the-job search.

McGuinness and Wooden (2007) use longitudinal data from Australia to test whether over-skilled workers are more likely to express a greater desire to quit their current job as well as a greater confidence in relation to their future employment prospects compared to well-matched workers. They also examine whether these intentions lead to higher rates of voluntary quits among over-skilled workers. Lastly, they test whether job search activities and the subsequent job separations end-up in an improved match. They find that, despite the fact that over-skilled workers have a greater desire to quit their job and are much more job mobile than matched workers, they are not any more confident than matched workers about their prospects of securing superior or even comparable jobs. More specifically, they find that some of the greater mobility observed among over-skilled workers is caused by involuntary job separations and that even when job separations are voluntary, the majority of job changes do not end in improved skills matches. According to the authors, the majority of mismatched workers either stay in jobs where their skills are not sufficiently utilised or exit the workforce completely. Finally, they note that even though findings of greater mobility among over-skilled workers are frequently taken as evidence in support of matching and/or career mobility theories, the results of their paper suggest that such strong inferences are not justified.

Similarly, Mavromaras et al. (2013) use the panel element of the Household Income and Labour Dynamics in Australia (HILDA) survey to examine the effect of mismatch on mobility and the stability of employment. More specifically, Mavromaras et al. (2013) estimate the probability of job change between two consecutive interviews depending on an individual's level of mismatch in the job that they left so as to examine if the likelihood of quitting or being laid off is higher for mismatched employees compared to their well-matched counterparts. Mavromaras et al. (2013) fail to find a significant direct effect of mismatch on mobility. More specifically, after controlling for individual unobserved heterogeneity, neither of the three categories of mismatch they define (only over-educated, only over-skilled or both over-educated and over-skilled) are found to significantly affect involuntary job mobility and it is just over-education on its own or together with over-skilling that increases voluntary mobility, nevertheless only for males. The authors note that many of the statistically significant estimates of mismatch derived from pooled probit models lose their significance once panel estimation is employed suggesting that some of that significance was due to unobserved heterogeneity bias and that previous results based on cross sectional and short panel data studies are prone to significant biases stemming from the lack of controls for unobserved heterogeneity. However, they conclude that given the fact that they do not control for employer-specific unobserved heterogeneity, which is anticipated to be relevant in the case of layoffs, the issue of job mobility and mismatch remains uncertain notwithstanding the above evidence. As mentioned earlier, the present chapter, even if it uses cross sectional data, addresses the first concern raised by the authors above by employing an IV regression methodology to reveal the causal effect of mismatch (in this case over-education) on the outcome of interest (in this case on-the-job search which can provide an indication of the intention to quit and of voluntary job mobility which was of interest in the above study).

Lastly, Congregado et al. (2016), analyse the incidence, effects, dynamics and routes out of over-qualification while incorporating distinctions in employment status. They base their measure of over-qualification on self-reported (i.e. subjective) information and using data from the European Community Household Panel for the EU-15 find, among other findings, that over-qualification is a permanent phenomenon and demonstrate that effective routes out of over-qualification vary according to employment status. They test the hypothesis that over-qualified workers are more likely to display withdrawal behaviours, such as absenteeism, on-the-job search and turnover than non over-qualified workers. In terms of on-the-job search, they use four different binary logit specifications (for all workers, self-employed, private employees and public employees) and find that, consistent with their beformentioned hypothesis as well as with previous studies, over-qualification has a real effect on on-the-job search. In particular, they find that workers who feel over-qualified are approximately 72% more likely to look for alternative jobs compared to those who do not feel over-qualified and that only small differences arise when separately exploring these effects by employment status. More specifically, they find that the highest propensity to search on-the-job is exhibited by private sector employees and the lowest by the self-employed. Lastly, discussing other determinants of on-the-job search they find that females, cohabiting individuals and those reporting good health are less likely to engage in on-the-job search while age has a non-linear, inverted U-shaped relationship. Moreover, higher education levels are found to have a positive and statistically significant effect for public and private workers but a negative effect for the self-employed while attending an education or training programme is found to have a positive effect on engaging in on-the-job search for all workers. Lastly, job search probabilities are found to be negatively affected by earnings and job tenure for all workers.

Following the discussion in this section, there seems to be a lack of consensus within the empirical literature as to the relation between mismatch and on-the-job search and this chapter attempts to close this gap. More specifically, the present chapter contributes to the above literature by examining the relationship between over-education and on-the-job search using a different methodology to the ones previously used in the on-the-job search literature. Via the use of IV regression, the possibility that the positive relationship between mismatch and on-the-job search, established in the literature, is due to the endogeneity of over-education will be tested. In other words, and as also discussed in the previous paragraphs, omitted unobserved heterogeneity rather than genuine over-education could be causing the often-found higher probability of being actively engaged in on-the-job search for the over-educated. This chapter addresses this issue and hence answers calls in the literature as to the need to disentangle the causal effect of over-education on on-the-job search. Moreover, the present chapter will add evidence from Cyprus to the existing literature of on-the job search and its relation to over-education.

3.3. Data and Descriptive Statistics

3.3.1 Data

Labour Force Survey

The European Labour Force Survey yearly cross sectional Cyprus files (CY-LFS) drawn from the EU-LFS for the period 2000-2015 are the main data source for this chapter. The CY-LFS quarterly survey on employment, unemployment and inactivity, covers private households and excludes individuals studying or working abroad. In each quarter, approximately 3,600 households are interviewed from all districts of Cyprus (Labour Force Survey, 2010). Personal and telephone interviews as well as computer input are employed for the data collection both at a household and at the individual level. The survey for Cyprus began in 1999 as a yearly survey and then gradually became a rolling survey with each household/person interviewed for 6 quarters but no possibility to connect observations from the same person. In order to guarantee consistency and to ensure that each observation in the analysis of the present chapter corresponds to a different individual, only the first interview wave for each person is used. The CY-LFS anonymised micro data contain information on current and past labour-force characteristics of individuals, personal details such as age, education, gender and region as well as occupation (ISCO) and industry (NACE) information.

3.3.2 Sample

The final sample of the present chapter comprises of people aged up to 64 years old who work as employees. As in Chapter 2, due to the complexities in the calculation of a reliable modal education level and hence over-education status in the case of those in self-employment, only people working as employees are kept in the sample. As mentioned in the literature review section, papers that have analysed on-the-job search by employment status have not found a significant difference between the two and this offers additional support to the choice of dropping the self-employed here. It also leaves us with a more homogeneous group of workers. Family workers are also dropped for similar reasons. Moreover, people below 20 years old or above 64 years of age as well as people not in employment are excluded from the analysis. The total number of observations following these adjustments is 32,252 and it is a pooled sample covering the period 2000-2015.

3.3.3 Descriptive Statistics

Table 3.1 below presents the means and standard deviations (SD) of all the variables used in the regressions.

Table 3.1: Descriptive and Summary Statistics

Variables Used in Regression Analysis	**Mean**	**SD**
On-the-Job Search for another Job	0.035	0.183
Over-education	0.289	0.453
Age Group		
25-29	0.142	0.349
30-34	0.148	0.355
35-39	0.141	0.348
40-44	0.135	0.342
45-49	0.125	0.331
50-54	0.107	0.309
55-59	0.078	0.268
60-64	0.037	0.189
Job Tenure (in months)	111.13	112.056
Married	0.699	0.459
Female	0.509	0.500
Temporary contracts	0.142	0.349
Part time work	0.056	0.231
Industry (NACE Rev 1-1 digit)		
Mining and quarrying	0.002	0.049
Manufacturing	0.094	0.291
Electricity, gas and water supply	0.015	0.120
Construction	0.097	0.296
Wholesale and retail trade	0.169	0.375
Hotels and restaurants	0.080	0.272
Transport, storage and communications	0.064	0.245
Financial intermediation	0.062	0.241
Real estate, renting and business activities	0.072	0.259
Public Administration and defence; compulsory social security	0.096	0.295
Education	0.084	0.277
Health and social work	0.045	0.208
Other community, social and personal services activities	0.036	0.184
Activities of private households	0.063	0.243
Extraterritorial organizations and bodies	0.007	0.086
Size of the firm		
11-19 persons	0.115	0.319
20-49 persons	0.147	0.354
50 persons or more	0.340	0.474
Do not know but less than 11 persons	0.007	0.085
Do not know but more than 10 persons	0.015	0.120
Macro level Variable used as Instrument		
Annualised change in the supply of graduates by educational category and sex	0.011	0.044

Firstly, job search is expected to depend on a number of demographic characteristcs. For example, gender is likely to play a role in on-the-job search. More specifically, according to Blau and Kahn (1981); Parsons (1991); Van Ophem (1991) and Keith and McWilliams

(1999), even though voluntary job mobility is a more probable exit route from employment for women, on-the-job search aiming at voluntary job mobility is more probable among men.

Moreover, it is expected that on-the-job search will be more probable amongst younger workers as well as among people with a lower job tenure (e.g. Campbell 1997). According to Wald (2005), the potential gain from on-the-job search is higher the younger the worker as they have more years left in the labour market and hence are expected to engage more in job seach. In a similar manner, as workers' job tenure increases, their specific job experience increases, potentially leading to promotion and higher earnings or a higher bargaining power and hence fewer reasons to search for another job. In other words, as on-the-job learning and the acquisition of job-secific skills which accumulate with tenure result in higher productivity and a higher wage if workers are paid according to their marginal products, outside jobs become less competitive (Parsons 1972; Jovanovic 1979). Moreover, according to Mortensen (1975), a number of firm characteristics only become known to the worker after they become employed and hence a worker may decide to quit if these characterisitcs, once learned, make the job unaccetable to them. Given that age and tenure are positively correlated, age is also expected to have a similar effect as tenure (Burdett, 1978).

Finally, marriage can be viewed as a constraint to job-search as partner considerations such as location of partner's job, working hours etc may affect the decision to search for another job and being unmarried has been linked with higher on-the-job search (e.g. Royalty, 1998). Similarly, being a female is also interacted with being married in this chapter's regressions as married females might act as tied stayers in the case where the male is the primary income earner of the household, in the same fashion discussed in Chapter 2 for over-education, but in this case applied to on-the-job search.

A number of job characteristics are also expected to have an effect on the probability of engaging in on-the-job search. Firstly, working on a part-time basis is expected to give rise to heightened job-search. Secondly, if a person has a temporary job or a contract of limited duration this is expected to cause an increased job search in the hope of a more stable and more permanent working arrangement. Firm size is expected to have a negative relation with on-the-job search as larger firms offer more opportunities for within firm progression and promotion and hence the need to search for outside options while employed is expected to be lower than in the case of smaller firms where progression opportunities within the firm are more limited. Tenure as discussed in the previous paragraph is another of the job characteristic examined[68].

Lastly, a number of controls, such as year dummies to control for all things happening in the economy that may affect the relationships being examined, are entered in all regressions. Moreover, the industry within which a person works is included as a control in the on-the-job search equation to control for any inter-industry variation, i.e. for any industry-specific effect that may affect all workers within an industry to the same extent.

3.4 Methodology and Model Specifications

3.4.1 Binary Choice Pooled Probit Model

In order to examine the determinants of on-the-job search, a binary response probit model is estimated. The dependent variable in the regression is a dummy variable equal to 1 if a person is looking for another job and zero otherwise. As discussed in the previous section, apart from the over-education dummy that is equal to 1 if a person is over-educated and equal to zero otherwise and of which the significance and coefficient are of primary interest in the job search equation, a number of other variables considered as possible determinants of on-the-job search are included in the regressions. These are split into demographic characteristics (gender, age, and marital status), job characteristics (part time work, temporary contract, job tenure, and firm size) and other control variables (year and industry dummies).

As mentioned in Chapter 2, probit models are commonly motivated as latent variable models where the (latent) variable/outcome of interest, cannot be observed or measured. The underlying latent variable in this case is the ‘propensity to engage in job search’ but what is observed is just the yes/no indicator of whether respondents do engage in on-the-job search, depending on whether the latent variable is above or below some cut-off.

The statistical specification of this binary choice model is given by the following expression:

---

[68] The data set used in the present chapter does not incorporate distinctions between private and public level employees and hence it is not possible to test which of the two groups is more likely to engage in on-the-job search.

$$\mathbf{Pr\ (JS = 1)\ =\ F\ (x\beta)\ =\ \Phi(x\beta)} \qquad \textbf{(3.1)}$$

Specifically, it is assumed that the model takes the form where Pr denotes probability, and Φ is the Cumulative Distribution Function (CDF) of the standard normal distribution. The parameters β are estimated by maximum likelihood[69].

3.4.2 Instrumental Variable (IV) Regressions

In order to identify whether the relationship between over-education and job search is a causal one rather than a spurious correlation due to job search being caused by the same factors that cause over-education, IV estimation is employed. This estimation method involves instrumenting the independent variable that causes the endogeneity bias. This endogeneity bias is caused by the correlation between the endogenous explanatory variable and the disturbance term. This is in turn due to factors that influence the endogenous explanatory variable (over-education) also affecting the outcome variable (on-the-job search) but not being controlled for and so being present in the disturbance term.

In this case, it can be expected that unobserved factors such as ability, which are not controlled for in the pooled probit regressions, might be affecting over-education. If this is the case, unobserved ability (or other unobserved individual heterogeneity) might be the factor causing heightened job search rather than the over-educated status itself. For example, ability is expected to have a negative effect on over-education while it can be expected to have a positive effect on on-the-job search as the more able are more likely to be looking out for a new job, while the less able are simply grateful for the job they have got. Given that ability has a negative effect on over-education and a positive one on on-the-job search, it is expected that the endogeneity bias will be negative so the base coefficient in the probit regressions is expected to be biased downwards, and the IV coefficient (which removes this bias) is expected to be larger.

[69] More details in relation to the pooled probit model can be found in section 2.4.1.1 of Chapter 2.

To test this, one of the macro variables found to significantly affect the probability of over-education in Chapter 2 is used as an instrumental variable in order to isolate the true causal effect of over-education on on-the-job search.

In order for an IV to be valid two conditions must be true:

Cov $(z,\varepsilon)=0$: Instrument exogeneity, that is the IV should be exogenous and

Cov $(z,x) \neq 0$: Instrument relevance, that is the IV should be correlated with the endogenous variable x, in this case whether the individual is over-educated.

The estimated equation with a single instrument is expressed as:

$$\mathbf{Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i} \qquad \textbf{(3.2)}$$

The IV regression breaks X into two parts: one potentially correlated with ε and one that is not, hence isolating the part that is not correlated with ε and making it possible to estimate β_1. This is done using an instrumental variable Z_i, which is uncorrelated with ε_i. The IV detects movements in X_i that are uncorrelated with ε_i and uses these to estimate β_1.

The IV regression is a two-stage Least Squares (TSLS) process that runs two different regressions: the first uses OLS and regresses X on Z, therefore isolating the part of X that is uncorrelated with ε and the second uses the predicted values of X_i, $\hat{X}_i$, computed in the first stage to regress Y on $\hat{X}$ using OLS.

More specifically, the first stage regression can be expressed as:

$$\mathbf{X_i} = \pi_{\mathbf{0}} + \pi_{\mathbf{1}}\mathbf{Z_i} + \mathbf{v_i} \qquad \textbf{(3.3)}$$

Due to the fact that Z_i is uncorrelated with ε_i, $\pi_0+\pi_1 Z_i$ is also uncorrelated with ε_i. Even though π_0 and π_1 are unknown they have been estimated above hence the predicted values of X_i, $\hat{X}_i$, are computed where:

$$\hat{\boldsymbol{X}}_{\mathbf{i}}=\hat{\pi}_{\mathbf{0}} + \hat{\pi}_{\mathbf{1}}\mathbf{Z}_{\mathbf{i}}, \qquad \mathbf{i} =, \dots, \mathbf{n} \tag{3.4}$$

In the second stage, X_i is replaced by $\hat{X}_i$ and the following regression is carried out using OLS:

$$\mathbf{Y}_{\mathbf{i}} = \boldsymbol{\beta}_{\mathbf{0}} + \boldsymbol{\beta}_{\mathbf{1}}\, \hat{\mathbf{X}}_{\mathbf{i}} + \boldsymbol{\varepsilon}_{\mathbf{i}} \tag{3.5}$$

Because $\hat{X}_i$ is uncorrelated with ε_i in large samples, β_1 can be estimated by OLS using the regression specification in (3.5) above.

In order for $\pi_0+\pi_1$ to be well estimated using regression (3.4), the sample must be large. The resulting estimator is called the "Two Stage Least Squares" (TSLS or 2SLS) estimator and $\hat{\beta}_1^{TSLS}$ is a consistent estimator of β_1.

As mentioned earlier, an instrument has to be an exogenous variation that affects X. In this case the variable chosen to serve as an instrument, is the percentage change in the supply of labour by educational category and sex whose coefficient was found to be 0.35 in Chapter 2[70]. The rationale for choosing this macro variable to serve as an instrument is the theoretical reasoning that, as the supply of labour by education level and sex increases, the probability of over-education also increases as there is more competition for matched jobs among people with the same education level who compete for the same jobs. This variable is expected to meet the condition of a valid instrument as it is expected to affect the probability of over-education yet not the probability of engaging in on-the-job

[70] More details on this variable can be found in Chapter 2 and in Appendix 3B

search[71] i.e. it offers an exogenous variation or a shock that affects X (i.e. over-education) but is uncorrelated with the disturbance term, ε_i.

To test the appropriateness of the chosen instrument, the weak identification test was run and the results showed that the Kleibergen-Paap Wald rk F statistic equals 31.49. This is above the Stock-Yogo critical values for the 10% maximal value (16.38) i.e. the maximum amount of the possible IV bias that can be tolerated, relative to the OLS bias. In other words, any weak instrument bias will be no more than 10% of the OLS bias (i.e. of the OLS with endogenous regressors) which means that the IV strategy is worth continuing and preferable when compared to OLS.

3.5. Results

3.5.1 Cyprus

The first column in the results table below presents the marginal effects from the binary probit model of the determinants of on-the-job search. The third column shows the IV regression results[72] when over-educated is treated as endogenous and the annualised change in the supply of people within each of the four educational categories: 1) never attended school, 2) primary education, 3) secondary education and 4) tertiary education and divided by sex is used as an IV. Given that the IV regression is a linear regression (i.e. a linear probability model), an OLS linear regression is also run with these results presented in the second column of the below table so as to show that treating the dependent variable as linear does not qualitatively change the results.

In all specifications, the dependant variable is a dummy variable equal to one if a person is looking for another job while employed and equal to zero otherwise.

[71] Finding a good Instrument is not without limitations and there may be cases when it could be argued that on-the-job search could be potentially affected by the annualised change in the supply of labour. Yet, looking at the correlation coefficient between the IV and on-the-job search, which is equal to 0.02, this is unlikely to be the case here.
[72] The 3rd column in Table 3.2 presents the Second Stage results. First Stage IV Regression results can be found in Appendix 3C.

Table 3.2: Probit, Ordinary Least Squares (OLS) and IV regression Results of the Determinants of On-the-Job Search

	Marginal effects from Probit	OLS Coefficients	IV Second Stage Coefficients
Over-educated [Instrumented]	0.017*** (0.002)	0.018*** (0.003)	0.043*** (0.010)
Age Group (20-24)			
25-29	0.003 (0.004)	-0.005 (0.006)	-0.006 (0.006)
30-34	-0.006* (0.004)	-0.022*** (0.006)	-0.023*** (0.007)
35-39	-0.017*** (0.004)	-0.034*** (0.006)	-0.035*** (0.007)
40-44	-0.010** (0.004)	-0.030*** (0.006)	-0.030*** (0.007)
45-49	-0.009* (0.004)	-0.027*** (0.006)	-0.027*** (0.008)
50-54	-0.017*** (0.004)	-0.033*** (0.006)	-0.032*** (0.007)
55-59	-0.022*** (0.005)	-0.035*** (0.006)	-0.034*** (0.008)
60-64	-0.026*** (0.004)	-0.043*** (0.006)	-0.040*** (0.008)
Married (Single, divorced/ widowed)	-0.013*** (0.002)	-0.006* (0.004)	-0.006 (0.004)
Female (Male)	-	0.006 (0.005)	-0.0002 (0.006)
Married*Female	-	-0.023*** (0.005)	-0.022*** (0.008)
Married Females	-0.017*** (0.003)	-	-
Single Females	0.003 (0.004)	-	-
Part Time work (Full time)	0.085*** (0.006)	0.127*** (0.009)	0.126*** (0.012)
Temporary Contract (perm.)	0.037*** (0.004)	0.067*** (0.006)	0.067*** (0.006)
Firm Tenure in months	-0.0002*** (0.00002)	-0.0001***(9.17e-06)	-0.0001*** (9.30e-06)
Size of the firm(1 -10 people)			
11-19 persons	-0.005* (0.003)	-0.007* (0.004)	-0.007* (0.004)
20-49 persons	-0.007** (0.003)	-0.009** (0.003)	-0.009** (0.003)
50 persons or more	-0.010*** (0.002)	-0.01*** (0.003)	-0.011*** (0.003)
Do not know but < 11 persons	0.003 (0.011)	0.002 (0.015)	0.003 (0.015)
Do not know but > 10 persons	-0.021*** (0.006)	-0.024*** (0.007)	-0.025*** (0.009)
Industry NACE Rev 1-1 digit (Agriculture etc.)			
Mining and quarrying	-0.0004 (0.020)	0.016 (0.016)	0.022 (0.017)
Manufacturing	0.030*** (0.007)	0.038*** (0.010)	0.043*** (0.010)
Electricity, gas and water supply	0.022** (0.011)	0.035** (0.011)	0.037** (0.016)
Construction	0.014** (0.006)	0.026*** (0.009)	0.032*** (0.010)
Wholesale and retail trade	0.025** (0.006)	0.036*** (0.009)	0.041*** (0.010)
Hotels and restaurants	0.029*** (0.007)	0.045*** (0.010)	0.047*** (0.011)
Transport, Stor.&Communication	0.018** (0.007)	0.029** (0.010)	0.032*** (0.010)
Financial intermediation	-0.002 (0.006)	0.017* (0.010)	0.016* (0.010)
Real estate, renting and business	0.019*** (0.007)	0.029*** (0.010)	0.032*** (0.011)
Public Administration and defence	-0.005 (0.006)	0.020** (0.009)	0.020** (0.010)
Education	0.023*** (0.007)	0.038*** (0.010)	0.045*** (0.010)
Health and social work	0.009 (0.007)	0.024*** (0.010)	0.029** (0.011)
Other community, social and	0.032*** (0.007)	0.044*** (0.011)	0.049*** (0.013)
personal services activities	-0.019*** (0.005)	-0.068*** (0.010)	-0.076*** (0.012)
Activities of private households	0.015 (0.008)	0.027** (0.012)	0.029** (0.012)
Extraterritorial organizations			
Constant	-	0.045*** (0.011)	0.031** (0.014)
R^2	0.199	0.077	0.074
N	32 252	32 252	32 252

Notes: *Significance at 10% **Significance at 5% ***Significance at 1%; Robust Standard errors for Probit and OLS in brackets; Clustered standard errors around a group variable of education category, sex and year for IV regressions in brackets; Omitted reference groups in brackets next to variable names; Year Dummies are also incorporated in all regressions; Pseudo R^2 reported in the case of Probit model and Centered R^2 in the IV regression result columns.

Looking at the results in the first two columns of the above table, it is evident that, as expected, over-education has a strongly significant positive relationship with on-the-job search. More specifically, over-educated employees are 1.7 percentage points more likely to engage in on-the-job search in the probit model and 1.8 percentage points more likely to engage in on-the-job search in the OLS regression compared to those who are not over-educated[73]. This is a significant number given that only 3.5% of workers conduct on-the-job search at any one time, as it means that on-the-job search is almost 50% more likely for the over-educated.

As also evident in the above table, the effect of over-education on the probability of engaging in on-the-job search is significantly higher when over-education is treated as endogenous. More specifically, those over-educated have a 4.3 percentage point higher chance of looking for another job than people who are not over-educated i.e. a probability of on-the-job search 2.5 times larger than when over-education is treated as exogenous in the probit/OLS regression. This is in line with the predictions about the negative OLS bias due to not controlling for unobserved factors such as ability that could be affecting both the explanatory variable of interest, i.e. over-education and the dependent variable, i.e. on-the-job search. In other words, taking the example of ability, if the more able are less likely to be over-educated but more likely to search on-the-job, then the OLS regressions would bias the coefficient of over-education downwards while instrumenting over-education provides an exogenous shock that only affects over-education enabling the true (and larger in this case) effect of over-education on on-the-job search to unveil.

In terms of the other independent variables, the overall trends in the IV regression results are similar to the Probit/OLS regression results. Firstly, in terms of age, the youngest age group, which is the reference category, has the highest probability of engaging in on-the job search and this propensity to engage in on-the-job search falls with age in all three specifications yet to a somewhat smaller degree in the probit compared to the OLS and IV regressions. This result is in-line with earlier predictions.

[73] As a robustness check and to be sure that the observed effect of over-education on on-the-job search is not driven by the required education part of the over-education variable, regressions including the modal education level by occupation as additional regressors were also run with results being robust to this addition. Hence, even if people in more skilled occupations (i.e. in occupations with a higher modal education level) are more likely to search on the job, there is still an over-education effect, over and above this. In other words, holding skill level of occupation constant, those who are over-educated within a skill level are more likely to search.

Similarly, being married reduces the chances of looking for another job in the probit and to a lesser extend in the OLS regression while being married does not seem to be affecting the probability of on-the-job search in the IV regressions. Being a female has a positive coefficient but does not seem to significantly affect the probability of on-the-job search as it is statistically insignificant in all three specifications.

In terms of the interaction between marital status and sex, which shows the additional effect of being married for a woman, the regression results above demonstrate that married women are less likely to engage in search than single women. In other words, single women, single men and married men all have a similar (insignificantly different from each other) rate of on-the-job search, while the rate for married women is lower. More specifically, married women are 1.7 percentage points less likely to search compared to single women in the probit and OLS regressions and 2.2 percentage points less likely to search on-the-job in the IV regression.

The rest of the independent variables have a very similar effect on on-the-job search across all three regression specifications. More specifically, working part time and under a temporary contract both significantly increase the probability of engaging in on-the-job search in all three specifications even though the effect of the former on on-the-job search is substantially larger than the effect of the latter. On the other hand, increasing job tenure i.e. months working with the current employer is found to have a negative effect on on-the-job search, however this effect is extremely small due to the fact that one month is a relatively short period for job tenure and hence the probability of on-the-job search cannot change much[74]. In terms of the size of the firm, as expected, the larger the firm the lower the likelihood of engaging in on-the-job search, probably because more opportunities for progression within the firm are available compared to smaller firms. The industry in which one is employed also seems to affect the probability of engaging in on-the-job search.

The following IV regression isolates the reason behind on-the-job search and provides a robustness check to the above findings. More specifically, the dependent variable of looking for another job is replaced with the looking for another job 'so as to get better working conditions (e.g. pay, working or travel time, quality of work)' variable which is

[74] A squared term for tenure was also included in the regressions with results showing a positive and significant coefficient meaning that the likelihood of job search falls with tenure, but at a decreasing rate.

more specific as to the reason behind on-the-job search. It is expected that the results will be equivalent to when no distinction as to the reason behind job search is made and such evidence will provide a confirmation as to the causal relationship between over-education and on-the-job search observed earlier.

Table 3.3: IV Regression Results of the Effect of Over-education on On-The-Job Search because of the Wish to have Better Working Conditions

Independent Variables	**IV Second Stage Coefficients**
Over-educated [Instrumented]	0.033*** (0.010)
Age Group (20-24)	
25-29	-0.007 (0.006)
30-34	-0.024*** (0.006)
35-39	-0.032*** (0.005)
40-44	-0.031*** (0.006)
45-49	-0.029*** (0.006)
50-54	-0.029*** (0.006)
55-59	-0.031*** (0.006)
60-64	-0.035*** (0.007)
Female (Male)	-0.001 (0.006)
Married (Single, divorced or widowed)	-0.009** (0.004)
Married*Female	-0.012* (0.007)
Part Time work [Full-time work]	0.068*** (0.009)
Temporary Contract (Permanent job)	0.027*** (0.005)
Firm Tenure in months	-0.00004***(6.99e-06)
Size of the firm (Firm size between 1 to 10 people)	
11-19 persons	-0.001 (0.003)
20-49 persons	-0.004 (0.003)
50 persons or more	-0.008*** (0.003)
Do not know but less than 11 persons	0.001 (0.014)
Do not know but more than 10 persons	-0.025*** (0.005)
Industry (NACE Rev 1-1 digit) (Agriculture etc.)	
Mining and quarrying	0.008 (0.008)
Manufacturing	0.026*** (0.008)
Electricity, gas and water supply	0.024** (0.011)
Construction	0.019** (0.008)
Wholesale and retail trade	0.029*** (0.008)
Hotels and restaurants	0.036*** (0.009)
Transport, storage and communications	0.024** (0.009)
Financial intermediation	0.010 (0.008)
Real estate, renting and business activities	0.022** (0.009)
Public Administration and defence; compulsory social security	0.012* (0.008)
Education	0.031*** (0.008)
Health and social work	0.022** (0.009)
Other community, social and personal services activities	0.033*** (0.010)
Activities of private households	-0.039*** (0.010)
Extraterritorial organizations and bodies	0.017** (0.008)
Constant	0.028** (0.012)
Centered R^2	0.045
N	32 252

Notes: *Significance at 10% **Significance at 5% ***Significance at 1%; Clustered standard errors around a group variable of education category sex and year in brackets next to coefficients; Omitted reference groups in brackets next to explanatory variable names; Year Dummies are also incorporated in the regression.

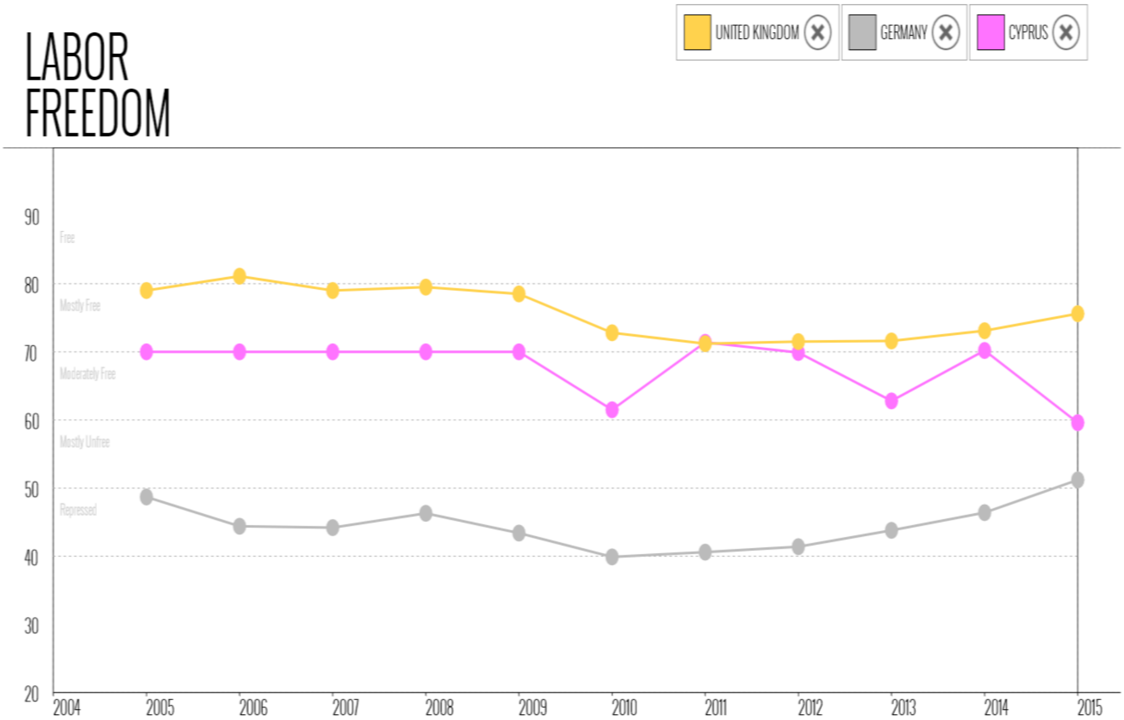
As expected, the results when the dependent variable is restricted to people who search because of the wish to have better working conditions are very similar to when the dependent variable captures on-the-job search in general irrespective of the reason for doing so. Therefore, the earlier results are robust to this change in the dependent variable. However, even though the coefficient of over-education is larger in the above regression than in the OLS and probit regressions, it is slightly lower than the effect of over-education in the IV regression when all reasons for looking for another job are examined jointly.

3.5.2 Replication of Analysis for the UK and Germany

In what follows, the same analysis undertaken for Cyprus above is replicated in another two countries, namely the UK and Germany, again using data from the annual EU-LFS files which contain comparable information for all EU countries. The aim of this analysis is to observe whether Cyprus behaves more like the more flexible UK labour market or more like the more rigid German labour market and to confirm whether the patterns of results of the causal effect of over-education on on-the-job search also apply to other countries apart from Cyprus.

The next graph demonstrates the degree of flexibility or freedom of the Cyprus, British and German labour markets for the period 2004-2015. Labour Freedom is in this case defined as: "*a quantitative measure that considers various aspects of the legal and regulatory framework of a country's labour market, including regulations concerning minimum wages, laws inhibiting layoffs, severance requirements, and measurable regulatory restraints on hiring and hours worked, plus the labour force participation rate as an indicative measure of employment opportunities in the labor market"*(https://www.heritage.org/index/labor-freedom). It consists of the following quantitative sub-factors: ratio of minimum wage to the average value added per worker; hindrance to hiring additional workers; rigidity of hours; difficulty of firing redundant employees; legally mandated notice period; mandatory severance pay and labour force participation rate. It is one of the components of regulatory efficiency that together with the rule of law, government size, and market openness form the Index of Economic Freedom[75].

Figure 3.1 Labour Freedom Index for Cyprus, the UK and Germany



Source:https://www.heritage.org/index/visualize?cnts=cyprus%7Cgermany&src=ranking; Date of extraction 05.04.18

As can be seen in the above figure, the UK scores higher in this index over the period of consideration, being described as having a free to mostly free labour market. This is followed by Cyprus that is found to have a (mostly) moderately free labour market whereas the German labour market is described as repressed. The fact that these three countries differ in the amount of labour market flexibility, makes the comparison of the results interesting.

Summary statistics from the LFS for the UK demonstrate that there is a larger percentage of people in part time work in the UK than in Cyprus (24% compared 5.6% in the case of Cyprus) and fewer temporary contract arrangements (4.5% compared to 14.2% in the case

[75] More details in relation to the Index and its components as well as how it is constructed can be found at the following link: https://www.heritage.org/index/pdf/2018/book/methodology.pdf

of Cyprus) with a shorter average firm tenure of 107.5 months against 111.1 months in Cyprus. In the case of Germany, the percentage of part time work is similar to the UK, i.e. 24.9% while the percentage of temporary work arrangements in Germany lies between the one of Cyprus and the UK i.e. 10.6%. The average job tenure on the other hand is substantially higher in Germany i.e. 133.6 months, in line with the picture of a more rigid labour market drawn above. In terms of on-the-job search, the UK ranks first with a 7.5% on-the-job search rate, which is substantially larger than the same statistic for Cyprus (3.5%) and for Germany (3.3%). The level of over-education is almost identical in the UK and Cyprus, 28.4% in the UK compared to 28.9% in the case of Cyprus, and somewhat lower in Germany where this figure stands at 22.4%[76].

[76] Detailed descriptive statistics for the UK and Germany can be found in Appendix 3D and 3E.

Table 3.4: Probit, OLS and IV Regression Results of the Effect of Over-Education on On-The-Job Search for the UK

	Marginal Effects from Probit	OLS Regression Coefficients	IV Second Stage Coefficients
Over-educated [Instrumented]	0.023*** (0.001)	0.025*** (0.001)	0.093*** (0.016)
Age Group (20-24)			
25-29	-0.001 (0.002)	-0.009*** (0.003)	-0.004 (0.003)
30-34	-0.006** (0.002)	-0.018*** (0.003)	-0.010*** (0.004)
35-39	-0.004* (0.002)	-0.018*** (0.003)	-0.010*** (0.003)
40-44	-0.002 (0.002)	-0.016*** (0.003)	-0.009** (0.003)
45-49	-0.004* (0.002)	-0.018*** (0.003)	-0.010*** (0.003)
50-54	-0.011*** (0.002)	-0.024*** (0.003)	-0.016*** (0.003)
55-59	-0.030*** (0.002)	-0.039*** (0.003)	-0.031*** (0.003)
60-64	-0.060*** (0.002)	-0.067*** (0.003)	-0.057*** (0.003)
Married (Single/divorced/widowed)	-0.021*** (0.001)	-0.010*** (0.001)	-0.008*** (0.002)
Female (Male)	-	-0.008*** (0.002)	-0.011*** (0.003)
Married*Female	-0.027*** (0.001)	-0.022*** (0.002)	-0.022*** (0.003)
Married Female	-0.007*** (0.002)	-	-
Single Female	-0.018*** (0.001)	-	-
Part Time work (Full time work)	0.023*** (0.001)	0.025*** (0.001)	0.018*** (0.004)
Temporary Contract (Permanent)	0.103*** (0.003)	0.132*** (0.004)	0.130*** (0.004)
Firm Tenure in months	-0.0003*** (6.59e-06)	-0.0002***(4.01e-06)	-0.0002***(5.29e-06)
Size of the firm (1-10 people)			
11-19 persons	0.002 (0.002)	0.002 (0.002)	0.0002 (0.002)
20-49 persons	-0.0001 (0.001)	-0.0003 (0.002)	-0.002 (0.002)
50 persons or more	0.0004 (0.001)	-0.0003 (0.001)	-0.003** (0.002)
Do not know < 11 persons	-0.005 (0.008)	-0.002 (0.010)	-0.003 (0.010)
Do not know> 10 persons	-0.004 (0.006)	-0.005 (0.006)	-0.005 (0.008)
Industry NACE Rev 1-1 digit (Agriculture etc.)			
Mining and quarrying	-0.006 (0.007)	-0.005 (0.007)	0.011 (0.009)
Manufacturing	0.017*** (0.005)	0.018*** (0.005)	0.032*** (0.007)
Electricity, gas and water supply	0.018*** (0.006)	0.019*** (0.006)	0.031*** (0.008)
Construction	0.003 (0.005)	0.004 (0.005)	0.022** (0.008)
Wholesale and retail trade	0.030*** (0.005)	0.033*** (0.005)	0.046*** (0.007)
Hotels and restaurants	0.042*** (0.005)	0.051*** (0.006)	0.067*** (0.009)
Transport, storage&communication	0.020*** (0.005)	0.021*** (0.005)	0.037*** (0.007)
Financial intermediation	0.025*** (0.005)	0.025*** (0.006)	0.045*** (0.009)
Real estate, renting and business	0.028*** (0.005)	0.030*** (0.005)	0.047*** (0.008)
Public Administration and defence	0.022*** (0.005)	0.021*** (0.005)	0.035*** (0.007)
Education	0.021*** (0.005)	0.021*** (0.005)	0.040*** (0.008)
Health and social work	0.026*** (0.005)	0.026*** (0.005)	0.041*** (0.008)
Other community, social and personal services activities	0.023*** (0.005)	0.024*** (0.006)	0.042*** (0.008)
Activities of private households	-0.005 (0.010)	-0.007 (0.012)	-0.0001 (0.011)
Extraterritorial organizations& bodies	-0.006 (0.012)	-0.005 (0.011)	0.012 (0.011)
Constant	-	0.085*** (0.006)	0.048*** (0.012)
R^2	0.0668	0.037	0.0237
N	317 712	317 712	317 712

Notes: *Significance at 10% **Significance at 5% ***Significance at 1%; Robust Standard errors for Probit and OLS and Clustered standard errors around a group variable of education category, sex and year for IV regressions in brackets next to coefficients; Omitted reference groups in brackets next to explanatory variable names; Year Dummies are also incorporated in all the regressions; Pseudo R2 reported in the case of Probit model and Centered R2 in the IV regression result columns.

The above table of results demonstrates that the pattern of results is similar in the UK and Cyprus hence providing more confidence that the Cyprus results represent a genuine pattern. The size of the effect is a bit larger in the UK, which can be attributed to its more flexible labour market. More specifically, over-educated workers are 2.5 percentage points more likely to engage in on-the-job search than those not over-educated and when over-education is instrumented using the supply of labour by education level and sex, this probability increases dramatically to 0.095. This means that the causal effect of over-education on on-the-job search is 9.5 percentage points compared to 4.3 percentage points in the case of Cyprus. In terms of the first stage IV regression results[77] and the weak identification tests in the case of the UK, the results show that the supply of graduates has a strongly significant effect of 1.93 on over-education and the Kleibergen-Paap Wald rk F statistic is equal to 34.30, which is significantly above the 10% maximal IV size (16.38). In other words, a 1 percentage point increase in the supply of labour by education category and sex causes a 1.93 percentage point increase in the probability of over-education.

The following table presents the German results.

[77] Provided in Appendix 3F

Table 3.5: Probit, OLS and IV Regression Results of the Effect of Over-Education on On-The-Job Search for Germany

	Marginal Effects from Probit	OLS Regression Coefficients	IV Second Stage Coefficients
Over-educated [Instrumented]	0.005*** (0.0004)	0.005*** (0.0003)	0.057 (0.149)
Age Group (20-24)			
25-29	0.011*** (0.0004)	0.024*** (0.001)	0.018 (0.016)
30-34	0.018*** (0.001)	0.031*** (0.001)	0.026 (0.016)
35-39	0.019*** (0.001)	0.031*** (0.001)	0.026* (0.015)
40-44	0.020*** (0.001)	0.032*** (0.001)	0.028** (0.013)
45-49	0.021*** (0.001)	0.034*** (0.001)	0.030** (0.011)
50-54	0.019*** (0.001)	0.035*** (0.001)	0.030*** (0.011)
55-59	0.016*** (0.001)	0.035*** (0.001)	0.033*** (0.008)
60-64	0.002*** (0.001)	0.027*** (0.001)	0.024*** (0.008)
Married (Single/divorced/widowed)	-0.017*** (0.0003)	-0.006*** (0.0004)	-0.008* (0.004)
Female (Male)	-	-0.001** (0.001)	-0.004 (0.007)
Married*Female	-	-0.027*** (0.001)	-0.024*** (0.008)
Married Female	-0.019*** (0.0004)	-	-
Single Female	-0.002*** (0.001)	-	-
Part Time work (Full time work)	0.046*** (0.001)	0.050*** (0.001)	0.053*** (0.010)
Temporary Contract (Permanent)	0.042*** (0.001)	0.067*** (0.001)	0.069*** (0.008)
Firm Tenure in months	-0.0002*** (2.32e-06)	-0.0001***(1.15e-06)	-0.0001*** (0.00002)
Size of the firm (1-10 people)			
11-19 persons	-0.005*** (0.001)	-0.007*** (0.001)	-0.007*** (0.001)
20-49 persons	-0.006*** (0.001)	-0.009*** (0.001)	-0.009*** (0.002)
50 persons or more	-0.010*** (0.0004)	-0.013*** (0.001)	-0.015*** (0.004)
Industry NACE Rev 1-1 digit (Agriculture etc.)			
Mining and quarrying	-0.007** (0.003)	-0.012*** (0.002)	-0.011** (0.006)
Manufacturing	-0.010*** (0.001)	-0.014*** (0.002)	-0.013*** (0.004)
Electricity, gas and water supply	-0.012*** (0.002)	-0.015*** (0.002)	-0.016** (0.007)
Construction	-0.011*** (0.001)	-0.015*** (0.002)	-0.012 (0.010)
Wholesale and retail trade	-0.009*** (0.001)	-0.013*** (0.002)	-0.013*** (0.004)
Hotels and restaurants	0.007*** (0.002)	0.018*** (0.002)	0.019*** (0.005)
Transport, storage&communication	-0.003* (0.002)	-0.007*** (0.002)	-0.006* (0.004)
Financial intermediation	-0.014*** (0.002)	-0.017*** (0.002)	-0.028 (0.031)
Real estate, renting and business	-0.001 (0.001)	-0.002 (0.002)	-0.005 (0.010)
Public Administration and defence	-0.015*** (0.001)	-0.016*** (0.002)	-0.021 (0.015)
Education	-0.021*** (0.001)	-0.025*** (0.002)	-0.029** (0.013)
Health and social work	-0.017*** (0.001)	-0.022*** (0.002)	-0.030 (0.021)
Other community, social and	-0.006*** (0.002)	-0.009*** (0.002)	-0.010* (0.006)
personal services activities	0.003 (0.002)	0.009** (0.004)	0.011 (0.008)
Activities of private households	0.006 (0.006)	0.001 (0.006)	-0.002 (0.010)
Extraterritorial organizations& bodies			
Constant	-	0.030*** (0.002)	0.024* (0.016)
R^2	0.1442	0.0454	0.031
N	1 604 961	1 604 961	1 604 961

Notes: *Significance at 10% **Significance at 5% ***Significance at 1%; Robust Standard errors for Probit and Clustered standard errors around a group variable of education category sex and year for IV regressions in brackets next to coefficients; Omitted reference groups in brackets next to explanatory variable names; Year Dummies are also incorporated in all the regressions; Pseudo R2 reported in the case of Probit model and Centered R2 in the IV regression result columns.

As demonstrated in the above table, the results for Germany[78] point to a positive but smaller effect of over-education on on-the-job search than in the case of Cyprus and the UK but this effect is only significant in the probit and OLS regressions. In other words, there is less on-the-job search overall in Germany compared to the UK and to a lesser extent compared to Cyprus and it is much less likely to be affected by job conditions such as over-education. It has to be noted that the instrument does not work well in the case of Germany as demonstrated by the very small Kleibergen-Paap Wald rk F statistic (0.08). Moreover, the first stage IV regression results[79] show that over-education itself does not seem to be a result of oversupply of educated labour as the labour supply variable is not found to significantly affect over-education. Therefore, it seems that the German labour market behaves differently in this respect compared to Cyprus and the UK. Wolbers (2003), finds that the effect of being mismatched on the probability of on-the job search is smaller in countries where the share of school-based vocational education is high compared to countries where this share is low. Given that Germany has a large share of vocational education, this could potentially explain why those over-educated do not have as large a probability to look on-the-job compared to those not over-educated and could also explain the lack of responsiveness of over-education to the supply of labour by education category.

Comparing the results of all three countries in terms of the other determinants of on-the-job search, age seems to behave differently in Germany compared to Cyprus and the UK as it seems that as workers age, they have a larger probability of engaging in on-the-job search than the 20-24 age group that serves as the reference category, whilst the opposite was true in the other two countries examined earlier. Similar to Cyprus and the UK, being a female, married or a married female reduces the chances of on-the-job search while the effect of the size of the firm, firm tenure, temporary contract arrangements and working on a part-time basis is similar in all three countries. In conclusion, it can be said that Cyprus seems to behave more like the flexible UK labour market than the more rigid German market.

[78] Whereas in the case of Cyprus and the UK , only observations of respondents in the first occasion they are interviewed (i.e. in interview wave number 1) are kept, in the case of Germany the variable distinguishing between interview waves was not available and hence observations of all interview waves were kept. Results for Cyprus are similar when all waves are used

[79] Available in Appendix 3G

3.6 Summary and Concluding Remarks

The present chapter provides an in-depth study of the phenomenon of on-the-job search with a special interest in its relationship with over-education. It firstly contributes to the literature on the determinants of on-the-job search in general and to the country-specific literature for Cyprus where, to the best of my knowledge, such research does not exist. It finds that, in line with previous empirical research, age appears to have a negative effect on on-the-job search, with workers more likely to search on-the-job when they are younger. Working on a part-time or a temporary basis is also found to increase the likelihood of on-the-job search while an employee's likelihood of searching on-the-job falls with job tenure as the acquisition of firm-specific capital reduces the potential gain from search (Pissarides and Wadsworth, 1994). Married females are also found to elicit less job search efforts than single females while being married also seems to hinder job search efforts irrespective of gender in the probit and OLS specifications while it is found to have a similar yet not statistically significant effect in the Instrumental Variable (IV) resgression.

The main objective of the present chapter was to shed light on the causal relationship between over-education and on-the-job search, contributing in this way towards closing the gap in the existing literature. This was done via the use of IV regressions, so as to isolate the true causal effect of over-education on on-the-job search and to eliminate concerns about the Probit/OLS results being driven by endogeneity. To this end, the use of the annualised change in the labour supply by education and sex, found to significantly affect the probability of over-education in Chapter 2, to instrument over-education, has revealed that treating over-education as endogenous makes the coefficient of this variable larger than when treating it as exogenous. As discussed in this chapter, this might be due to the fact that over-education as well as on-the-job search could be simultaneously affected by unobserved worker heterogeneity which cannot be taken into account in models employing cross sectional data such as the pooled Probit or OLS. This means that the over-education variable is likely to be correlated with the disturbance term in such regressions. For example, the coeffiecient of over-education could be biased downwards if the more able are less likely to be over-educated and more likely to search on-the-job while workers with lower ability are more likely to be over-educated and less likely to search on-the-job if for example they are content with the job they have and doubt whether they could easily find another commensurate one.

By isolating the effect of over-education on on-the-job search and excluding the possible endogeneity bias, the results in the present chapter imply that over-education caused by an over-supply of similarly educated individuals is not voluntary. This is so, as even when the possibility of not searching on-the-job due for example to a worker being of a lower ability (or other reasons such as mobility or other family constraints or worker characteristics such as risk aversion) is taken into consideration, those who are over-educated still search more on the job compared to those not over-educated. These results are also robust to a change in the dependent variable to reflect the reason of job search and more specifically to reflect job search because of the wish to have better working conditions.

The analysis for Cyprus has also been replicated in the UK and Germany and it has been shown that Cyprus is more similar to the UK than to Germany, where the IV regressions did not seem to fit the data very well. More specifically, an over-educated worker was found to be between 2.3-2.5 percentage points (1.7-1.8 in Cyprus) more likely to search on-the-job in the UK compared to a non over-educated worker when over-education is treated as exogenous and 9.3 percentage points more likely to search on-the-job when over-education is treated as endogenous (4.3 in Cyprus). In Germany, this effect is found to be much smaller in the Probit and OLS regressions with over-education causing a mere 0.5 percentage point increase in the probability of employed search while this effect in the IV regressions is found to be considerably larger yet not statistically significant. Therefore, the effect of over-education on on-the-job search is the largest in the UK which has the most flexible or free labour market compared to the other two countries while this effect is the smallest in Germany which has the more rigid labour market, with Cyprus lying in the middle of the two. All in all, it seems that those over-educated do look more on-the-job not only in Cyprus but also in the UK and Germany even though the evidence is weaker for Germany. It also appears to be the case that the relationship between over-education and on-the-job search goes hand in hand with the level of freedom of a country's labour market.

In conclusion, the finding that over-educated workers are indeed more likely to engage in on-the-job search, appears to be in line with the matching theory of over-education which suggests that over-education is sub-optimal from the worker's perspective. Given that on-the-job search is a form of withdrawal behaviour that is likely to be linked to lower productivity levels via lower commitment, over-education is also sub-optimal for the firm and as an extension for the economy as a whole due to lower output levels. In other words, the analysis in this chapter suggests that via its effect on on-the job search, over-education could have a real negative productivity penalty not only for the worker but also for the firm and the economy as a whole and therefore the creation of policies to prevent and/or reduce the level of over-education should be placed high in the political agenda[80].

[80] These are discussed further in Chapter 5.

Appendix 3A: Correspondence Tables for Variables that Change throughout the Years of Analysis

1) NACE REV1.1 to NACE REV 2

NACE industry classification changes in the LFS files starting in 2009 from Rev 1.1. to Rev 2[81]. In order to solve the discrepancy, Rev2 is re-coded back to Rev 1.1 as per the below table. As it can be seen, the changes at the 1 digit level are not that many or major. Each row corresponds to one industry category.

Table A3.1: Correspondence table between sections of NACE Rev 1.1 and NACE Rev. 2

NACE Rev. 1.1		**NACE Rev. 2**	
Section	**Description**	**Section**	**Description**
A **B**	Agriculture, Hunting and Forestry Fishing	**A**	Agriculture, Forestry and Fishing
C	Mining and quarrying	**B**	Mining and quarrying
D	Manufacturing	**C**	Manufacturing
E	Electricity, gas and water supply	**D** **E**	Electricity, gas, steam and air conditioning supply Water supply, sewerage, waste management and remediation activities
F	Construction	**F**	Construction
G	Wholesale and retail trade: repair of motor vehicles, motorcycles and personal and household goods	**G**	Wholesale and retail trade; repair of motor vehicles and motorcycles
H	Hotels and restaurants	**I**	Accommodation and food service activities
I	Transport, storage and communications	**H** **J**	Transportation and storage Information and communication
J	Financial intermediation	**K**	Financial and insurance activities

[81] NACE was also changed before from REV 1 to Rev 1.1. (Rev 1.1 was only used from 2005-2007 (or 2008 in my files) but no differences were found at the 1 digit aggregation level.

K	Real estate, renting and business activities	**L** **M** **N**	Real estate activities Professional, scientific and technical activities Administrative and support service activities
L	Public Administration and defence; compulsory social security	**O**	Public administration and defence; compulsory social security
M	Education	**P**	Education
N	Health and social work	**Q**	Human health and social work activities
O	Other community, social and personal services activities	**R** **S**	Arts, entertainment and recreation Other service activities
P	Activities of private households as employers and undifferentiated production activities of private households	**T**	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use
Q	Extraterritorial organizations and bodies	**U**	Activities of extraterritorial organizations and bodies

Based on the above correspondence table, the below changes were made:

Category A was an exception as it was taken from Rev 1.1 to Rev 2 instead of the other way around. For the years 1999-2008, Section B Fishing (from Rev. 1.1) was merged with Section A: Agriculture, Hunting and Forestry due to fishing being a very small cell throughout the years of analysis. From 2009 when Rev 2 takes over, Section A becomes Agriculture, Forestry and Fishing hence the two revisions become consistent and Section A is kept as it is for the years 2009 onwards.

To go back to Rev 1.1 from Rev 2, from years 2009 onwards the following reclassification took place:

a) Section D (Electricity, gas, steam and air conditioning supply) and E (Water supply, sewerage, waste management and remediation activities) from Rev. 2 were merged back to the Rev 1.1 single category E named Electricity, gas and water supply.

b) Categories H and J from revision 2 were merged into category I of Rev 1.1 (Transport, storage and communications).

c) Categories L, M and N were merged together into category K of Rev 1.1 (Real estate, renting and business activities).

d) Lastly, categories R and S were also merged into category O of Rev. 1.1. (Other community, social and personal services activities).

The rest of the categories were either unchanged at the one digit level or had a slightly different wording that was not altering their context.

Detailed NACE correspondence tables at a higher aggregation level and more information about NACE can also be found at the following link: http://ec.europa.eu/eurostat/web/nace-rev2/correspondence_tables

2) International Standard Classification of Education (ISCED) 1997: For years 2014 and 2015 the ISCED 1997 changed to ISCED 2011 in the LFS files. In these years, the below correspondence table provided by UNESCO was used to convert ISCED11 back to ISCED97 to guarantee continuity.

Table A 3.2: Correspondence between ISCED 2011 and ISCED 1997 levels

ISCED 2011	**ISCED 1997**
ISCED 01	-
ISCED 02	ISCED 0
ISCED level 1	ISCED level 1
ISCED level 2	ISCED level 2
ISCED level 3*	ISCED level 3
ISCED level 4*	ISCED level 4
ISCED level 5	
ISCED level 6	ISCED level 5
ISCED level 7	
ISCED level 8	ISCED level 6
*Content of category has been modified	slightly

Source:http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf

3) **ISCO-88 occupational classification**: The ISCO-88 occupational classification changed to ISCO-08 starting in 2011. In order to be able to have an uninterrupted and comparable set of classifications throughout the whole period of analysis, correspondence tables provided by the ILO (available at the following link: http://www.ilo.org/public/english/bureau/stat/isco/isco08/) were used to take ISCO-08 back to ISCO-88 in the 2011-15 period. Due to their length they are omitted here, but they can be accessed in the link above.

Appendix 3B: Description of Variables Used in Chapter 3

Age: Age in the LFS is disseminated in 5 year age bundles, e.g. 15-19, 20-24, 25-29 etc.

Sex: This is a binary variable equal to 1 for women and zero for men.

On-the-Job Search: This is a binary variable equal to 1 if the person is employed and looking for another job and equal to zero if he/she is employed but is not looking for another job at the time of the survey.

Main reason for looking for another Job: The original variable as disseminated in the LFS consists of the following categories:

1 -Risk or certainty of loss or termination of present job

2 - Actual job is considered as a transitional job

3 - Seeking an additional job to add more hours to those worked in present job

4 - Seeking a job with more hours worked than in present job

5 - Seeking a job with fewer hours worked than in present job

6 - Wish to have better working conditions (e.g. pay, working or travel time, quality of work)

7 - Other reasons

This variable was recoded for the purposes of the present Chapter so that all categories were pulled together except for category 6 above that is the main category of interest. The resulting variable is a binary variable equal to 0 if the reason for searching for another job fell within any of the categories numbered 1-5 or 7, listed above and equal to 1 if the reason for looking for another job is number 6 above, i.e. wish to have better working conditions (e.g. pay, working or travel time, quality of work).

Over-education: The over-education variable in this Chapter is calculated by taking the (minimum) modal education level (2 digit International Standard Classification of Education (ISCED) 1997 as it appears in the LFS variable HATLEVEL=highest level of education or training successfully completed)) by occupation (ISCO-88 3 digit) categories. The education categories based on which the modal value is calculated are the following:

00 No formal education or below ISCED 1

11 ISCED 1

21 ISCED 2

22 ISCED 3c (shorter than two years)

31 ISCED 3c (two years and more)

32 ISCED 3 a, b

30 ISCED 3 (without distinction a, b or c possible, 2 y+)

41 ISCED 4a, b

42 ISCED 4c

43 ISCED 4 (without distinction a, b or c possible)

51 ISCED 5b

52 ISCED 5a

60 ISCED 6

Economic activity of the local unit NACE REV.1.1: These are the 1-digit industry classifications as classified using the NACE Rev. 1.1:

A) Agriculture, Hunting and Forestry
B) Fishing
C) Mining and quarrying
D) Manufacturing
E) Electricity, gas and water supply
F) Construction
G) Wholesale and retail trade: repair of motor vehicles, motorcycles and personal and household goods
H) Hotels and restaurants
I) Transport, storage and communications
J) Financial intermediation
K) Real estate, renting and business activities
L) Public Administration and defence; compulsory social security
M) Education

N) Health and social work

O) Other community, social and personal services activities

P) Activities of private households as employers and undifferentiated production activities of private households

Q) Extraterritorial organizations and bodies

Marital status: The original LFS variable for marital status consisted of the following categories: 1 Single; 2 Married; 3 Widowed; 4 Divorced or legally separated. For the purposes of the present chapter this variable was recoded to be equal to 0 if Single, Widowed and/or Divorced or legally separated (i.e. categories 1, 3 and 4 pooled together) and equal to 1 if Married.

Working Part time VS Full time: This is a binary variable equal to 0 if the person has a full-time job and equal to 1 if he/she has a Part-time job.

Permanency of the job: This is a binary variable equal to 0 if a person has a permanent job or work contract of unlimited duration and equal to 1 if a person has temporary job/work contract of limited duration.

Firm tenure: Time in months since the person started current employment. This variable, available in the LFS is derived by the reporting of the year and month a person started working with their current employer.

Size of the firm: Number of persons working at the local unit. This variable has the following categories:

01-10 Exact number of persons, if between 1 and 10

Category number 11: 11 to 19 persons

Category number 12: 20 to 49 persons

Category number 13: 50 persons or more

Category number 14: Do not know but less than 11 persons

Category number 15: Do not know but more than 10 persons

Annualised change in the Labour Supply by educational category and Sex (used as an Instrumental Variable for over-education): Description of this variable is available in Appendix 2D.

Appendix 3C: First Stage IV Regression Results Estimating Predicted Value of Over-Education in Cyprus

Independent Variables	IV Coefficients	
Annualised change in the supply of labour by education category and sex	1.88***	(0.335)
Age Group (20-24)		
25-29	0.062***	(0.013)
30-34	0.046***	(0.015)
35-39	0.031**	(0.015)
40-44	0.022*	(0.015)
45-49	0.009	(0.017)
50-54	-0.016	(0.020)
55-59	-0.036*	(0.024)
60-64	-0.056**	(0.028)
Female (Male)	0.139***	(0.047)
Married (Single, divorced or widowed)	-0.003	(0.012)
Married Female (Single, divorced or widowed Male)	-0.027*	(0.016)
Part Time work (Full-time work)	0.011	(0.013)
Temporary Contract (Permanent job or work contract of unlimited duration)	-0.010	(0.009)
Firm Tenure in months	-0.0003***	(0.00004)
Size of the firm(Firm size 1-10 people)		
11-19 persons	0.004	(0.009)
20-49 persons	0.010	(0.009)
50 persons or more	0.010	(0.010)
Do not know but less than 11 persons	-0.056**	(0.023)
Do not know but more than 10 persons	0.037*	(0.020)
Industry (NACE Rev 1-1 digit) (Agriculture etc.)		
Mining and quarrying	-0.281***	(0.046)
Manufacturing	-0.240***	(0.039)
Electricity, gas and water supply	-0.111**	(0.040)
Construction	-0.252***	(0.040)
Wholesale and retail trade	-0.200***	(0.049)
Hotels and restaurants	-0.113**	(0.049)
Transport, storage and communications	-0.134***	(0.047)
Financial intermediation	-0.020	(0.049)
Real estate, renting and business activities	-0.164***	(0.042)
Public Administration and defence; compulsory social security	-0.055	(0.049)
Education	-0.371***	(0.051)
Health and social work		
Other community, social and personal services activities	-0.207***	(0.045)
Activities of private households	-0.226***	(0.041)
Extraterritorial organizations and bodies	0.301***	(0.066)
	-0.098**	(0.049)
Constant		
Centered R^2	0.558***	(0.103)
N	0.0742	
	32 252	

Notes: *Significance at 10% **Significance at 5% ***Significance at 1%; Omitted reference groups in brackets next to variable names; Year Dummies are also incorporated in the regressions; Clustered Standard errors in brackets next to coefficients

Appendix 3D: Descriptive Statistics for the UK

Variable	Mean	Standard Deviation
On-the-Job Search for another Job	0.075	0.264
Over-education	0.284	0.451
Age Group		
25-29	0.111	0.314
30-34	0.122	0.327
35-39	0.128	0.334
40-44	0.139	0.346
45-49	0.134	0.341
50-54	0.122	0.327
55-59	0.098	0.297
60-64	0.058	0.233
Job Tenure (in months)	107.537	105.447
Married	0.559	0.496
Female	0.512	0.500
Temporary contracts	0.045	0.208
Part time work	0.242	0.429
Industry (NACE Rev 1-1 digit)		
Mining and quarrying	0.005	0.069
Manufacturing	0.126	0.332
Electricity, gas and water supply	0.013	0.113
Construction	0.054	0.225
Wholesale and retail trade: repair of motor vehicles, motorcycles and personal and household goods	0.135	0.342
Hotels and restaurants	0.041	0.198
Transport, storage and communications	0.079	0.270
Financial intermediation	0.045	0.207
Real estate, renting and business activities	0.108	0.310
Public Administration and defence; compulsory social security	0.079	0.270
Education	0.116	0.321
Health and social work	0.149	0.356
Other community, social and personal services activities	0.041	0.198
Activities of private households as employers and undifferentiated production activities of private households	0.002	0.039
Extraterritorial organizations and bodies	0.001	0.035
Size of the firm		
11-19 persons	0.197	0.397
20-49 persons	0.535	0.499
50 persons or more	0.003	0.053
Do not know but less than 11 persons	0.006	0.079
Do not know but more than 10 persons		
Macro level Variable used as Instruments:		
Annualised change in the supply of graduates by educational category and sex	0.008	0.032
N	**323 655**	

Appendix 3E: Descriptive Statistics for Germany

Variable	Mean	Standard Deviation
On-the-Job Search for another Job	0.033	0.177
Over-education	0.224	0.417
Age Group		
25-29	0.102	0.303
30-34	0.119	0.324
35-39	0.129	0.335
40-44	0.140	0.347
45-49	0.146	0.352
50-54	0.130	0.336
55-59	0.099	0.299
60-64	0.050	0.218
Job Tenure (in months)	133.61	125.315
Married	0.574	0.494
Female	0.478	0.500
Temporary contracts	0.107	0.309
Part time work	0.249	0.433
Industry (NACE Rev 1-1 digit)		
Mining and quarrying	0.006	0.076
Manufacturing	0.226	0.418
Electricity, gas and water supply	0.013	0.115
Construction	0.065	0.246
Wholesale and retail trade: repair of motor vehicles, motorcycles and personal and household goods	0.134	0.341
Hotels and restaurants	0.031	0.172
Transport, storage and communications	0.069	0.254
Financial intermediation	0.036	0.187
Real estate, renting and business activities	0.088	0.283
Public Administration and defence; compulsory social security	0.087	0.282
Education	0.068	0.251
Health and social work	0.121	0.326
Other community, social and personal services activities	0.041	0.199
Activities of private households as employers and undifferentiated production activities of private households	0.005	0.068
Extraterritorial organizations and bodies	0.001	0.028
Size of the firm		
11-19 persons	0.117	0.321
20-49 persons	0.151	0.357
50 persons or more	0.549	0.498
Macro level Variable used as Instruments:		
Annualised change in the supply of graduates by educational category and sex	-0.003	0.030
N	**1 635 846**	

Appendix 3F: First Stage IV Regression Results Estimating Predicted Value of Over-Education in the UK

Independent Variables	**IV First stage coefficients**
Annualised change in labour supply by education category and sex [Instrument]	1.929*** (0 .329)
Age Group (20-24)	
25-29	-0.074*** (0.007)
30-34	-0.110*** (0.009)
35-39	-0.117*** (0.009)
40-44	-0.111*** (0.009)
45-49	-0.113*** (0.010)
50-54	-0.110*** (0.011)
55-59	-0.109*** (0.013)
60-64	-0.111*** (0.013)
Female (Male)	0.023 (0.020)
Married (Single, divorced or widowed)	-0.030*** (0.006)
Married Female (Single, divorced or widowed Male)	0.010 (0.008)
Part Time work (Full-time work)	0.109*** (0.007)
Temporary Contract (Permanent job/work contract of unlimited duration)	0.034*** (0.009)
Firm Tenure in months	-0.0002***(0.00002)
Size of the firm(Firm size between 1 to 10 people)	
11-19 persons	0.018*** (0.004)
20-49 persons	0.023*** (0.005)
50 persons or more	0.035*** (0.007)
Do not know but less than 11 persons	0.013 (0.015)
Do not know but more than 10 persons	0.014 (0.010)
Industry (NACE Rev 1.1 digit) (Agriculture hunting, forestry, fishing)	
Mining and quarrying	-0.261*** (0.032)
Manufacturing	-0.200*** (0.026)
Electricity, gas and water supply	-0.186*** (0.028)
Construction	-0.270*** (0.030)
Wholesale and retail trade: repair of motor vehicles, motorcycles and personal and household goods	-0.193*** (0.027)
Hotels and restaurants	-0.240*** (0.036)
Transport, storage and communications	-0.245*** (0.027)
Financial intermediation	-0.313*** (0.034)
Real estate, renting and business activities	-0.267*** (0.028)
Public Administration and defence; compulsory social security	-0.220*** (0.033)
Education	-0.322*** (0.030)
Health and social work	-0.244*** (0.030)
Other community, social and personal services activities	-0.278*** (0.027)
Activities of private households	-0.103*** (0.036)
Extraterritorial organizations and bodies	-0.267*** (0.043)
Constant	0.529*** (0.043)
Centered R^2	0.0237
N	**317 712**

Notes: *Significance at 10% **Significance at 5% ***Significance at 1%; Omitted reference groups in brackets; Year Dummies are also incorporated in the regressions; Omitted Reference Categories in brackets next to variable names; Clustered Standard errors in brackets next to coefficients

Appendix 3G: First Stage IV Regression Results Estimating Predicted Value of over-education in Germany

Independent Variables	**IV First stage coefficients**
Annualised change in the supply of labour by education category and sex [Instrument]	0.312 (1.110)
Age Group (20-24)	
25-29	0.105*** (0.010)
30-34	0.106*** (0.014)
35-39	0.103*** (0.016)
40-44	0.083*** (0.018)
45-49	0.071*** (0.019)
50-54	0.057*** (0.020)
55-59	0.049** (0.022)
60-64	0.053** (0.026)
Female (Male)	0.039 (0.040)
Married (Single, divorced or widowed)	0.027** (0.011)
Married Female (Single, divorced or widowed Male)	-0.052*** (0.016)
Part Time work (Full-time work)	-0.048*** (0.007)
Temporary Contract (Permanent job/work contract of unlimited duration)	-0.045*** (0.005)
Firm Tenure in months	-0.0002*** (0.00002)
Size of the firm(Firm size between 1 to 10 people)	
11-19 persons	0.001 (0.003)
20-49 persons	-0.008* (0.005)
50 persons or more	0.024*** (0.004)
Industry (NACE Rev 1-1 digit) (Agriculture hunting, forestry, fishing)	
Mining and quarrying	-0.023 (0.019)
Manufacturing	-0.010* (0.006)
Electricity, gas and water supply	0.035*** (0.011)
Construction	-0.059*** (0.012)
Wholesale and retail trade: repair of motor vehicles, motorcycles and personal and household goods	-0.003 (0.008)
Hotels and restaurants	-0.025** (0.010)
Transport, storage and communications	-0.005 (0.010)
Financial intermediation	0.212*** (0.012)
Real estate, renting and business activities	0.062*** (0.010)
Public Administration and defence; compulsory social security	0.097*** (0.008)
Education	0.082*** (0.026)
Health and social work	0.141*** (0.021)
Other community, social and personal services activities	0.031*** (0.009)
Activities of private households	-0.032** (0.012)
Extraterritorial organizations and bodies	0.054*** (0.012)
Constant	0.101 (0.075)
Centered R^2	0.031
N	**1 604 961**

Notes: *Significance at 10% **Significance at 5% ***Significance at 1%; Omitted reference groups in brackets; Year Dummies are also incorporated in the regressions; Omitted Reference Categories in brackets next to variable names; Clustered Standard errors in brackets next to coefficients

Chapter 4: Job Polarisation and its Effect on Job Mobility in Cyprus: A Pseudo Panel Analysis using LFS Data for the Period 1999-2014.

4.1 Introduction

The term job polarisation, or hollowing-out, employed in the literature to refer to a range of related phenomena, was initially used to describe the increased employment and labour demand for jobs at the high and low ends of the wage spectrum. At the same time, the fall in employment in mid-level jobs has raised concerns about the potential disappearance of middle-skill jobs (Goos and Manning, 2007). Technological change and more specifically the theory of Task-Biased Technological Change (TBTC) (Autor, Levy, and Murnane, 2003) (hereafter referred to as ALM) has been frequently employed to explain the existence of polarisation. According to AML's routinisation hypothesis, jobs in the middle of the job distribution are the ones most likely to be replaced by technology as they largely comprise of routine tasks. On the other hand, jobs at the high and low ends of the job continuum involve tasks that are complementary rather than replaceable by technology and hence increase in numbers following the drop in the costs of technology over time (Nordhaus, 2007). This creates a U-shaped jobs distribution.

The polarisation hypothesis has provided a prominent justification for recent employment developments in the US and other developed countries (Goos and Manning 2007; Acemoglu and Autor 2011; Goos et al. 2014) and according to Cedefop (2012) prognoses, it is likely to continue to play an important role in Europe in the coming years, even though research in relation to Cyprus is very limited. One of the implications of job polarisation is that, apart from a surge in jobs at the high end of the job continuum, numerous jobs at the opposite end that offer lower wages and disadvantageous terms and conditions are also created, translating into concerns for policy-makers in relation to equity and social cohesion (Cedefop, 2009). This makes it imperative to study the quantitative evolution of jobs over the years so as to properly anticipate future skill needs and job opportunities which is essential for sound economic policy (Kampelmann and Rycx, 2011). This is done in the first part of the present chapter.

Moreover, an important aspect of polarisation that has not attracted much attention in the literature is the effect of this hollowing-out of the job distribution on job mobility. In fact, most of the literature on the effects of polarisation has been focused on wage inequality (Autor et al. 2006a; Goos and Manning 2007). However, changes in the occupational structure can also have a significant impact on job mobility. Firstly, as employment in mid-level jobs declines, job polarisation may result in reduced job opportunities for low-wage workers to progress to. As Holmes (2011) points out, job polarisation may result in a bottleneck with openings of well-paid, high-level jobs above the bottleneck being filled with increasingly well-qualified new entrants, instead of via the career progression of older workers in low or mid-level jobs (Holmes, 2011). Secondly, looking into job mobility of displaced mid-level workers following polarisation is very important in order to understand whether they move up to high-level jobs or end up in low-level jobs and what are the aspects that determine these transitions. If they move down the jobs distribution and into low-level jobs this means that their labour market position deteriorates and such a finding is a cause of concern for policy makers that would require policy interventions so as to correct it. This latter aspect of job mobility, i.e. job mobility of displaced mid-level workers, is the focus of the second part of the present chapter.

4.1.1 Research Questions and Chapter Objectives

The present chapter endeavours to offer a twofold contribution to the literature. First, it will establish whether job polarisation has indeed taken place in Cyprus, a country where the phenomenon has not been examined at the individual country level, during the period of 1999-2014. In this chapter, polarisation will be defined as the growth in the shares of total employment in high-ranked and low-ranked jobs relative to middle-ranked jobs over time. Instead of using wages as the sole indicator to characterise the quality of the jobs whose quantitative evolution is analysed, a secondary indicator based on the average educational level of workers within each job will also be added. To this end, and using the jobs approach, a novel methodology in the analysis of job change, jobs defined as occupations within industries, will be ranked both according to their average wage and according to their modal education level at the start of the period.

Once the pattern of job change has been established, the second objective of this chapter will be to look at job transitions of workers displaced from mid-level jobs due to routinisation in Cyprus between 1999 and 2014 and to answer calls in the literature for more research in relation to job mobility of displaced mid-level workers. In other words, it will look at transitions in and out of the various job groups to examine whether displaced workers from the middle of the wage distribution end up at the lower or higher ends of the wage distribution or in unemployment/inactivity. This will be done using repeated cross sections to form pseudo cohorts of workers who are then followed over time. More specifically, pseudo cohorts of workers based on age group and education level will be created and regressions will be run at the cohort level so that each observation in this pseudo panel will be a cohort at a point in time. In other words, having drawn the aggregate picture across all cohorts, the chapter will go on to demonstrate, using Instrumental Variable (IV) regression analysis, the effects of those polarisation shifts within cohorts. To my knowledge, there do not exist other studies in the literature of job polarisation that use pseudo cohorts to establish the impact of routinisation on job mobility of workers displaced from mid-paid jobs.

The present chapter is structured as follows: Section 4.2 discusses the related literature on job polarisation, Section 4.3 describes the data and methodology and presents the descriptive statistics of the relevant variables, Section 4.4 presents the econometric results and lastly Section 4.5 summarises and concludes.

4.2 Literature Review

4.2.1 Theories of Job Polarisation in the Literature

Three labour demand mechanisms trying to explain polarisation are found in the literature. Firstly, globalisation and the higher propensity to offshore labour services in many production jobs that are found in the middle of the wage distribution (see Hijzen 2007; Blinder 2009; Oldenski 2012) has been put forward as an explanation of the hollowing in the middle of the job distribution. More specifically, it has been often argued that in industrialised countries, it is more profitable and easier, hence more likely, for the labour-intensive parts of production to be offshored, so as to allow production at home to concentrate on more capital-or skill-intensive production (e.g. Glass and Saggi, 2001). For example, theoretical models based on the view that jobs involving routine tasks are more offshoreable than jobs involving non-routine tasks have been developed by Antras et al. (2006) and Grossman and Rossi-Hansberg, (2008) while Blinder's (2007) and Jensen and Kletzer's (2010) empirical work for the US analyses the jobs' task content to evaluate their potential for international offshoring.

Secondly, income inequality has been argued to lead to an increased demand for certain low-paid service jobs. In other words, the increase in income of the top earners is expected to drive up the demand for and hence employment of low-skill service workers (Gadrey 1996; Manning 2004; Autor and Dorn 2009) creating a U-shaped job distribution with jobs in the high and low ends increasing while those in the middle remain stagnant.

Even though both the offshorability of jobs and wage inequality have a definite impact on specific occupations, empirical studies come to the conclusion that these factors are of secondary importance for the overall evolution of the occupational employment structure (see Goos et al., 2009). On the other hand, the third explanation of polarisation, developed by Autor, Levy, and Murnane (2003), based on a refinement of the Skill-Biased Technical Change (SBTC) hypothesis has been far more successful in accounting for polarisation. The idea behind the SBTC theory is that technological developments lead to greater demand for high-skilled workers (upskilling/upgrading of the labour force) as they are the ones that are complementary to technological change and hence benefit the most from computers and information and communication technology (ICT). While the SBTC has provided the main explanation in relation to shifts in demand for workers of different skill levels (Violante, 2008), predicting a linear shift in employment demand in favour of the high rather than the low-skilled and has been hence consistent with the upgrading trend in the employment structure observed in many countries over the past decades, it cannot account for the simultaneous relative growth in lower-level jobs observed in numerous countries.

ALM (2003) argue that the impact that new technologies have on occupations is determined to a large extent not on their skill content as such but on the tasks that they perform. More specifically, the theory of Task-Biased Technological Change (TBTC) stresses the substitutability between routine tasks and technology (Autor et al., 2003) and goes further than the standard theory of SBTC in predicting real labour market changes in the middle and lower parts of the wage distribution in recent years.

Routine tasks are defined by ALM (2003) as tasks that follow clear rules and procedures that can be "specified in computer code and executed by machines" (ALM 2003, 1283).

Given the significant drop in the costs of technology over time (Nordhaus, 2007), it follows that firms are expected to substitute labour with technology in routine tasks, leading to job losses (gains) in occupations carrying out routine (non-routine) tasks. In other words, according to ALM (2003), the falling price of computer power or falling automation costs should cause a drop in the relative demand for labour performing routine tasks (e.g., bookkeepers, repetitive production work). On the other hand, machines or computers cannot be programmed to carry out non-routine and interactive tasks such as those carried out by a hairdresser or a waiter for example, and hence the shares of these lower-paid, low-skill but non-routine service sector jobs are expected to grow as has been the case in recent years. In other words, human labour performing routine tasks, i.e. tasks that can be expressed by rules or step-by-step procedures, is the one mostly at risk of being replaced by computer-driven technology, while this is not the case for human capital in jobs that involve non-routine tasks which are relatively resistant to the advances of technological change.

Goos and Manning (2007) pointed out that routine tasks are mostly found in the middle of the wage distribution. It follows that the substitutability of routine tasks in the middle of the wage distribution with machines, combined with the complementarity of non-routine cognitive tasks found at the high paying end of the wage distribution (e.g. engineers, economists) with computers and the non-complementary nor-substitute nature of the non-routine manual tasks found at the bottom of the wage distribution (e.g. cleaners, waiters) could provide a possible explanation for the hollowing-out of the jobs distribution. In other words, ALM's (2003) theory seems to fit well with the pattern of observed polarisation where there is a disproportionate growth at both ends of the job spectrum.

4.2.2 Review of the Empirical Literature

4.2.2.1 Job Polarisation

In a review of the literature, Acemoglu and Autor (2011) state that US and EU labour markets have endured "systematic, non-monotonic changes in the structure of employment across occupations" giving rise to quick instantaneous increase of both high education, high wage occupations and low education, low wage occupations (Acemoglu and Autor 2011, 1046). Similarly, in another literature review, McIntosh (2013) notes that job polarisation when jobs are ranked according to initial wage is indeed a real phenomenon in all the developed countries in which it has been studied.

For the US, Autor and Dorn (2009) find that after 1980 employment and wages in labour markets initially specialised in routine-intensive occupations both polarised and study both theoretically and empirically the drivers of the changing shape of low-wage and low-skill employment in the US labour market. Considering US commuting zones, they examine the change in employment between 1980 and 2005 in routine-intensive, high-skill non-routine jobs and low-skill non-routine jobs expressing the change in employment as the change in each occupation's employment share within one of three age groups: 16-29, 30-54 and 55-64. They explain these changes as a function of the commuting zone's initial share of routine occupations (as an indicator of its potential routinisation) with their results pointing to the fact that the observed changes in jobs and wages were related to this routinisation measure, in the ways predicted by the TBTC model. Moreover, Autor and Dorn (2009) find employment and wage polarisation within regional labour markets that is equivalent to the employment polarisation observed at the aggregate level in the US, UK and West Germany. All in all, their findings suggest that the displacement of routine task activities that causes changes in labour specialisation results in increasing employment and wages in service occupations and to polarisation of employment and wage growth more generally (Autor and Dorn, 2009).

Autor et al. (2006), ranking occupations according to their initial average education level instead of their initial wages, again find a polarisation pattern with increases in employment shares in high and low-level occupations for the period 1990-2000, while for the period 1980-1990 they find that the trend has been one of upgrading with higher employment growth in occupations with higher levels of initial education.

On an EU level and using data from the European Labour Force Survey for a number of EU countries, the next three studies deviate in their results. Firstly, Goos et al. (2009), looking at the connection between initial wages and the changing patterns of employment shares for a panel of European countries between 1993 and 2006, find evidence of job polarisation in Europe as a whole. More specifically, looking at ISCO-88 two-digit occupations, they find that the four lowest-paying and the eight highest-paying occupations see a growth in their employment shares, while the nine middling occupations see a fall in their employment shares. Accounting for the task content of occupations, Goos et al.'s model makes a distinction between three types of tasks: abstract (intense in non-routine cognitive skills), service (intense in non-routine non-cognitive skills), and routine (intense in both cognitive and non-cognitive routine skills). They then run a cross-country regression controlling for the offshoreability and educational composition of occupations and find employment change between 1993 and 2006 to be is positively associated with abstract and service task importance, but negatively correlated with routine task importance.

Even though Goos et al.'s (2009) expanded analysis of European Union (EU15) countries before the 2004 enlargement, described above, found a consistent pattern of job polarisation between 1993 and 2006 in all countries except Portugal, Fernandez-Macias (2012) and Oesch and Rodriguez-Menes (2011) show that there are differences across EU countries in relation to employment change with countries following different polarisation patterns. More specifically, Fernández-Macías (2012) shows that, for the period 1995-2007, instead of a homogeneous pattern of change in employment structures, there was a variety of patterns of structural employment change across EU countries and argues that this can be better understood if, along with factors such as technology (routinisation), other factors such as institutions (i.e. skill supply evolution and wage-setting institutions) also play a role in the structure of employment and its evolution over time. Similar to the present chapter, he also uses the jobs approach, describing such an approach as an innovation in labour market structural analysis. Moreover, Oesch and Rodriguez-Menes (2011) examining the pattern of occupational change in Britain, Germany, Switzerland and Spain, four countries with different labour market institutions, find considerable occupational upgrading across all four countries (particularly in Germany) while they find significant cross-country variations in employment change especially in low but also in mid-level jobs. They argue that this finding is in contrast to claims of a uniform technical change and that wage-setting institutions filter the pattern of occupational change. They add that countries are more likely to experience polarisation if such institutions enable the creation of low-paid jobs with their results suggesting that this is more the case in Britain and Spain rather than in Germany and Switzerland.

Overall, although the majority of studies have found job polarisation in the US and the UK during the 1990s (e.g. Wright and Dwyer 2003; Goos and Manning 2007), there does not seem to be a consensus as to the patterns of job change in countries like Germany, Netherlands, Sweden, Spain as well as a few other European and developed Asian countries. As mentioned earlier, evidence on job polarisation for Cyprus is almost inexistent. Given that studying trends in the structure of employment over time can provide an indication of future changes and can hence be very useful in informing policy makers and other labour market participants (Wilson, 2008), the present chapter will add evidence for Cyprus based on two proxies of job quality.

4.2.2.2 Polarisation and Job Mobility

McIntosh(2013) in his review of the literature on job mobility in and out of the various job groups notes that evidence addressing questions of job mobility is relatively scarce, presumably due to data restrictions as such analyses require panel data.

Holmes (2011) also notes that, up to this point, very little attention has been paid to the effects of the change in the occupational structure on mobility and studies the fortunes of workers displaced from routine occupations. More specifically, he attempts to track individuals who are displaced from routine occupations to observe their career mobility using UK panel data from the National Child Development Study (NCDS). Holmes (2011) observes individuals who started off in routine jobs in his initial year of data, 1981, and analyses their likelihood of transition into other categories of employment between 1981 and 2004 separating his time frame into 4-5 year bundles. The other employment categories he considers are Professional, Managerial, Intermediate/Associate professional, Service, Manual non-routine, or Remaining in a routine job. Holmes (2011) uses a separate binary logit model specification for estimating the probability of transition from routine jobs to each of the other employment categories listed above while controlling for a number of background characteristics such as age (capturing general work experience), gender, ethnicity, specific experience and education/qualifications. In order to establish causation of the flows from routine jobs to other job groups specifically due to routinisation, Holmes (2011) adds another explanatory variable that measures ‘routinisation’ i.e. the overall change in employment in routine jobs in each period. This allows him to distinguish between job mobility for technological reasons from job mobility for normal career mobility reasons, for example, because of career progression from a routine role to a managerial role. Therefore, there are two forms of occupational moves in his model: career mobility and routinisation-driven mobility. Finally, by allowing for new entrants into the labour market, his model looks at the experiences of different waves of labour market entrants.

Holmes' (2011) main result is that his routinisation variable has a positive and significant effect on mobility in all equations except the Managerial one which means that routinisation drives job mobility from Routine jobs both upwards to Professional and Intermediate occupations as well as downwards to Service occupations but not to Managerial occupations. So, there is a higher mobility out of routine jobs and into the before-mentioned occupations in periods of great technological change. Holmes (2011) moves on to add some interaction terms between the routinisation variable and education and qualifications as well as work experience, both general and specific. In summary, he finds that having a higher level of qualifications increases the chances of moving from declining Routine occupations up to Professional occupations while having intermediate level qualifications makes it more possible to move from Routine to Managerial and Intermediate occupations. According to Holmes (2011), higher qualifications do not seem to be protecting individuals from moving to lower-level jobs following routinisation. Moreover, Holmes (2011) finds that the more routine specific experience a person has, the less likely they are to move out from Routine jobs compared to those with less routine-specific skills who are more likely to move out of Routine jobs when these decline. General work experience follows a similar path, however older workers seem to have an advantage over younger workers in case they are displaced from Routine jobs as they are more likely to move to Professional and Managerial jobs compared to younger workers.

Similar to the above work by Holmes (2011), is another paper from Holmes et al. (2011) that adds another (younger) cohort from the British Cohort Study (BCS), born in a specific week in April 1970. Holmes et al. (2011) perform the same analysis for people originally in routine occupations as in Holmes (2011), now covering two periods, 1981-2004 with the NCDS cohort and 1996-2008 with the BCS cohort and find that without routinisation, the younger cohort are more upwardly mobile. However, the reverse is true when a 10% routinisation takes place as the older NCDS cohort is found to be much more upwardly mobile compared to the younger BCS cohort and therefore even though the younger cohort seems to have more occupational mobility, it is the older cohort who moves more when faced with technological change. Moreover, in this paper, Holmes et al. (2011) examine the likelihood of entry into routine jobs of their two cohorts and find no evidence that the fall in the numbers of younger workers entering routine jobs is any larger than the general decline in routine jobs.

Smith (2013) also points out that, even if polarisation has been well documented in the literature, the dynamics and associated drivers behind these employment trends as well as the fortunes of workers displaced from middle-skill jobs are not entirely understood. He uses US data from the matched monthly CPS, the March CPS supplement, and the Displaced Worker Survey to examine the employment states or occupations in which unemployed persons previously employed in low, middle, or high-skill occupations move to. He also examines the changes over time of transitions between job types and employment states and how these have contributed to trends in employment change within the various job levels. He finds that the fall in the employment share in middle-skill jobs is a result of both a drop in inflows into these jobs (mainly from non-employment and for younger workers) as well as due to an increase in outflows from these jobs (to non-employment and to other jobs). He also finds that the upsurge in the employment share in low and high-skill jobs seems to be due to an increase in worker transitions from other job types. For the low skilled he notes that this is apparent within all demographic groups while for the increase in the share of workers in high-skill jobs he notes that this is also compositional in nature due to the growth in the numbers of college-educated workers.

The present chapter attempts to contribute to this limited literature on job mobility using a different methodology and a different level of aggregation to the one used in the above-mentioned papers so as to isolate the causal effect of routinisation on moves from mid to either lower or higher-level jobs or out-of-employment. Given that panel data sets with the relevant variables are still scarce, this chapter offers a novel method of analysis at the cohort level, derived from individual level data and data information from different sources. This is clearly a contribution to the literature as it offers a new opportunity to study job mobility and related phenomena even in the absence of real panel data. Even though the results and contributions are at the cohort rather than the individual level, the conclusions from such analyses are very useful in observing the fate of different cohorts over the years and informing policy makers as to where to target their efforts.

4.3 Data, Methodology and Descriptive Statistics

4.3.1 Data

The Labour Force Survey cross sectional Cyprus files (CY-LFS) drawn from the European Labour Force Survey (EU-LFS) for the period 1999-2014 are the main data source for this chapter. In order to guarantee consistency and to ensure that each observation in the analysis of the present chapter corresponds to a different individual, only the second quarter file of each year is used. Detailed information about the EU-LFS can be found in Section 3.3.1 of Chapter 3.

As mentioned in Chapter 3, apart from personal and past and current labour-force characteristics, the CY-LFS anonymised micro data also contains occupation (ISCO) and industry (NACE) information, at the 3 and 1 digit level respectively, indispensable for creating the job variable following the jobs approach methodology. Data on the highest level of attained education by job as well as levels of employment by job are also drawn from the CY-LFS.

Unfortunately, the EU and CY-LFS wage data that are available from 2009 onwards, are only disseminated in income deciles, something that is not ideal for the purposes of studying job polarisation where a continuous and more detailed variable is preferred. For this reason, in order to analyse job polarisation when jobs are ranked according to wage group, income data from the 2005 and 2006 cross sectional EU-SILC files which report gross annual wages were merged with the LFS jobs data. A description of the EU/CY-SILC data set is available in Chapter 2.

4.3.2 Methodology

4.3.2.1 The Jobs Approach

The jobs approach, first originated by Nobel laureate Joseph Stiglitz and later refined by Wright and Dwyer (2003), involves characterising a job as a specific occupation in a specific sector using standardised international classifications. As noted in Eurofound (2013), while the sector depicts the type of economic value created, the structure of occupations indicates the way that this is being created and hence these two concepts of occupation and sector reveal the two principal dimensions of structural change.

According to Fernández-Macías (2012), the jobs approach allows the shift of the unit of analysis from the individual to the job level, perceiving jobs as specific occupations within specific sectors, so as to observe the change pattern of advanced economies' employment structures. Moreover, and according to the same author, "it allows the linking of quantitative information (on employment numbers by job) and qualitative information (on wage and educational levels by job), constructed through different processes and from different sources" (Fernández-Macías 2012, 11). Lastly, the use of a jobs-based methodology to identify net employment changes permits the linkage of existing empirical data directly to strategic policy commitments (Eurofound , 2013).

In practice, as was done in this chapter, the job variable is formed by creating a matrix with all possible combinations of occupation (defined by the ISCO classification) and sector (defined by the international NACE classification). In order to do so, only respondents in employment[82] during the reference week were kept in the sample. Given that the sector classification, NACE Rev.1.1 changes to NACE Rev. 2 in 2009 and the occupation classification, ISCO-88 changes to ISCO-08 in 2011, correspondence tables provided by the ILO were used in order to take NACE Rev. 2 back to Rev. 1.1 and ISCO-08 back to ISCO-88 and to guarantee consistency throughout the analysis[83]. Due to numerous small cells in the resulting jobs approach group variable, some of the problematic cells were rearranged consistently in all years so as to contain at least 20 observations each in every year. This was done by taking occupation back to the 2 or 1 digit level in cases of small cell size. For this reason, the final jobs approach job variable is a combination of industry at the 1 digit level and occupations at the 1, 2 and 3 digit levels. Categories in the jobs approach variable that still had fewer than 20 observations after the above rearrangement were dropped[84].

The resulting jobs were then ranked based on the highest level of educational attainment and wages and placed into ordered groups, in order to study the evolution in the number of workers, i.e. the employment share, in each of those job quality groups. In other words, after ranking jobs according to the variables described above and grouping them into education tiers and wage quintiles, the employment shifts and the change in the structure

[82] Belonging to one of the two categories: "Did any work for pay or profit during the reference week" or "Was not working but had a job or business from which he/she was absent during the reference week"
[83] Details as to how this was done can be found in Appendix 4C.
[84]More details as to how the jobs approach variable was constructed as well as a complete list of the resulting jobs can be found in Appendix 4B.

of jobs during the period of consideration were tracked so as to examine the shift of employment and to determine whether net employment growth has been concentrated in the top, middle or bottom of the employment structure (Eurofound, 2013) and hence whether job polarisation has indeed taken place in Cyprus during the period of consideration. This analysis of jobs permits the documentation of patterns of polarisation and how they have changed.

4.3.2.2 Calculating Employment Change when Jobs are Ranked According to Attained Level of Education

The first job characteristic based on which patterns of employment change are going to be ranked is the highest level of education or training successfully completed. In order to do this, cross sectional data from the first year of data i.e.1999 EU-LFS were used so as to calculate the modal level of education by job in the initial period of observation. To this end the derived education variable[85] available in the LFS was used which categorises attained education into 3 different levels: low, mid and high, based on the ISCED international education classification. Jobs were then ranked into tiers: Tier 1-Low level education jobs, Tier 2-Middle level education Jobs and Tier 3-High level education jobs.

4.3.2.3 Calculating Employment Change when Jobs are Ranked According to Wages

In order to study employment change when jobs are ranked according to their mean wage, data from the 2005 and 2006 EU-SILC cross sectional files were used. More specifically, two income variables from the EU-SILC were used in order to derive the mean wage by job. The first variable, PY010G, refers to the gross personal income, total and components at the personal level for employees. The second variable, PY050G, refers to cash benefits or losses from self-employment or self-employment income for the self-employed[86]. The reference period for both of the above variables is a twelve-month period and hence they refer to annual income. Even though hourly wages can be argued to be the best wage

[85] More details can be found in Appendix 4A

[86] This information as well as information in relation to other variables from the EU-SILC dataset can be accessed at the following link: http://ec.europa.eu/eurostat/web/income-and-living-conditions/methodology/list-variables

measure, so as to get an unbiased average at the job level irrespective of the share of part-timers in the job (Eurofound, 2013), they depend on having good hours data and unreliable hourly data could produce a worse wage measure than annual wages. In order to avoid noise in the reporting of working hours, an indicator of the hourly wage is not derived for the purposes of the present chapter, and the above variables are used as they are. For this reason, all people working part time are dropped. This is because part-timers' earnings would be influenced by how many hours they work, and so would not be an accurate indicator of the value of the job. Family workers were also dropped from the analysis, albeit very small in number.

The reason for measuring wages in the middle of the sample period rather than at the beginning, as is usual in the literature, is the fact that wage information in the EU-SILC only becomes available in 2005. Lastly, the reason for using two cross sections instead of just one is to ensure reliability of the results given that some of the resulting job cells in 2005 were small in size. For this reason and in order to have more reliable results derived from more observations, data from 2006 were also added in order to generate the mean wage by job. This mean wage by job data was then merged with the main dataset of this chapter, i.e. the LFS cross sectional files for the period 1999-2014[87]. After working out the mean wage by job, the next step was to group jobs into quintiles based on their mean wage in 2005 and 2006. More specifically, jobs were grouped into the lowest-paid 20 percent (quintile 1: low-level[88] jobs), the second lowest-paid 20 percent (quintile 2: low to mid-level jobs), the mid-paid 20 percent (quintile 3: mid-level jobs), the second highest-paid 20 percent (quintile 4: mid to high-level jobs) and the highest-paid 20 percent (quintile 5: high-level jobs) based on their mean wage and cell size in 2005 and 2006.

In both cases, i.e. when ranking jobs according to their initial education level in 1999 and according to the wage they were paying between 2005-06, the changes in the employment

[87] Given that occupation data in the EU-SILC is available at a higher level of aggregation than in the EU-LFS, i.e. ISCO-88 2 instead of 3 digits, a new matrix occupation-industry variable was created for the purposes of merging the two data sets together. This jobs approach variable consists of 80 categories compared to the original jobs approach variable, described earlier which contains 92 occupation-industry combinations. The wage information for these 80 different jobs from the EU-SILC was then merged with the more detailed job data (92 jobs) from the LFS. This means that, for the purposes of ranking jobs according to wage, those job cells in the EU-LFS that were made up of a 3 digit ISCO code for occupation, take the same 2 digit wage information derived from the EU-SILC, since this is the highest level of ISCO-88 aggregation available in this dataset. Even if this seems unfortunate, the level of information lost is not great, given that the job information in the EU-LFS was adjusted so as to deal with small cell size in some job cells which as a result meant that a great number of job cells ended up being made from ISCO at the 2 and 1 digit rather than the 3 digit level.

[88] Low-level, low-to-mid-level etc. and low-paid, low-to-mid-paid etc. are used interchangeably in the present chapter. For example, low-level and low-paid refer to the same jobs. Similarly, low-to-mid level and low-to-mid-paid etc.

shares of jobs within the various job levels are computed– i.e. the numbers of individuals in 1999 who have jobs that are in each tier or quintile of the education and wage distribution respectively are compared to the number of individuals in the same jobs in 2014[89].

4.3.2.4 Pseudo Cohorts

In order to be able to study job mobility of displaced mid-level workers over time, a number of pseudo cohorts were constructed. Pseudo cohorts can be defined as "artificially created data sets constructed from repeated cross-sections" (McIntosh 2005, 3). Even though the survey respondents at each point in time differ, they are still representative of the full cohort in the population. For example, and as noted by McIntosh (2005), even if the actual 25-29 year olds interviewed in 2004 will be different to those 20-24 year olds interviewed in 1999, the fact that the dataset used is nationally representative, means that the resulting pseudo cohort will be representative of the true cohort of this age in the population as a whole and estimates of the proportions at each qualification level employing data from these pseudo cohorts will be unbiased estimates of the real proportions in the actual cohort of this age (McIntosh, 2005).

Comparing repeated cross-sectional data to genuine panel data, Verbeek (2008) notes that the major drawback of the former is that the same individuals are not followed over time and hence individual histories cannot be incorporated in a model. In contrast, repeated cross-sections are less prone to panel dataset issues such as attrition and non-response and are usually considerably larger, both in numbers of respondents and in the time period they cover. The requirements for modelling transitions from a repeated cross-sectional data set using pseudo cohorts are that each individual should be member of exactly one cohort which does not vary with time and that the variables chosen to define cohorts are observed for all individuals in the sample (Verbeek, 2008). According to Verbeek (2008), the grouping variables used have to satisfy the applicable conditions for an instrumental variable to be consistent (Verbeek 2008, 370). In other words, the variable(s) based upon

[89] Using a fixed reference point for defining job quality means that results are conditional on the observed quality at the start of the studied period (in the case of the education ranking) and in the middle of the studied period (in the case of wages). This means that the assumption that job quality does not evolve throughout the studied period is made as is common in the polarisation literature (e.g. Goos and Manning, 2007).

which the cohorts are going to be defined have to be exogenous or pre-determined and relevant.

In the present chapter, the use of pseudo cohorts enables the analysis of job mobility and career progression of the middle-level workers following the restructuring of employment in the past years. More specifically, job mobility from mid-level jobs towards other job levels as well as out-of-employment is tested by using a cross sectional dataset that is pooled together at the cohort level, sorting each cohort by age group and education level and observed every five years starting in 1999 up to 2014, i.e. at four different points in time. For the purpose of the present chapter, eighteen cohorts are defined in 1999 based on five year age bundles and three levels of attained education[90]. Given that the EU-LFS anonymised data delivers the age variable in five year bundles instead of a continuous variable, these eighteen cohorts are observed every five years. More specifically, the cohorts are defined as follows and are each observed in 1999, 2004, 2009 and 2014: Cohorts 1, 2 and 3 aged 20-24 in 1999 with low, mid and high level education respectively; Cohorts 4, 5 and 6 aged 25-29 in 1999 with low, mid and high-level education respectively; Cohorts 7, 8 and 9 aged 30-34 in 1999 with low, mid and high-level education respectively; Cohorts 10, 11 and 12 aged 35-39 in 1999 with low, mid and high-level education respectively; Cohorts 13, 14 and 15 aged 40-45 in 1999 with low, mid and high-level education respectively and Cohorts 16,17 and 18 aged 45-49 in 1999 with low, mid and high-level education. To give an example, Cohort 1 aged 20-24 with low level education in 1999 will be aged 25-29 in 2004, 30-34 in 2009 and 35-39 in 2014. Similarly, Cohort 4 low educated and aged 25-29 in 1999 will be aged 30-34 in 2004, 35-39 in 2009 and 40-45 in 2014 and so on while the education level based upon on which the cohorts are defined remains unchanged throughout the years of analysis.

Aggregating all observations at the cohort level, the resulting model for the effect of a change in the proportion in mid-level jobs on the change in the proportion of a cohort in jobs at the other levels or out-of-employment can be written as:

$$\mathbf{\overline{Y}_{ct} = \overline{X}_{ct}\boldsymbol{\beta} + \overline{u}_{ct}} \qquad \mathbf{c = 1, \ldots, C; \quad t = 1, \ldots, T,} \qquad \mathbf{(4.1)}$$

[90] The LFS derived aggregated three-level education variable, also used in defining job change when jobs are ranked by education, is used here.

where $\overline{Y}_{ct}$ is the average value of all observed y_{it}'s in cohort c in period t, and similarly for the other variables in the model[91]. This creates a pseudo panel or synthetic panel with recurrent observations over T periods and C cohorts.

4.3.2.5 Instrumental Variable (IV) Regressions

Linear pseudo cohort panel regressions like the one in (4.1) above, demonstrate the flows of workers towards the various job groups, but do not prove causation. In other words, even if they establish whether the relationship between the proportion of workers in mid-level jobs is negatively or positively related to the proportion of workers in low or high-level jobs, for example by looking at the coefficient of the mid-level group in the regression in (4.1) above, this does not prove that the change in mid-level jobs is responsible for the change in the proportion of low or high-level jobs. For example, if a negative relation between the change in mid-level jobs and the change in the proportion in low-level jobs is found, there could be a reverse causality whereas the above-mentioned relation between the two is caused by low-level workers progressing to mid-level jobs and not by displaced mid-level workers forced to move down to low-level jobs as a result of routinisation. In other words, changes in the share in level 1 low-level jobs can show up as a change in the share at level 3 mid-level jobs. Therefore, it could be argued that the specification in (4.1) above suffers from a simultaneous causality bias. In other words, changes in the explanatory variable, $\overline{X}_{ct}$, the proportion of a cohort employed in mid-level jobs, could either lead to changes in the proportion of a cohort in low-level jobs, low-to-mid-level jobs, mid-to-high level jobs or high-level jobs or be a result of such changes. Hence, it is not possible to be certain that the observed relation between $\overline{X}_{ct}$ and $\overline{Y}_{ct}$ is a causal one as $\overline{X}_{ct}$ could be causing changes in $\overline{Y}_{ct}$ and vice versa. Such concerns can be eliminated via the use of Instrumental variables (IV).

As explained in the previous chapter, the idea behind using IVs is to identify a random exogenous variation in the explanatory variable of interest, which in this case is the change in the proportion of a cohort in mid-level jobs (the proportion in mid-level jobs at time t minus the proportion at time t-1). To create such an instrument, task data from Goos et al. (2009) [92] were used to construct a 'routinisation' variable referring to the prior

[91] These variables are discussed in detail in section 4.3.3 below
[92] See Appendix for details 4F

proportion of a cohort in routine jobs or in jobs that score high in routine task importance. More specifically, each occupation was assigned a routine task score and the lag of the proportion of each cohort in the various occupations, ranked according to this score, was entered as an IV for the change in the proportion of a cohort in mid-level jobs. In other words, the endogenous variable, in this case the change in the proportion of a cohort in mid-level jobs, was instrumented by this newly created measure of routinisation which is assumed to be exogenously determined sometime in the past and is used to identify exogenous change in the share of mid-level jobs this period.

It is expected that, due to routinisation, the proportion of a cohort in mid-level jobs will change more when the cohort are more involved in routine jobs in the previous period, i.e. it is expected that there is going to be a greater fall in mid-level jobs for those cohorts where there was a higher proportion in jobs involving routine tasks in the previous period, while this is not expected to have an effect on the proportions of the different cohorts in other job levels or out-of-employment. This follows from the TBTC theory of polarisation discussed earlier which stresses the substitutability between routine tasks and technology (Autor et al., 2003) and which postulates that the more routine tasks a job involves, the more likely it is to be computerised (and routine tasks are typically found within jobs in the middle of the job distribution). In other words, if an exogenous fall in the proportion of people in mid-level jobs takes place solely due to the fact that they happened to be working in jobs with a high routine task intensity (i.e. due to routinisation) and if this random change in the proportion in mid-level jobs is still negatively or positively associated with the change in the proportion of a cohort in the other job groups or those out of employment i.e. the dependent variable of interest in each regression, this would provide support for the argument that changes in the former lead to changes in the latter and to establish causation of the flows from mid to lower or higher-level jobs or out-of - employment.

Linear IV regressions are then run at the cohort level looking at the effect of changing employment in the middle group, i.e. mid-level jobs, on the proportion of the cohort in the low and high ends of the distribution of jobs as well as the effect on those not in employment. It follows that, if a significant effect of the change in the proportion in mid-level jobs on the change in the proportion in other job groups is still observed, then this means that a causal effect exists i.e. since the focus is on that part of the variation in the proportion of mid-level jobs that is exogenously determined by the routine nature of the tasks involved in jobs in that part of the distribution. Moreover, via the use of IV regressions, the risk of capturing moves of mid-level workers due to career motivations (for example upward mobility as a result of promotion) rather than moves specifically due to the routinisation of job tasks, present in the case of a mere observation of the occupational mobility of routine workers over time, is also eliminated.

As discussed in Chapter 3, the second stage IV regression of the "Two Stage Least Squares" (TSLS or 2SLS) estimator is the following:

$$\mathbf{Y_i = \beta_0 + \beta_1 \hat{X}_i + \varepsilon_i} \qquad \textbf{(4.2)}$$

In the present chapter, given the fact that the data set is a repeated cross section pooled together to form pseudo cohorts, observations for N individuals exist in each year. These individuals are not the same every year and therefore i does not range from 1 to N in each t as in the case of genuine panel data where the same individuals are followed over a number of t periods. Therefore, the regression in (4.2) above takes the following form

$$\mathbf{\overline{Y}_{ct} = \beta_0 + \beta_1 \hat{\overline{X}}_{ct} + \varepsilon_{ct}} \qquad \mathbf{c = 1, \ldots, C; \quad t = 1, \ldots, T,} \qquad \textbf{(4.3)}$$

Where $\mathbf{\overline{Y}_{ct}}$ and $\mathbf{\hat{\overline{X}}_{ct}}$ are the change in Y and X from one year to the next and $\hat{\overline{X}}_{ct}$represents a K-dimensional vector of other explanatory variables measured at both time t and t-1 and β is the associated vector of coefficients. The subscripts c and t are the cohort and the year that each cohort is observed. C=18 and T=4 in this case.

A detailed discussion in relation to Instrumental Variable regressions can be found in Section 3.4.2 of Chapter 3.

4.3.3 Descriptive Statistics of variables used in Job Mobility regressions

Besides the proportion of people in mid-level jobs, Equation (4.3) controls for a range of additional variables that relate to flows in and out of jobs. More specifically, in this specification, the vector of control variables $\widehat{\widehat{X}}_{ct}$ contains the following information: the proportion of each cohort under temporary employment contracts (temporary $\text{contracts}_{i,t}$); the proportion of people in part time jobs ($\text{part-time}_{i,t}$); the proportion of females ($\text{female}_{i,t}$); the 5-year age group (age $\text{group}_{i,t}$) and the cohort group's education level, ($\text{education}_{r,i,t}$) available in the LFS. Moreover, in order to establish whether being over-educated affects whether a person previously working in a job in the middle of the wage distribution will end up in a low or a high-level job, a dummy variable equal to 1 for those whose education level is above the modal education in their job and equal to zero otherwise is generated and variables corresponding to the proportion of over-educated workers in the cohort are entered into the regression ($\text{overeducated}_{i,t}$)[93].

Table 4.1 below provides summary statistics of the explanatory variables used in the regression analysis.

[93] Descriptions of all the variables used in this chapter can be found in Appendix 4A

Table 4.1: Descriptive and Summary Statistics

	Mean	Standard Deviation
Proportion in low-level jobs	0.228	0.130
Proportion in low-to-mid-level jobs	0.194	0.087
Proportion in mid-level jobs	0.213	0.107
Proportion in mid-to-high level jobs	0.185	0.083
Proportion in high-level jobs	0.180	0.223
Proportion out-of-employment	0.261	0.142
Age Group		
20-24	0.042	0.201
25-29	0.083	0.278
30-34	0.125	0.333
35-39	0.167	0.375
40-44	0.167	0.375
45-49	0.167	0.375
50-54	0.125	0.333
55-59	0.083	0.278
60-64	0.042	0.201
Female	0.470	0.096
Education		
Highly Educated	0.333	0.475
Medium level education	0.333	0.475
Low Educated	0.333	0.475
Over-education	0.158	0.139
Temporary contracts	0.120	0.098
Part time work	0.070	0.048

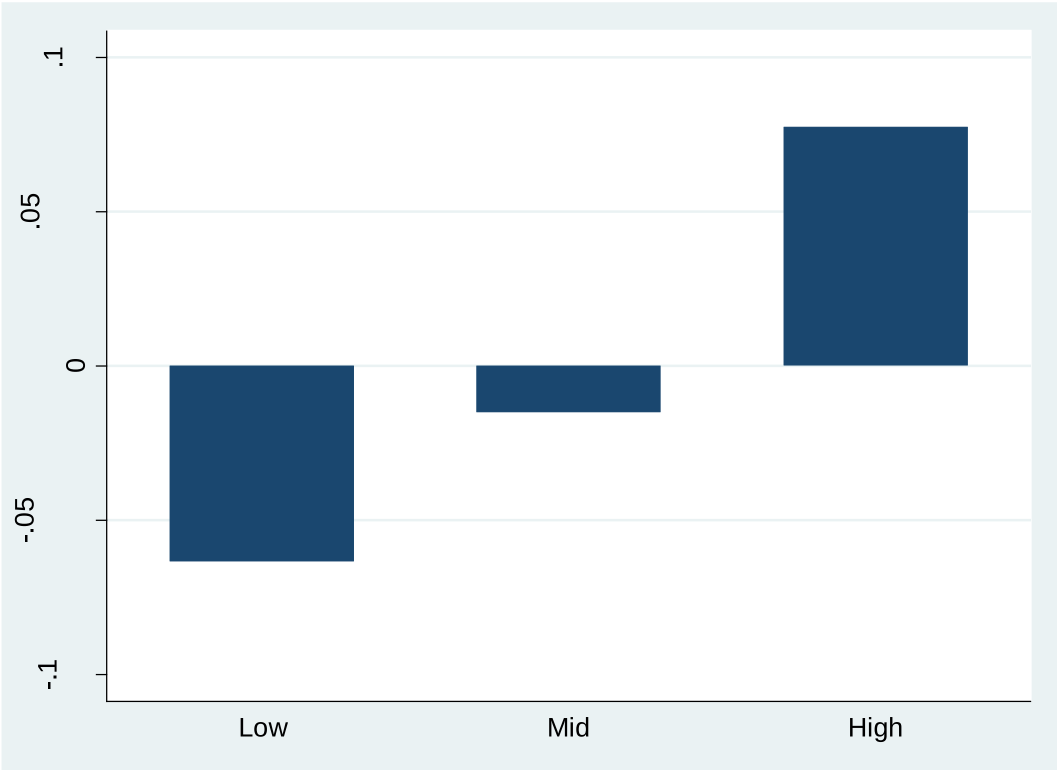
Note: The proportions in the various job levels are calculated based on the employee sample only, whereas the proportion of people out of employment is based on the full sample (employed plus out-of employment)

4.4 Results

4.4.1 Trends in Job Change

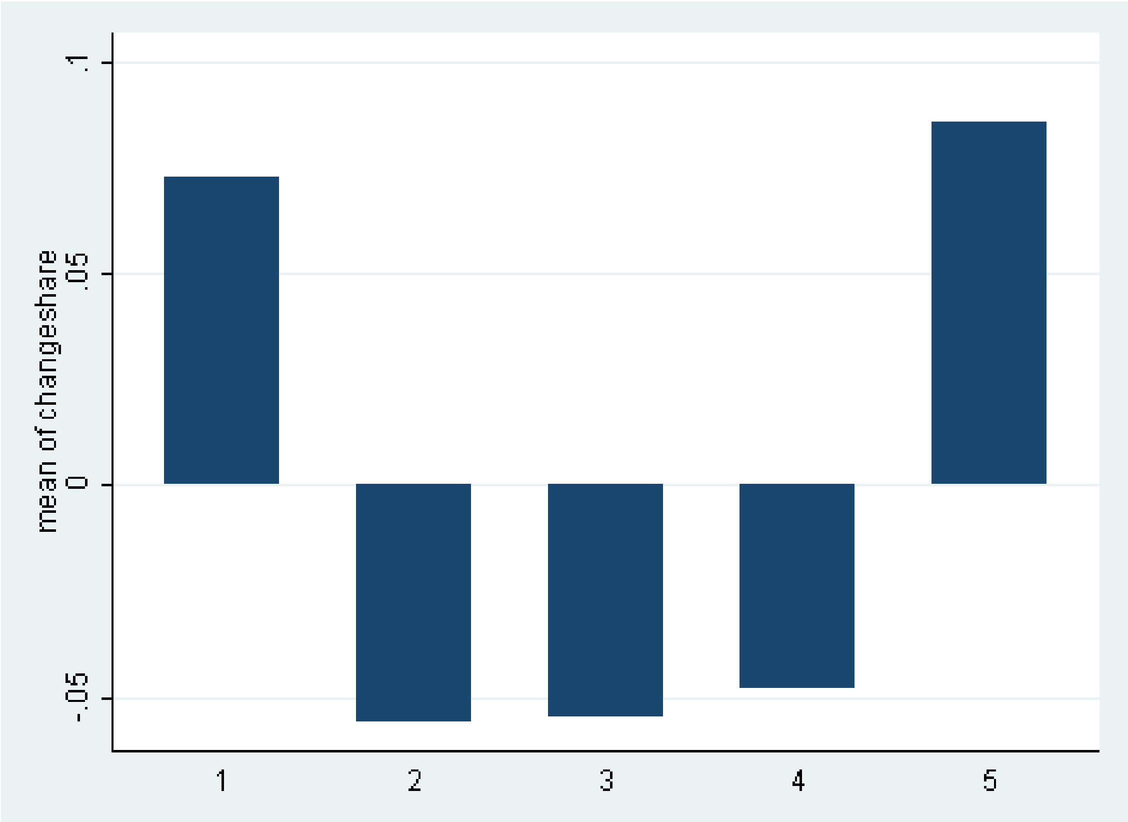
The below graphs demonstrate the change in the employment share by job when jobs are ranked according to: 1) education tiers and 2) wage quintiles for the period 1999-2014.

Figure 4.1: Mean Change in the Employment Share when Jobs are ranked according to the Modal Education Level of Job Holders for the Period 1999-2014.



As depicted in Figure 4.1, when jobs are ranked according to education tiers, a pattern of clear upgrading seems to have taken place in Cyprus between the period of 1999 and 2014. More specifically, when jobs are ranked according to their initial education level in 1999, high-level jobs i.e. those jobs that required a high level of education in 1999 have increased their employment share. At the same time, jobs with a low and mid modal education attainment level saw a drop in their employment shares, with low-level jobs showing a larger fall compared to mid-level jobs. These trends are in line with the SBTC theory discussed earlier which postulates that due to their complementarity with technology, high-skill jobs are the ones expected to increase their employment share the most following advances in technological change. At the same time, they are also in line with the skill evolution on the labour supply side and the rise in the overall education level of the population over the past years, discussed in Chapters 1 and 2.

Figure 4.2: Mean Change in the Employment Share when Jobs are ranked according to the Mean Wage of Job Holders for the period 1999-2014.



On the other hand, when jobs are ranked according to the mean wage they were paying between 2005 and 2006, as presented in Figure 4.2, a pattern of clear polarisation is evident. This confirms that, when jobs are ranked according to mean wage and similar to numerous industrialised countries, job polarisation has also taken place in Cyprus, a finding in line with the TBTC theory. The routinisation hypothesis, can potentially explain the observed hollowing-out in the middle of the job distribution in Figure 4.2 as mid-level jobs are the ones most likely to be substituted by technology as they score high in terms of routine task intensity compared to jobs at the lower and higher ends of the jobs distribution. Another important observation from Figure 4.2 above, is that the increase in high-level jobs is slightly overcoming the increase in low-level jobs. This means that, over the fifteen year period under examination here, more better- rather than worse-paying jobs have been created. Moreover, it seems that the low-to-mid-level group is the one that loses the largest employment share compared to the other middle groups yet this difference is very small especially when compared to the mid-level job group[94].

The above findings are consistent with similar studies of the Eurofound (e.g. Eurofound 2013; Eurofound 2015). Nevertheless, there are a number of differences that distinguish the present study from the above. Specifically, when ranking jobs according to the attained education of job holders, the present chapter uses the modal level of education instead of the mean. Moreover, when jobs are ranked according to wage, a different and single data source is used, i.e. the EU-SILC, whereas Eurofound (2013) and Eurofound (2015) use wage data from multiple other EU-level sources. Lastly, the present chapter uses correspondence tables from the International Labour Organisation (ILO) to ensure continuity in the occupation ISCO and industry NACE classifications given that they change during the period of analysis whereas the above mentioned studies break down the analysis based on these changes in the classifications, something that does not allow an undisturbed comparison of trends in job polarisation. In other words, by using two separate time periods to examine polarisation trends because of the change in the ISCO variable, it becomes difficult to establish whether changes in the distribution of jobs are due to the change in the classifications of occupations or due to polarisation. The period under examination is also slightly different with Eurofound (2013) running over a shorter period and not including 2013 which was the year of the collapse of the Cypriot banking system.

As evident in the above figures, and even though education and wages are oftentimes used as correlates or proxies of job quality, there is a clear discrepancy in employment change when jobs are ranked according to the two methods. This finding is not unusual in the literature as a similar result was also found in Eurofound (2013), with the authors noting that the reason for this is that a significant share of jobs in the middle of the wage distribution (in EU countries) have a higher comparative position in terms of wages than education. According to the authors, a great number of these jobs (those in manufacturing) remained stagnant in the 1995-2007 expansion period and numerous have been destroyed in the following recession resulting in wage polarisation and a depression of the bottom of the education structure. Similarly, descriptive analysis by industry in this case revealed that this divergence between employment change trends when jobs are ranked according

[94] A trend of job polarisation also exists when the three middle groups in figure 4.2 (low-to-mid, mid and mid-to-high level) are pooled together to form one single middle group.

to education vs wages is in large part due to the fact that mid-paid occupations within the various industries appear to be of a lower education level than those in the middle of the educational distribution.

More specifically, occupations within the Agriculture, forestry, fishing and hunting industry that are found at the lowest end of the attained education distribution occupy a substantial share of employment in the middle of the wage distribution when jobs are ranked according to mean wage. In the same way, in Manufacturing, even though Technicians and associate professionals and Clerks are in both cases found in the middle of the job distribution, more low-level occupations falling within Craft and related trades workers, Plant and machine operators and assemblers and Elementary occupations, that have a low educational ranking, seem to have a higher mean wage. In other words, even though these occupations occupy the lower end of the job distribution when they are ranked according to education, they are paid a higher wage that places them in the middle of the wage distribution. This is also the case in Construction and to a lesser extent in terms of numbers in Transport, storage and communication and Public administration and defence. The two rankings are completely different when it comes to Real estate where mid-paid workers are found to be exclusively Technicians and associate professionals whereas they are found to be Clerks in mid-educated jobs. Similarly, in the Other community social and personal service activities industry, where mid-educated occupations are concentrated solely within clerical posts, mid-paid jobs are also spread across higher level ISCO-88 Group 2 (Professionals) and ISCO-88 Group 3 Technicians and associate professionals. In all other industries, the occupations in both rankings and employment shares are more or less the same. The above trends apply to all years of analysis.

All in all, what can be said is that discrepancies in employment change patterns when jobs are ranked according to education vs wages seem to be explained mostly by large differences in the composition of employment within the Agriculture hunting and forestry and fishing; Construction and Manufacturing industries and to a lesser extent within Transport, storage and communication and Other community social and personal service. More specifically, in the above industries, occupations that are considered as low-level in terms of attained education are found in the middle of the wage distribution. Therefore, even though found at different points of the education and wage distributions, the jobs whose employment share declines are the same in both cases, leading to an upgrading pattern in terms of attained education and to a polarisation pattern in terms of wages. Given that workers in such jobs, like for example skilled production workers who are employed in jobs in the middle of the job distribution, might not have very high levels of formal education but may acquire their job specific skills through work experience and on-the-job training, mean wage is deemed as a better proxy of skill than education, for these jobs. For this reason and given that the purpose of this chapter is to shed light on the declining middle part of the job distribution, the rest of the chapter proxies skill/job quality using wages rather than education.

Taking a closer look at changes in the jobs distribution when jobs are ranked according to wages confirms that, similar to European analyses, Cyprus has also experienced ongoing shifts away from the primary sector (particularly agriculture) and manufacturing industries and towards services and the knowledge intensive economy with the addition of shifts away from the construction industry[95]. The mid-level jobs that are found to have experienced declines from the beginning to the end of the examination period are Animal producers and related workers in Agriculture; Precision, handicraft, craft printing in Manufacturing; Building frame and related trades and Building finishers and related trades in Construction; Electrical and electronic equipment workers in Wholesale and Retail Trade as well as Customer service clerks in Transport.

For the low to mid-level jobs, falls in employment shares are mostly concentrated within Manufacturing, in jobs such as Wood treaters, Machine operators and Manufacturing labourers as well as in Elementary occupations within Construction and Public administration and defence. Lastly, Technicians within Health and social work also experience falls in their employment shares. For mid-to-high level jobs, falls in the proportion of people employed have been concentrated within Manufacturing in jobs like Technicians and associate professionals, within Transport storage and communications in jobs such as Drivers and mobile plant operators and Elementary occupations as well as within Financial intermediation in jobs such as Office and numerical clerks. Falls in the proportion of people employed in mid-to high-level jobs can also be found in Protective services workers within Public administration and defence as well as in Technicians within Health and social work.

[95] This is consistent with the overview of the changing industrial structure for Cyprus described in Chapter 1.

At the lower end of the job distribution, the greatest increase in employment share in low-level jobs has been in Domestic and related helpers, cleaners and launders within the Activities of private households as employers and undifferentiated production activities of private households sector. Smaller increases can also be found in Customer services clerks within Wholesale and retail trade and within Health and social work as well as in Elementary occupations within Other community, social and personal services and within Real estate and Education. Service workers and Shop and market sales workers within education have also increased their employment shares.

Lastly, in terms of the composition of high-level jobs that increase their employment proportions, these are more dispersed in terms of individual industries and more concentrated in terms of occupation. Nevertheless, they are all concentrated within the service sector. For example, Legislators and senior officials within Hotels and restaurants, Transport and Real estate; Professionals within Transport storage and communication, Business professionals in Real estate, Other teaching professional in Education as well as Nursing and midwifery professionals within Health and social work.

Given that the data in this chapter encompass both the pre-recession expansion and recession period, the 1999-2014 period was also broken down into two separate periods and the above analysis when jobs are ranked according to mean wage was replicated in both periods so as to observe whether the patterns of structural employment change differ in expansions and recessions and whether economic crises create job polarisation (Eurofound, 2013). The results from this analysis[96] demonstrated that polarisation is not limited to the recession period[97].This means that treating the period as a whole, and looking for polarisation patterns across the full period is justified in the present chapter. This is the case (polarisation pre- and post-recession), whether the period is split by the time of recession in most of Europe[98] as well as by the time of the recession in Cyprus itself which saw its peak in 2013 with the collapse of its banking system. Lastly, polarisation when jobs are ranked according to wage is also present across all age groups with very small variations as demonstrated in the various graphs in Appendix 4E.

[96] These are available in Appendix 4D

[97] This finding is in accordance with Eurofound's (2013) country level results for Cyprus, which find that for the periods 1999-2007 and 2008-2010, Cyprus was among the countries that experienced more or less a clear polarisation process.

[98] Eurofound (2013) breaks down their 1995-2012 period into: a) the pre-recession employment expansion (1995-2007); b) the Great Recession (2008-2010) and c) the stalled recovery (2011-2012). According to the same source, in Cyprus, sharp falls in employment were reported in the period from 2011 to 2012 and hence this period is also marked as one of recession instead of recovery for Cyprus.

4.4.2 Patterns of Employment by Age Group and Cohort over Time

Having shown above the overall polarisation trends when jobs are ranked according to wages, this section is a first attempt to show descriptively where people end up working and whether younger vs older age groups and cohorts are now found in greater numbers in different wage quintiles of the jobs distribution, compared to previously. The tables that follow show the raw proportions of the relevant group found in the low, mid and high[99] wage quintile, by age group as well as broken down by age group and education (i.e. cohort, as defined in this chapter) for each reference year so as to observe how the different cohorts have changed their shares across the various wage quintiles over the years of analysis[100]. In the same manner, the raw proportions of those out of employment are also presented. Here, just like in the job polarisation graphs presented above, jobs are consistently ranked to one of the wage quintiles based on wage information derived from the EU-SILC in years 2005 and 2006 and hence it is not the case that the observed changes in the tables that follow are due to movements of jobs into new quintiles due to relative wage changes. In other words, the tables below present genuine flows of workers across a pre-defined set of jobs, ranked according to the mean wage they were paying in 2005-06. It also has to noted that the various age and/or education groups whose proportions in the different job levels at different points in time are being analysed in this section refer to different groups of people who are the same age and/or within the same educational category at different points in time (i.e. comparing different stocks) rather than discussing changes over time for given groups of people (i.e. flows).

[99] The low-to-mid-level and mid-to-high-level job groups are omitted from the analysis in this section and that is why the proportions across the different job levels do not sum up to 1.
[100] Even though all 15 years are available, the proportions are only shown at these four distinct time periods to match the previous analysis in this chapter.

Table 4.2(a): Proportions of People in Low-Level[101] Jobs (All Education Groups)

	1999	2004	2009	2014
20-24	0.133	0.123	0.140	0.140
25-29	0.109	0.158	0.153	0.189
30-34	0.091	0.168	0.176	0.171
35-39	0.128	0.182	0.206	0.186
40-44	0.098	0.191	0.184	0.166
45-49	0.104	0.144	0.177	0.191
50-54	0.088	0.135	0.133	0.11
55-59	0.06	0.101	0.11	0.115
60-64	0.064	0.059	0.091	0.071

Table 4.2(b): Proportions of High-Educated People in Low-Level jobs

	1999	2004	2009	2014
20-24	0.092	0.113	0.151	0.139
25-29	0.071	0.091	0.105	0.155
30-34	0.066	0.073	0.089	0.085
35-39	0.047	0.098	0.092	0.105
40-44	0.022	0.109	0.138	0.072
45-49	0.031	0.087	0.118	0.063
50-54	0	0.051	0.089	0.052
55-59	0.024	0.07	0.044	0.059
60-64	0	0.016	0.021	0.041

Table 4.2(c): Proportions of Mid-Educated People in Low-Level Jobs

	1999	2004	2009	2014
20-24	0.140	0.127	0.126	0.116
25-29	0.117	0.223	0.229	0.192
30-34	0.119	0.205	0.185	0.226
35-39	0.135	0.221	0.234	0.218
40-44	0.112	0.189	0.178	0.195
45-49	0.113	0.155	0.168	0.245
50-54	0.077	0.14	0.158	0.131
55-59	0.025	0.071	0.112	0.148
60-64	0.065	0.042	0.095	0.071

Table 4.2 (d): Proportions of Low-Educated People in Low-Level Jobs

	1999	2004	2009	2014
20-24	0.157	0.121	0.183	0.286
25-29	0.146	0.197	0.15	0.341
30-34	0.067	0.288	0.384	0.339
35-39	0.202	0.218	0.419	0.360
40-44	0.138	0.272	0.272	0.312
45-49	0.128	0.177	0.236	0.234
50-54	0.121	0.169	0.143	0.152
55-59	0.081	0.124	0.143	0.129
60-64	0.073	0.073	0.11	0.086

[101] Low level jobs refer to jobs in the lowest pay quintile

The first table above, Table 4.2 (a), demonstrates the overall trends in the proportions of each age group in low-paid jobs over the 15 year period under consideration as observed once every 5 years. Comparing the first and last year of observation, all age groups have increased their proportion in low-paid jobs between 1999 and 2014 with small variations up and down from one period to the next. For example, the youngest age group, has increased from 13.3% in 1999 to 14% in 2014 with a small drop to 12.3% in 2004 while for example the proportion of 25-29 olds in low-paid jobs has followed a steeper increase from 11% in 1999 to 18.9% in 2014 with a very small drop of 0.5 percentage points between 2004 and 2009. All in all, the largest increase in the proportion in low-paid jobs seems to be exhibited by the 45-49 age group with an 8.7 percentage points increase in the proportion in low-paid jobs between 1999 and 2014, closely followed by the 25-29 and 30-34 age groups with an increase of 8 percentage points from 1999 to 2014. The smallest increase in the proportion of people in low-paid jobs has been experienced by the youngest and oldest age groups with a mere 0.7 percentage points increase between the first and last year of observation. Looking at the above table vertically, it can be seen that there is no single age group that has the greatest proportion in low level jobs in all years as this changes in every period of observation.

The three tables that follow, i.e. tables 4.2 (b)-(d), break down the low-level job group into high, mid and low educated sub-groups so as to observe how the proportions of each age group across the years of analysis vary for the various education groups. In this manner the variations in the employment positions of each cohort as defined by age group and education can also be observed.

Firstly, looking at the overall trends of high-educated workers in low-level jobs, i.e. Table 4.2(b), it seems that all age groups have increased their shares in low-level jobs across the four periods of observation. Mid-educated workers have also increased their shares in low-level jobs, as depicted in Table 4.2(c), with the exception of the 20-24 age group that slightly decreased their share. Similarly, low-educated workers of all age groups (Table 4.2(d)) have increased their shares in low-level jobs. This is reflective of the overall trend in the jobs distribution, where low-level jobs increased in numbers and hence employment in these jobs goes up for all education groups but mostly for low and mid-educated workers who occupy these jobs in larger numbers as can be seen by the absolute numbers of the proportions in the above three tables.

Looking vertically at Table 4.2(b), the youngest high-educated age group is more likely to be employed in low-level jobs compared to all other age groups, in all years except in 2014 when the proportion of the 25-29 age group is slightly higher than the proportion of 20-24 year olds. For the mid-educated group, there does not seem to be a persistent trend in terms of a specific age group occupying low-level jobs in higher numbers. Lastly, in terms of the low-educated, the 35-39 group occupies the highest proportion in low-level jobs in all years except in 2004 when the 30-34 age group is represented in higher numbers in low-level jobs.

Looking diagonally at the above tables, the proportions of the high, mid and low-educated cohorts aged 20-24, 25-29,30-34,35-39, 40-44 and 45-49 years of age in 1999 in low-level jobs can be followed across all years of analysis. Firstly, comparing the first and the last year of observation, all high educated cohorts except the cohort aged 30-34 in 1999 experience a very slight increase in their proportion in low-level jobs, probably reflecting the fact that once they enter low-level jobs it is difficult to escape following the findings of Chapter 2. However, taking the in-between fluctuations into account it can be seen that, with the exception of the youngest cohort, the proportions of all other cohorts in low-level jobs go up substantially, before falling again by 2014 –probably reflecting the fact that older high-educated workers are less likely to be in low-level jobs in general (e.g. they may choose to retire earlier rather than stay in such jobs long term). The increasing trend in the proportions employed from the first to the last year of data is also true for mid-educated cohorts, however the increases in their proportions in low-level jobs are larger than in the case of their highly educated counterparts. In the case of low-educated cohorts, again their shares in low-level jobs increase to a larger extent than mid and high-educated cohorts, except in the case of the cohorts aged 35-39 and 40-44 in 1999, where even though they experience an initial increase between 1999 and 2009 they then fall again by 2014, possibly reflecting the lower proportions of older workers in such jobs, perhaps due to early retirement, since both of these cohorts are aged 50+ by 2014.Therefore, looking diagonally and following individual cohorts, yet different stocks of workers at the same age and with the same level of education, it looks like the proportion of people in low-paid jobs does not fall much as they get older, at least until the oldest working age categories.

Table 4.3(a):
Proportions of People in Mid-Paid Jobs (All Education Groups)

	1999	2004	2009	2014
20-24	0.166	0.166	0.148	0.041
25-29	0.165	0.165	0.154	0.086
30-34	0.174	0.148	0.144	0.119
35-39	0.150	0.157	0.144	0.141
40-44	0.175	0.159	0.150	0.112
45-49	0.151	0.160	0.151	0.133
50-54	0.112	0.169	0.163	0.132
55-59	0.140	0.125	0.122	0.096
60-64	0.077	0.127	0.101	0.069

Table 4.3(b):
Proportions of High-Educated People in Mid-Level Jobs

	1999	2004	2009	2014
20-24	0.092	0.076	0.079	0.022
25-29	0.097	0.065	0.064	0.064
30-34	0.103	0.095	0.058	0.077
35-39	0.084	0.08	0.089	0.082
40-44	0.086	0.079	0.052	0.061
45-49	0.076	0.067	0.062	0.073
50-54	0.010	0.080	0.052	0.073
55-59	0	0.105	0.044	0.048
60-64	0.044	0.066	0.052	0.041

Table 4.3(c):
Proportions of Mid-Educated People in Mid-Level Jobs

	1999	2004	2009	2014
20-24	0.149	0.127	0.163	0.053
25-29	0.172	0.234	0.271	0.115
30-34	0.211	0.178	0.248	0.173
35-39	0.171	0.160	0.205	0.225
40-44	0.194	0.183	0.184	0.136
45-49	0.118	0.159	0.188	0.169
50-54	0.071	0.190	0.171	0.139
55-59	0.1	0.101	0.096	0.121
60-64	0.032	0.095	0.061	0.055

Table 4.3(d):
Proportions of Low-Educated People in Mid-Level Jobs

	1999	2004	2009	2014
20-24	0.306	0.329	0.233	0.048
25-29	0.254	0.268	0.206	0.110
30-34	0.195	0.194	0.143	0.152
35-39	0.187	0.253	0.129	0.123
40-44	0.219	0.202	0.221	0.172
45-49	0.203	0.231	0.173	0.128
50-54	0.164	0.195	0.248	0.196
55-59	0.179	0.142	0.184	0.109
60-64	0.096	0.149	0.134	0.093

In terms of mid-level jobs, as demonstrated by Table 4.3 (a) above and as expected given the job polarisation trend demonstrated in the previous section, there was an overall drop in the proportions employed in mid-paid jobs for all age groups except the 50-54 age group who slightly increased their share in such jobs. The largest drop in the proportion of people in mid-paid jobs between the first and last year of observation was experienced by the youngest age group and the smallest drop by the oldest age group.

Looking at Table 4.3(b) set against Table 4.3 (d), it can be seen that the proportion of the two youngest age groups in mid-level jobs declined to a much smaller extent when they possessed a high rather than a low level of education, yet the proportion could not drop as much for the higher educated given that it was lower to start with. More specifically, high educated 20-24 years old see a decline of just 0.07 in mid-level jobs from the first to the last year of observation compared to low-educated 20-24 year olds whose proportion in mid-level jobs declines by 0.258. Similarly, the proportion of highly educated 25-29 year olds declines by 0.033 compared to 0.144 for low educated 25-29 year olds. Another interesting point from Tables 4.3(b) and 4.3(d) is that the proportion of the oldest age group in mid-level jobs declines by exactly the same amount irrespective of the level of education. This could possibly be explained by the fact that education is more important for young age groups who do not possess work experience while the opposite is true for older age groups whose work experience may play an important role in the case of declining mid-level jobs. For example, 20-24 year olds can use their high education level as an advantage and enter or stay in mid-paid jobs during a time of polarisation when the number of these jobs declines, whereas older age groups are less affected by polarisation irrespective of their education level as what matters is their on-the job experience that puts them in an advantageous position when it comes to keeping or entering mid-level jobs when employment opportunities in these decline. A high education level also seems to shelter all other age groups as their proportions in mid-paid jobs decline in a much smaller fashion than their low-educated counterparts, while high educated 50-54 and 55-59 year olds increase their proportion in mid-level jobs. In other words, even if employment shares in mid-level jobs shrink because of the polarisation, these two age groups increase their proportions in such jobs when they are highly educated in the case of 55-59 year olds and when they are either low or high-educated in the case of the 50-54 age group.

Focusing on Table 4.3(d) and on those low-educated individuals in mid-paid jobs whose employment share declines due to polarisation, the trends are the same as when no distinction by education level is made. However, the magnitude of the declines in the proportions for the youngest age groups is almost double in the case of low education than when no distinction according to education level is made, while the decline in the proportion of older age groups is much smaller in the case of low education. All other groups experience smaller declines while the 50-54 age group is the only one that exhibits a small increase of 0.032 in their proportion employed in mid-paid jobs. Overall, the above table shows that the two youngest groups are the ones who are most negatively affected by the decline in mid-paid jobs.

Lastly, mid-paid mid-educated workers seem to have a mixed pattern of change depending on the different age groups when Table 4.3(c) is observed horizontally. Looking at the mid-educated table vertically, it can be seen that in 1999 the 30-34 age group had the highest share of people in mid-paid jobs while the 25-29 age group takes this place in 2004 and 2009. In 2014 it is the 35-39 age group that has the greatest number of people in mid-level jobs.

Looking at the previous three tables diagonally to examine the cohort effect of changes in employment numbers, high-educated cohorts found in mid-paid jobs, experience a decline in their proportions for all age groups in all years of observation. This could mean that as the employment share in mid-level jobs drops, there are fewer positions for highly educated people in any age group, however it is not possible to confirm this simply by looking at these descriptive statistics in this section. In terms of the mid-educated cohorts in mid-level jobs, these see a fall in their proportions over the years of analysis except in the case of the cohort aged 20-24 years old in 1999. Lastly, low-educated cohorts in mid-paid jobs aged 20-24 to 45-49 in 1999, i.e. the first occasion they are observed, appear to experience drops in their numbers in mid-paid jobs as they grow older over the years of observation except of the cohort aged 30-34 in 1999 whose proportion in mid-paid jobs increases.

**Table 4.4(a):
Proportions of People
in High-Level Jobs
(All Education Groups)**

	1999	2004	2009	2014
20-24	0.054	0.053	0.057	0.053
25-29	0.085	0.181	0.186	0.179
30-34	0.109	0.134	0.215	0.206
35-39	0.143	0.147	0.197	0.236
40-44	0.165	0.143	0.169	0.217
45-49	0.139	0.158	0.133	0.154
50-54	0.141	0.144	0.174	0.189
55-59	0.106	0.097	0.163	0.118
60-64	0.020	0.045	0.056	0.071

**Table 4.4(b):
Proportions of High-Educated
People in High-Level Jobs**

	1999	2004	2009	2014
20-24	0.250	0.255	0.187	0.139
25-29	0.260	0.395	0.344	0.307
30-34	0.300	0.319	0.464	0.381
35-39	0.437	0.436	0.430	0.478
40-44	0.527	0.419	0.438	0.473
45-49	0.527	0.486	0.427	0.466
50-54	0.574	0.504	0.495	0.481
55-59	0.667	0.488	0.519	0.367
60-64	0.089	0.197	0.219	0.214

**Table 4.4(c):
Proportions of Mid-Educated
People in High-Level Jobs**

	1999	2004	2009	2014
20-24	0.003	0.012	0.007	0.007
25-29	0.010	0.010	0.028	0.014
30-34	0.031	0.012	0.013	0.027
35-39	0.041	0.025	0.048	0.023
40-44	0.078	0.057	0.062	0.051
45-49	0.103	0.043	0.056	0.041
50-54	0.154	0.106	0.075	0.060
55-59	0.133	0.083	0.113	0.027
60-64	0.043	0.095	0.068	0.060

**Table 4.4(d):
Proportions of Low-Educated
People in High-Level Jobs**

	1999	2004	2009	2014
20-24	0	0	0	0
25-29	0	0	0	0.012
30-34	0	0	0	0
35-39	0	0	0	0
40-44	0	0	0	0
45-49	0	0	0	0.008
50-54	0	0.007	0.009	0.005
55-59	0.007	0	0.004	0.016
60-64	0.003	0	0.003	0

As demonstrated by Table 4.4 (a) above, and in accordance with trends in the job distribution, all age groups except the 20-24 group, increased their shares in high-level jobs. Looking at the table vertically, for example looking at the first column on the left, 40-44 year olds seem to be most represented in high-level jobs in 1999, while in 2004 the 25-29 group has the highest proportion of people in high-paid jobs; the 30-34 group in 2009 and the 35-39 group in 2014.

Looking at Table 4.4 (b), the first thing to point out is that most of the highly educated age groups experience a drop instead of an increase in their proportions in high-level jobs. The only age groups that do experience an increase in their numbers in high-level jobs are the 25-29, 30-34 and 60-64 groups. Similarly, as can be seen in Table 4.4(c), the trends in the proportion of people in high-level jobs who are mid-educated are mixed with the various groups exhibiting very slight increases and decreases over the four periods of observation while there do not seem to be many low-educated people in high level jobs as in most years their proportions are zero as per Table 4.4(d).

Looking at the high-educated table i.e. Table 4.4(b) diagonally, it can be observed that young cohorts up to the age of 35-39 in 1999 increase their shares in high-paid jobs as they age while the two older high-educated cohorts, 40-44 and 45-49 year olds, become less likely to work in high-level jobs as they approach retirement age. This is also the case for their mid-educated counterparts in Table 4.4(c).

The last set of tables presents the proportions of the various age groups out-of-employment over the four years of reference both in total as well as broken down by education level in the same way as in the earlier tables in this section.

Table 4.5(a): Proportions of People Out-Of-Employment (All Education Groups)

	1999	2004	2009	2014
20-24	0.355	0.360	0.381	0.584
25-29	0.255	0.153	0.193	0.281
30-34	0.241	0.167	0.147	0.215
35-39	0.228	0.175	0.167	0.193
40-44	0.262	0.184	0.166	0.234
45-49	0.290	0.211	0.223	0.258
50-54	0.353	0.242	0.249	0.326
55-59	0.478	0.454	0.364	0.444
60-64	0.695	0.631	0.584	0.692

Table 4.5(b): Proportions of High-Educated People Out-Of-Employment

	1999	2004	2009	2014
20-24	0.30	0.245	0.317	0.494
25-29	0.143	0.110	0.169	0.240
30-34	0.122	0.103	0.093	0.175
35-39	0.098	0.076	0.102	0.137
40-44	0.118	0.096	0.086	0.140
45-49	0.084	0.096	0.118	0.173
50-54	0.198	0.080	0.094	0.197
55-59	0.214	0.209	0.215	0.298
60-64	0.822	0.607	0.583	0.607

Table 4.5(c): Proportions of Mid-Educated People Out-Of-Employment

	1999	2004	2009	2014
20-24	0.410	0.459	0.432	0.660
25-29	0.301	0.157	0.197	0.322
30-34	0.198	0.182	0.151	0.230
35-39	0.238	0.193	0.186	0.206
40-44	0.252	0.153	0.178	0.290
45-49	0.345	0.190	0.208	0.251
50-54	0.327	0.235	0.271	0.345
55-59	0.517	0.464	0.358	0.435
60-64	0.688	0.642	0.561	0.692

Table 4.5(d): Proportions of Low-Educated People Out-Of-Employment

	1999	2004	2009	2014
20-24	0.241	0.207	0.283	0.452
25-29	0.315	0.236	0.262	0.366
30-34	0.503	0.266	0.277	0.313
35-39	0.350	0.271	0.282	0.333
40-44	0.372	0.309	0.265	0.320
45-49	0.340	0.314	0.327	0.390
50-54	0.415	0.319	0.357	0.462
55-59	0.502	0.514	0.451	0.559
60-64	0.678	0.633	0.595	0.739

Looking at Table 4.5(a) above vertically, it can be seen that in all years the oldest age group is the one with the largest share of people out-of-employment when compared with other age groups followed by the 55-59 group in 1999 and 2004 but not in 2009 and 2014 where the 20-24 group has the highest percentage of people out-of-employment. In terms of high, mid and low-educated people out-of-employment, and similar to when no education distinction is made, the oldest age group is the one that takes up the largest share of people out-of-employment. The second largest proportion of people out-of-employment in the case of those highly educated is found to be the youngest age group, whereas in the case of those low-educated the second highest proportion of people out-of-employment is found to be the 55-59 age group. In the case of mid-educated, the second largest proportion of people out-of-employment is found to be in the youngest age group in some years and the 55-59 age group in others. In general, older people who are closer to retirement age as well as the youngest workers who are probably still in education or struggling to find their first job, are the ones found to be out-of-employment in greater numbers in all education level groups and when no education distinction is made. The first trend might be explained by the fact that being close to the retirement age, the 60-64 group might find it hard to be re-employed once not in employment or that a number of these workers have retired early. The second trend can be explained either by the fact that this youngest age group may also encompass those still in education or by the crisis and post crisis-induced rise in youth unemployment, meaning that young people find it hard to enter the labour market, especially given their lack of on-the-job experience.

Looking at Table 4.5 (a) horizontally, it can be observed that the 20-24 group exhibited a steady and relatively large increase in their proportion out-of-employment from 1999 to 2014. This is also the case for the 25-29 age group though to a smaller degree and with fluctuations up and down over the years of analysis. All other older age groups have seen a fall in their numbers out-of-employment. Looking at the high-educated table, i.e. Table 4.5 (b), horizontally, all age groups except for the 50-54 and 60-64 groups saw their proportions out-of-employment increase from 1999 to 2014 with small fluctuations between the start and end of this period. In the case of mid-educated people out-of-employment, trends when looking at table 4.5(c) horizontally are mixed but fluctuations do not seem to be too large in magnitude. Lastly, looking at the low-educated table, i.e. Table 4.5 (d), horizontally, it can be seen that the two younger together with the four older age groups, all increase their proportions out-of-employment while the middle three age groups see a drop in the proportions out-of-employment.

Looking at the highly educated cohorts, i.e. looking at Table 4.5(b) above diagonally, it can be seen that the cohorts aged 20-24 and 25-29 in 1999 have slightly reduced their shares out-of-employment between 1999 and 2014 as they aged, while all older cohorts increased their shares out-of-employment. However, this is not the case with low-educated cohorts aged 20-24 and 25-29 in 1999 as even if they experienced small fluctuations in between, they end up with a higher proportion of people out-of-employment in 2014 compared to 1999. This is also the case for all other low-educated cohorts. For the mid-educated, trends when the table is looked at diagonally demonstrate that the cohorts aged 20-24 and 25-29 in 1999 decrease their share out-of-employment, most probably because they manage to enter the labour market and start gaining work experience, while older cohorts observed in 1999 increase their proportions out-of-employment during the period of consideration.

In summary, this section has shown descriptively, how the raw proportions of different stocks of high, mid and low-educated age groups in high, mid and low-level jobs as well as out-of-employment have changed across the four points in time examined in this chapter, namely 1999, 2004, 2009 and 2014. Firstly, in terms of the proportions of workers in low-level jobs, the analysis in this section has demonstrated that when looking at the tables both horizontally and vertically, raw proportions have followed a generally upward trend with some small exceptions and that the share of low-educated people in low-level jobs has risen more substantially between 1999 and 2014 than for the mid and highly educated. Moreover, when looking at the tables diagonally, hence following cohorts over time (i.e. different stocks of workers with the same education and age at each point in time), each cohort of new 20-24 year old entrants begin with a higher proportion of people finding themselves in low-level jobs, which they then struggle to escape as they get older. In addition, again observing the cohort effect by looking diagonally in the low-level job tables, it seems as though each cohort is worse off than the last.

In terms of mid-level jobs, as expected, there is an overall fall in the proportions for all age groups. Young, low-educated workers in mid-level jobs see a sharper fall compared to older workers, which means that young people in 2014 are less likely to work in mid-level jobs than young people in 1999. This could be either because fewer young people start in such jobs over time because for example they choose to stay in education longer so as to compete for the increasing number of high-level jobs, or as a result of the rising higher education level of the population over the years, or that employers, faced with declining numbers of mid-level jobs, prefer to hire (or not lay off) older workers with more work experience than younger candidates with no or little work experience. Moreover, with the exception of the 50-54 and 55-59 age groups, the proportions of high-educated workers in mid-paid jobs also declined over the years. Nevertheless, it was observed that the two youngest age groups in mid-level jobs declined to a much smaller extent when they possessed a high rather than a low-level of education. Similarly, having a high rather than a low-level education seems to be sheltering all age groups who see a decline in their raw proportions in mid-level jobs as those with a low education experienced sharper declines in the raw proportions employed in mid-level jobs than the high-educated age groups. Mid-educated workers seem to have a mixed pattern of change.

For high-level jobs, most age groups increase their shares between 1999 and 2014 when no educational distinction is made. When broken down by education, most of the high-educated groups see a decline in their raw proportions in high-level jobs except for the 25-29, 30-34, 35-39 and the oldest age group, while mid-educated workers do not substantially change their proportions in high-level jobs between 1999 and 2014.

Lastly, trends in the proportions of the different age groups out-of-employment demonstrate that, when no education distinction is made, all groups except for the youngest one and to a lesser extent the 25-29 age group experience a fall in their proportions out-of-employment. In terms of the cohort effect of people out-of-employment, only the two youngest highly educated cohorts in 1999 reduce their shares out-of-employment. In other words, highly educated 20-24 and 25-29 year olds in 1999 are found in greater numbers out-of-employment than equivalently educated 20-29 year olds in 2014 while the opposite is true for all other high-educated groups. This could be explained by the fact that younger cohorts might have had a tough start but have found jobs over the years, while the older cohorts will begin to move out of the labour force, before retirement in some cases. In contrast, low educated 20-24 and 25-29 cohorts are found in greater numbers out of employment over the years. In other words, low-educated 20-29 year olds in 1999 are found in smaller proportions out-of-employment than in 2014, a trend that applies to all other low-educated cohorts as defined by their age group starting in 1999. Therefore, it seems that, as expected, low educated cohorts are in a worse position when it comes to being out of employment than highly educated cohorts.

Given that this section only provides descriptive information in relation to changes in employment patterns across the various parts of the jobs distribution for the different stocks of workers who are the same age and have the same attained education level at different points in time, it is not possible to pinpoint with certainty that the observed trends are the consequences of job polarisation due to routinisation. In addition and in particular, the analysis in this section does not show where workers go afterwards, when they leave mid-level jobs. This is attempted in the following section.

4.4.3 Job Mobility of Displaced Mid-level Workers Regression Results

This section takes a closer look into job mobility of people previously working in mid-level routine jobs. More specifically, the econometric analysis in this section will demonstrate in which of the other job groups, members of each cohort are more/less likely to be found when the proportion of the cohort in the middle of the jobs distribution falls due to routinisation, as postulated by the TBTC theory. In addition, the effect of a drop in the proportion of a cohort employed in mid-level jobs on the proportion of people not in employment (unemployed plus inactive/in early retirement) is also examined so as to observe whether the declining employment shares in mid-level jobs drive some of the people previously working in such jobs out of the labour market.

As mentioned earlier, for the IV regressions, the change in the proportion of a cohort in low-level, low-to-mid-level, mid-to-high-level, high-level jobs and out-of-employment from one period to the next, as opposed to the actual proportion at any one year is used as a dependent variable. The same is true for the independent variable referring to the change in the proportion of mid-paid jobs. Given that employment shares can be slow to change and would not change that much from one year to the next, the fact that each cohort is observed every five years and hence the change in the proportions in mid-level jobs refer to the change during the past five years can be seen as a virtue.

The regression results presented below demonstrate job mobility of the people previously employed in jobs in the middle of the jobs distribution. As discussed in detail earlier, the dependent variable is instrumented using the lag of the proportion of a cohort in jobs that score higher in terms of routine task importance.

Table 4.6: First Stage IV Regression Results for the First-Difference in the Proportion of a Cohort in Mid-Level Jobs

Independent Variables[102]	**Estimated Coefficients**	**Standard Error**
Routinisation at t-1[Instrument]	-0.250***	(0.076)
Age Group (60-64)		
25-29	0.070	(0.064)
30-34	-0.039	(0.048)
35-39	-0.007	(0.047)
40-44	-0.034	(0.043)
45-49	-0.027	(0.039)
50-54	-0.014	(0.034)
55-59	-0.08***	(0.026)
Female	-0.043	(0.096)
Education (Medium level education (i.e. ISCED level 3 and 4))		
Highly Educated	-0.22***	(0.060)
Low Educated	0.290***	(0.081)
Over-educated (Not over-educated)	-0.40**	(0.153)
Part time (Full time)	-0.108	(0.293)
Temporary work (Permanent job/unlimited duration contract)	-0.523***	(0.156)
Year Dummies(2005)		
2009	0.025*	(0.014)
2014	0.041*	(0.027)
Constant	0.048	(0.040)
N	54	

Notes: Dependent Variable is the Change in the Proportion of People in Mid-level Jobs; Significance is denoted by: *** $p < 0.01$, ** $p < 0.05$, ** $p < 0.10$; Omitted reference categories in brackets next to variables names; Clustered Standard errors in brackets next to coefficients.

[102] Given that the IV regressions in this section are run at the cohort, rather than at the individual level, the listed independent variables presented in the pseudo panel IV regression result tables in this section, refer to the mean or proportion of a cohort in possession of each characteristic.

Given that estimators can perform poorly when instruments are weak, i.e. when excluded instruments are only weakly correlated with the endogenous regressors, the weak identification test was run. The test's results in this case have a Kleibergen-Paap Wald rk F-statistic[103] of 10.88 which is above the Stock-Yogo critical values for the 15% maximal value (8.96) i.e. the maximum amount of the possible IV bias that can be tolerated, relative to the OLS bias. In other words, any weak instrument bias will be no more than 15% of the OLS bias which means that the IV strategy is worth continuing and preferable when compared to OLS.

In terms of the first stage results presented in Table 4.6, the effect of the lag of the proportion of a cohort in routine jobs on the change in the proportion of people in mid-level jobs is negative and strongly significant and has a coefficient of -0.25. This result provides support for the TBTC theory meaning that, as hypothesised, the observed polarisation has been caused by technological change in routine jobs. More specifically, each percentage point increase in the proportion of the cohort working in jobs involving routine tasks in the previous period causes a 0.25 percentage point lower change (smaller increase/larger fall) in the proportion of the cohort in mid-paid jobs.

In terms of the effect of age on the change in the proportion of a cohort in mid-level jobs, the insignificant coefficients on all age group variables except the 55-59 group mean that there do not seem to exist significant differences in what is happening to them-except in the case of the 55-59 age group, for whom the proportion working in mid-level jobs is falling faster than for the omitted group – in this case the oldest category, aged 60-64.

Table 4.6 also demonstrates that the increase in mid-level jobs is smaller (or the decrease larger) amongst the highly educated cohorts than the middle-educated cohorts and that the opposite is true for the low-educated cohorts. Similarly, the increase in mid-level jobs is also smaller (or the decrease is larger) in cohorts that have a higher number of a) over-educated people and b) people on temporary work contracts.

Lastly, the year dummies indicate that the change in the proportion in mid-level jobs is larger in 2009 and 2014 (and so maybe less negative given the observed decline in mid-level jobs in section 4.4.1) than the change in 2005 which is the reference period.

[103] According to the help file for ivreg2: when the i.i.d. assumption is dropped and ivreg2 is invoked with the robust, bw or cluster options, the Cragg-Donald-based weak instruments test is no longer valid. ivreg2 instead reports a correspondingly-robust Kleibergen-Paap Wald rk F statistic (Baum, Schaffer and Stillman, 2010)

In the table that follows, the second stage IV regression results are presented when the change in the proportion of a cohort in mid-level jobs is instrumented using the proportion of a cohort employed in routine jobs in the previous period. All the independent variables used in the regressions appear in the left column in the below table and the dependent variable in each IV regression appears in the top row of each column, starting from column 2. The dependent variable in each regression refers to the change in the proportion of a cohort in the various job levels, i.e. the second column refers to the effect of a change in the proportion of mid-level jobs on the change in the proportion in low-level jobs, the third column on the change in the proportion of a cohort in low-mid-level jobs and so on[104].

[104] Results based only on the employee sample

Table 4.7: Second Stage Results of IV Regressions for Job Mobility due to Routinisation

Independent Variables	Change in Low-Level Jobs	Change in Low-Mid-Level Jobs	Change in Mid-High-Level Jobs	Change in High-Level Jobs
Change in the proportion in mid-level jobs [instrumented]	-1.00*** (0.351)	0.360 (0.306)	-0.086 (0.173)	-0.279* (0.180)
Age Group(60-64)				
25-29	-0.138*** (0.040)	0.047 (0.057)	-0.021 (0.038)	0.112*** (0.038)
30-34	-0.106** (0.047)	0.098* (0.052)	-0.036 (0.030)	0.044 (0.046)
35-39	-0.070* (0.051)	0.061 (0.053)	-0.062** (0.026)	0.071* (0.045)
40-44	-0.122*** (0.040)	0.109** (0.046)	-0.028 (0.027)	0.041 (0.043)
45-49	-0.121*** (0.038)	0.118** (0.048)	-0.016 (0.024)	0.019 (0.042)
50-54	-0.105*** (0.030)	0.088** (0.038)	-0.015 (0.015)	0.032 (0.041)
55-59	-0.104** (0.042)	0.112** (0.054)	-0.023 (0.022)	0.015 (0.057)
Female (Male)	0.038 (0.127)	-0.171* (0.117)	-0.017 (0.103)	0.150*** (0.046)
Education(Mid-educated)				
Highly- Educated	-0.041** (0.020)	0.050* (0.029)	-0.047** (0.021)	0.038** (0.019)
Low- Educated	0.007 (0.027)	-0.045 (0.035)	0.039* (0.022)	-0.001 (0.019)
Over-education at t-1	-0.125* (0.081)	0.202 (0.144)	-0.241** (0.093)	0.163* (0.109)
Part time (Full time)	0.346 (0.277)	-0.217 (0.263)	-0.049 (0.110)	-0.081 (0.156)
Temporary work (Permanent job/unlimited duration cont.)	0.037 (0.137)	0.175 (0.132)	0.068 (0.087)	-0.280*** (0.093)
Year Dummies (2005)				
2009	-0.038* (0.021)	-0.002 (0.020)	0.003 (0.013)	0.037*** (0.013)
2014	-0.080** (0.035)	0.061** (0.025)	-0.033* (0.018)	0.052** (0.019)
Constant	0.140*** (0.047)	-0.077 (0.062)	0.062* (0.040)	-0.125** (0.048)
N	54	54	54	54
Centered R^2	0.651	0.371	0.444	0.308

Notes: Significance is denoted by: $***$ $p < 0.01$, $**$ $p < 0.05$, $**$ $p < 0.10$; Omitted reference categories in brackets next to variables names; Clustered Standard errors in brackets next to coefficients[105]

[105] Standard errors are clustered within cohorts to correct for the fact that even if the observations are independent across groups (clusters) this may not be the case within groups (i.e. cohorts in this case) and hence for the fact that the disturbance terms may be correlated for the same cohort in different periods.

First, Table 4.7 demonstrates that as the proportion of a cohort in mid-paid jobs falls, this causes a rise in the proportion of that cohort mostly in low but also in high-paid jobs from one period to the next. This means that, even if people displaced from mid-level jobs move to low as well as high-level jobs, they are more likely to experience a worsening in their labour market position by being forced to move down the jobs distribution in greater numbers than those displaced from mid-level jobs who move up to high-level jobs. It seems that a fall in the proportion of a cohort in mid-paid jobs does not have a causal effect on the propotion of the cohort moving into low-to-mid-paid or mid-to-high-paid jobs as the coefficients for these job groups are not statistically significant. In other words, looking at the variation in the proportions employed in each job group and the job mobility out of mid-paid jobs that is due to routinisation, the IV regressions show that if the fall in the proportion in mid-level jobs is 1 percentage point larger, then the increase in the proportion in low-level jobs will be 1 percentage point higher. Similarly, the increase in the proportion in high-level jobs will be 0.279 percentage points higher while this fall does not significantly affect the change in employment levels in other job groups[106].

In terms of the other controls affecting job mobility, as presented in Table 4.7 above, all age coefficients in the low-paid job column are negative and significant. This means that the changes in the low-paid job employment share are larger (more positive or less negative) for 60-64 year olds which is the reference category than for any other age group. On the other hand, changes in the low-mid-paid job employment share are smaller (more negative or less positive) for all age groups except the 25-29 and 35-39 year olds than they are for 60-64 year olds. For mid-to-high paid jobs, only the 35-39 age group seems to have a more positive employment share change than the reference category while the results above suggest that the increases in employment share in high-paid jobs are larger for the 25-29 and 35-39 age groups than for 60-64 year olds.

Moreover, the higher the proportion of females in a cohort, the less likely it is for the cohort members to move downwards to low-mid paid jobs while the more likely it is for the cohort members to move upwards to high-paid jobs as a higher proportion of the cohort being female corresponds to a larger increase in the proportion of the cohort in

[106] It has to be acknowledged that the sample size here is small and this can lead to imprecise estimates. Nevertheless, the fact that results are significant, despite having small sample sizes and hence large standard errors, suggests that the relationships are strong.

high-level jobs. As also pointed out by Eurofound (2013), the fact that women have increased their employment share in "good" jobs i.e. those in higher quantiles could be in part due to the fact that they are overrepresented in some growing sectors like health, while being underrepresented in declining sectors like construction as well as the rising education level of women. This could explain the finding here that the higher the proportion of females in a cohort, the larger the move to high-paid jobs as women account for a significant part of recent employment growth in the high end of the jobs distribution and men for a larger share of employment decline in mid-level jobs (Eurofound, 2013).

Table 4.7 also shows that, set against the mid-educated reference category, the higher the proportion of a cohort with a high education level, the smaller the change in the proportion in low, and mid-to-high-level jobs while the larger the change in the proportion in low-to-mid, and high-level jobs. On the other hand, the larger the proportion of a cohort who are low educated, the larger the change in the proportion in mid-to-high-level jobs.

Table 4.7 further demonstrates that the higher the proportion of a cohort who is over-educated in the previous period, the larger any increase in the proportion in high-level jobs will be, while the opposite is true for the proportion of the cohort in low and mid-to-high-paid jobs. In other words, those who were over-educated (and so educated to do higher level jobs) are more likely to move to high-level jobs.

The proportion of a cohort in part time jobs does not seem to affect the change in the proportions in any of the job levels/groups studied and the proportion of a cohort working under temporary contract arrangements only seems to affect the change in the proportion in high-paid jobs in a negative way.

Lastly, the year dummies pick up changes over time (for example general changes in the labour market) that affect all age and education groups. Following the results presented in the above table, year dummies seem to be suggesting smaller increases in low-paid jobs over time, and larger increases in high-paid jobs. However, the coefficients on 2014 for low-mid paid and mid-high paid jobs do not seem to fit within this pattern.

The table that follows presents the second stage IV regression results of the effect of a change in the proportion of a cohort in mid-level jobs due to routinisation (instrumented as in the preceding regressions) on the change in the proportion of a cohort out-of-employment (dependent variable).

Table 4.8: Job Mobility from Mid-Level Jobs to Out-of-employment[107]

Independent Variables	**Second Stage IV Coefficients**
Change in the Proportion in Mid-Level Jobs [Instrumented]	-0.135 (0.296)
Age Group(60-64)	
25-29	-0.318*** (0.067)
30-34	-0.243*** (0.035)
35-39	-0.250*** (0.049)
40-44	-0.237*** (0.037)
45-49	-0.224*** (0.029)
50-54	-0.212*** (0.025)
55-59	-0.160*** (0.035)
Female(Male)	-0.343** (0.143)
Education (Mid-Educated)	
Highly Educated	0.020 (0.036)
Low Educated	-0.005 (0.045)
Over-education at t-1	-0.014 (0.191)
Part time	-0.002 (0.158)
Temporary work	0.128 (0.157)
Year Dummies(2005)	
2009	0.091*** (0.021)
2014	0.166*** (0.034)
Constant	0.311*** (0.055)
N	54
Centered R^2	0.890

Notes: Significance is denoted by: ∗∗∗ p < 0.01, ∗∗ p < 0.05, ∗∗ p < 0.10; Omitted reference categories in brackets next to variables names; Clustered Standard errors in brackets next to coefficients

[107] Whereas the regressions in Table 4.7 were run on the "in employment" sample, this regression was run on the whole sample, i.e. employed plus out of work.

As can be seen in Table 4.8 above, the change in the proportion of the cohort out-of-employment does not change significantly with changes in the proportion of a cohort in mid-level jobs. This result suggests that job mobility due to routinisation seems to be causing movement across the jobs distribution, rather than movement out of the labour market.

On the other hand, age seems to play a significant role for movements out of employment. More specifically, since all the age coefficients are negative, this means that each of the age cohorts have a smaller proportion moving into unemployment/early retirement than the 60-64 year old group. Moreover, since the coefficients are more negative for the younger age cohorts, this means that, as expected, these groups are particularly less likely to move out of the labour market.

Similarly, a higher proportion of the cohort being female is associated with a smaller movement of the cohort into unemployment while all other control variable coefficients turn out to be insignificant.

4.5 Summary and Concluding Remarks

In the first part of the present chapter, jobs defined following the jobs approach methodology as occupations within sectors, were ranked both according to their average education level in 1999 as well as by the mean wage they were paying in 2005-06. Plotting the jobs' quantitative evolution over the 15-year period between 1999 and 2014, this chapter has demonstrated that the jobs distribution has upgraded when jobs are ranked according to education but that it has polarised when jobs are ranked according to mean wages, both during the pre-recession and recession periods. This divergence across the two job quality proxies seems to be driven by the fact that an important proportion of jobs in the middle of the wage distribution, whose employment shares decline, has a higher relative position in terms of wages than in terms of education. This is true in sectors such as Agriculture, hunting and forestry and fishing; Construction and Manufacturing and to a lesser extent within Transport, storage and communication and Other community, social and personal service. In these sectors, occupations that are considered as low-level in terms of attained education are found in the middle of the wage distribution.

The observed polarisation when jobs are ranked according to wages means that the employment share in high and low-level jobs increased while employment in jobs in the middle shrank. This job polarisation can be explained by the Task Biased Technological Change (TBTC) theory which postulates that jobs in the middle of the jobs distribution typically involve routine tasks that are the ones that are mostly replaceable by technology following technology advances. On the other hand, human capital in low-level jobs is not easy to replace by technology as the tasks that such jobs encompass require interaction and physical presence by workers (Maselli, 2012), while tasks in high-level jobs are also complementary rather than substitutable by technology. These changes result in a U-shaped jobs distribution.

Following the finding of job polarisation, the raw proportions in each job level (i.e. in low, mid and high-level jobs) by age and year, as well as broken down by education level, were presented so as to observe how the different workers' groups have changed their shares across the various job levels over the years of observation. This section has descriptively shown that each cohort of new 20-24 year old entrants begin with a higher proportion of people finding themselves in low-level jobs, with this proportion remaining high, suggesting that they then struggle to escape such jobs as they get older. Similarly, younger low-educated workers see a sharper fall in their employment shares in mid-level jobs compared to older workers meaning that young people in 2014 are less likely to work in mid-level jobs than young people in 1999.

Lastly, in order to examine job mobility outcomes of workers displaced from mid-level jobs, eighteen different pseudo cohorts based on age and education level were constructed and followed over four distinct periods of time. In order to disentangle the effect of routinisation on the observed mobility of mid-level workers from job mobility for other reasons, an Instrumental Variable (IV) methodology was employed. More specifically, the proportion of a cohort working in jobs involving routine tasks in the previous period, derived from the occupation's routine task intensity score available in Goos et al. (2009), was used as an IV in regressions at the cohort level. According to the routinisation hypothesis, routine tasks are replaced by technology causing an exogenous variation to the proportion of people in mid-level jobs which are the jobs that mostly comprise of such tasks. Therefore, the true effect of the change in the proportion of people employed in mid-level jobs on the change in the proportion employed in each of the other job groups as well as on the change in the proportions out-of-employment that is specifically due to routinisation is revealed. Results from these IV regressions, demonstrated that workers who are displaced from mid-level jobs due to routinisation go on to work mostly in low and to a lesser extent in high-level jobs while it does not seem to be the case that routinisation-induced job polarisation forces displaced mid-paid workers out of the labour market. This finding suggests that, as a result of routinisation, the largest part of workers previously working in mid-level jobs experience a worsening in their labour market position as they move to jobs of a lower quality.

All in all, this chapter has contributed to the job polarisation literature by plotting changes in the employment share across the different job levels for an uninterrupted 15 years period from 1999 to 2014 and for a country where such analyses are, to the best of my knowledge, very scarce. It has also attempted to descriptively examine how the workforce has changed its shares in the different job groups based on age and education characteristics over the years of observation. Most importantly, the present chapter is the first attempt to examine career movements of people previously working in mid-level jobs who have been displaced as a result of routinisation for Cyprus as no other study tackling this question exists for this specific country. Moreover, the use of pseudo cohorts as well as the use of the occupation's routinisation score as an Instrumental Variable, is to my knowledge the first of its kind in the polarisation literature. This methodology has proven to be a reliable way to isolate job movements specifically due to routinisation as the TBTC theory postulates. The implications of this chapter's findings are discussed in Chapter 5 of this thesis.

Appendix 4A: Description of Variables Used in Chapter 4

a) ISCO-88

The 3 digit ISCO-88 occupation classifications are used in this chapter in the derivation of the Jobs Approach variable. Full ISCO classifications used in this chapter can be found at the following link: http://www.ilo.org/public/english/bureau/stat/isco/isco88/major.htm

b) NACE Rev. 1.1

The 1 digit NACE Rev. 1.1 industry classifications are used in this chapter for the derivation of the Jobs Approach Variable

These classifications can be found at the following website: http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_CLS_DLD&StrNom=NACE_1_1

c) **Age Group*:** As mentioned in the main text and Appendix3B, age in the LFS is disseminated in five year intervals. Given that the interest of the present chapter is people of working age, the following age groups are used in the analysis: 20-24; 25-29; 30-34; 35-39; 40-44; 45-49; 50-54; 55-59; 60-64.

d) **Sex*:** This is a binary variable with being male serving as the reference category in the regressions in this chapter.

e) **Highest level of attained education*:** In this Chapter, the derived 3-category education variable of the LFS (HATLEV1D) is used. It is based on the 2-digit, ISCED-97 education variable (HATLEVEL=highest level of education or training successfully completed) described in Appendix3B.

Its categories are as follows:

1) Low: Lower secondary or more (ISCED-97 categories 00,11,21,22)
2) Medium: Upper secondary (ISCED-97 categories 30,31,32,41,42,43)
3) High: Third level (ISCED-97 categories 51,52,60)

f) **Temporary Contract***

This is a binary variable with the reference category being a person has a permanent job or work contract of unlimited duration. Employees with a temporary job/work contract of limited duration according to the EU-LFS User Guide published by Eurostat, Luxembourg are: employees whose main job will terminate either after a period fixed in advance, or after a period not known in advance, but nevertheless defined by objective criteria, such as the completion of an assignment or the period of absence of an employee temporarily replaced.

g) **Part time Vs Full time***

This is a binary variable with the reference category being working full time.

h) **Over-education at t-1***

This refers to the lag of the proportion of a cohort who possess a higher level of attained education than the mode within their job. It was defined and calculated in the same way as in Chapters 2 and 3 but this time within occupation-industry combinations (jobs) rather than just occupations.

*Given that the IV regression analysis in this chapter is performed at the cohort, rather than at the individual level, the variables used as independent variables in the IV regressions refer to the mean, or the proportion, of a cohort in possession of each of the listed characteristics presented in the variable descriptions in this Appendix.

Appendix 4B: The Jobs Approach Job Variable

The job variable in the present chapter was created as a matrix variable with all possible combinations of industry (defined according to NACE Rev. 1.1 1-digit level classifications) and occupation (defined according to ISCO-88 3-digit classification[108]).

Given that some resulting industry-occupation cells were very small, i.e. had less than 20 observations in the job combination variable, these were merged together. In order to deal with the fact that some groups existed in one year but not in another because in some years they had more than 20 observations while in others they had less, these where treated on a one- to- one basis and handled accordingly. For example, if a cell had more than 20 observations in most of the years but in just one year less than 20 then it was kept in all years.

The following NACE industries had less than twenty observations and were hence dropped altogether: Mining and quarrying; Electricity, gas and water supply and Extraterritorial organisations and bodies. Moreover, Fishing is a very small category and given that in ISCO Rev. 2 this category is pooled with Agriculture, hunting and forestry, Fishing was merged with the previous industry.

The final job categories after merging together small cells are presented in the table below. In total, there are 92 occupation-industry combinations. For presentation purposes, the job cells are broken down by industry. For example, the first job in the table below is: Market gardeners and crop growers in Agriculture, the second job is Animal producers and related workers in Agriculture and so on. The numbers in front of each occupation correspond to the ISCO-88 code as provided by the ILO and are reflective of the level of aggregation of each category. For example, single numbers reflect 1-digit level categories, while double and triple numbers reflect 2 and 3-digit levels of aggregation respectively.

To give an example of the small cell correction discussed earlier, the first two jobs in the below table are both a combination of ISCO-88 at the 3 digit level and NACE Rev. 1.1 at the 1 digit level whereas the third job in the table that refers to Other market oriented skilled agricultural and fishery workers in Agriculture pulls together the various small 3-digit ISCO-88 that fall within category 61 into one bigger 2-digit category. In general,

[108] Available at: http://www.ilo.org/public/english/bureau/stat/isco/isco88/major.htm

categories that start with the word "Other" refer to categories that have pooled together more than one occupation at the specific level of aggregation (due to small sample size) like in the aforementioned example of Other 61 market oriented skilled agricultural and fishery workers in Agriculture which contains all the small 3-digit categories within this 2-digit level occupation except groups 611 and 612 that were large enough to stay in a category on their own. Similarly the fourth job, Elementary occupations in Agriculture, pools all 3 and 2 digit ISCO-88 group 9 occupations together at the 1 digit level. It does not have the word "other" in front as there are no categories within the one digit (category 9 Elementary occupations) that are large enough to stay in their original 3 or 2 digit code. Therefore, and as mentioned above, this resulting Jobs Approach variable contains combinations of NACE Rev. 1.1. 1-digit level with ISCO-88 3, 2 and 1-digit occupations.

NACE Rev. 1.1 (1-digit level)	ISCO-88 reclassified groups by industry (Jobs Approach combination variable when combined with industries from left column)
Agriculture, Hunting and Forestry and Fishing	1) 611 Market gardeners and crop growers 2) 612 Market-oriented Animal producers and related workers 3) Other 61 market oriented skilled agricultural and fishery workers 4) 9 Elementary occupations
Manufacturing	5) 1 Legislators and senior officials 6) 2 Professionals 7) 3 Technicians and Associate professionals 8) 4 Clerks 9) 71 Extraction and Building trades workers 10) 72 Metal, machinery and related trades workers 11) 73 Precision, handicraft, craft printing 12) 741 Food processing and related 13) 742 Wood treaters, cabinet-makers and related trades workers

Manufacturing (Cont.)	14) 743 Textile, garment and related trades workers 15) 82 Machine operators and assemblers 16) Other 8 Plant and machine operators and assemblers 17) 932 Manufacturing labourers
Construction	18) 2 Professionals 19) 3 Technicians and Associate professionals 20) 4 Clerks 21) 712 Building frame and related trades workers 22) 713 Building Finishers and related trades workers 23) 714 Painters, building structure cleaners and related trades workers 24) 72 Metal, machinery and related trades workers 25) 8 Plant and machine operators and assemblers 26) 9 Elementary occupations
Wholesale and retail trade: repair of motor vehicles, motorcycles and personal and household goods	27) 1 Legislators and senior officials 28) 2 Professionals 29) 341 Finance and sales associate professionals 30) Other 3 Technicians and Associate professionals 31) 41 Office clerks 32) 42 Customer services clerks 33) 522 Shop, stall and market salespersons 34) 723 Machinery mechanics and fitters 35) 724 Electrical and electronic equipment 36) Other group 7: Craft and related trades workers

Wholesale and retail trade: repair of motor vehicles, motorcycles and personal and household goods (Cont.)	37) Group 8 Plant and machine operators and assemblers 38) 91 Sales and services elementary occupations 39) 93 Labourers in mining, construction , manufacturing and transport
Hotels and restaurants	40) 1 Legislators and senior officials 41) 4 Clerks 42) 512 Housekeeping and restaurant service 43) 913 Domestic and related helpers, cleaners and launderers 44) Other 9 Elementary occupations
Transport, storage and communications	45) 1 Legislators and senior officials 46) 2 Professionals 47) 3 Technicians and Associate professionals 48) 41 Office clerks 49) 42 Customer services clerks 50) 5 Service workers and shop and market Sales Workers 51) 83 Drivers and mobile plant operators 52) 9 Elementary occupations
Financial intermediation	53) 1 Legislators and senior officials 54) 2 Professionals 55) 34 Other associate professionals 56) Other 41 Office clerks 57) 412 Numerical clerks 58) 42 Customer services clerks
Real estate, renting and business activities	59) 1 Legislators and senior officials 60) 21 Physical, mathematical and engineering science professionals 61) 24 Business professionals 62) 3 Technicians and Associate professionals 63) 4 Clerks 64) 9 Elementary occupations

Public Administration and defence; compulsory social security	65) 0 Armed Forces 66) 2 Professionals 67) Other 3 Technicians and Associate professionals 68) 34 Other associate professionals 69) 41 Office clerks 70) 516 Protective Services workers 71) 7 Craft and related trades workers 72) Elementary occupations
Education	73) 232 Secondary education teaching professionals 74) 233 Primary and pre-primary education teaching professionals 75) Other 2 Professionals 76) 3 Technicians and Associate professionals 77) 4 Clerks 78) 5 Service Workers and Shop and Market Sales Workers 79) Elementary occupations
Health and social work	80) 222 Health professionals (except nursing) 81) 223 Nursing and midwifery professionals 82) Other 2 Professionals 83) 3 Technicians and Associate professionals 84) 4 Clerks 85) 5 Service Workers and Shop and Market Sales Workers 86) 9 Elementary occupations
Other community, social and personal services activities	87) 2 Professionals 88) 3 Technicians and Associate professionals 89) 4 Clerks

Other community, social and personal services activities (Cont.)	90) 51 Personal and Protective services workers 91) 9 Elementary occupations
Activities of private households as employers and undifferentiated production activities of private households	92) 913 Domestic and related helpers, cleaners and launderers

Appendix 4C: Variables that Change Throughout the Years of Analysis

a) NACE Rev 1.1. changes to Rev 2 in my files starting in 2009[109]. In order to solve the discrepancy, Rev2 was recoded back to Rev 1.1 as per Appendix 3A.

b) ISCO-08 to ISCO-88

As mentioned in the main text of this chapter, the ISCO classification changed in 2008 from ISCO-88 to ISCO-08. As in the case of Chapter 3, the correspondence table provided by the ILO on its website http://www.ilo.org/public/english/bureau/stat/isco/isco08/ was used to take ISCO-08 back to ISCO-88. Given that the ILO table provides classification correspondences at the 4 digit level while the occupation variable in my data is provided at the 3-digit level, the analogous 3-digit level corresponding codes from ISCO-08 back to ISCO-88 were worked out based on the correspondence table found in the above mentioned link. Of course, it has to be acknowledged that such an exercise can never be perfect, and assumptions or compromises have been made.

[109] It was also changed before from REV 1 to Rev 1.1. (Rev 1.1 was only used from 2005-2007 (or 2008 in my files) but no differences were found at the 1 digit aggregation level.

Appendix 4D: Job Polarisation by Pre-and Post-Recession Periods

Figure 4.A1: Mean Change of the Employment Share when Jobs are Ranked According to the Mean Wage of Job Holders for the Period 1999-2007

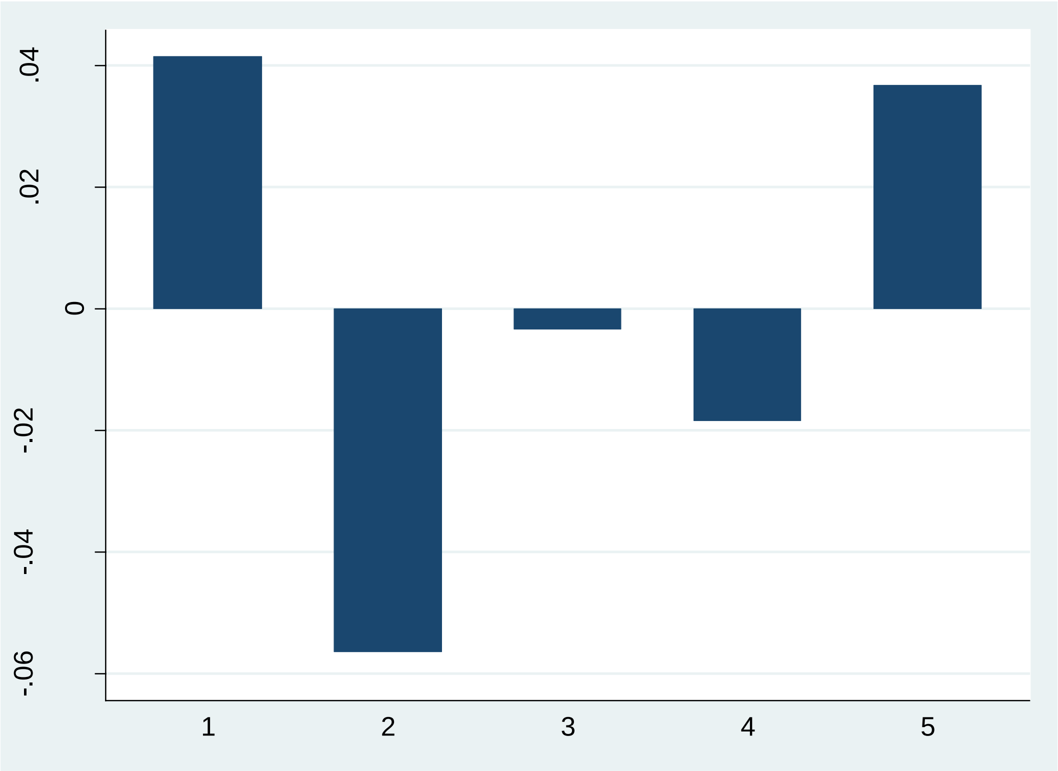


Figure 4.A2: Mean Change of the Employment Share when Jobs are Ranked According to the Mean Wage of Job Holders for the Period 2008-2014

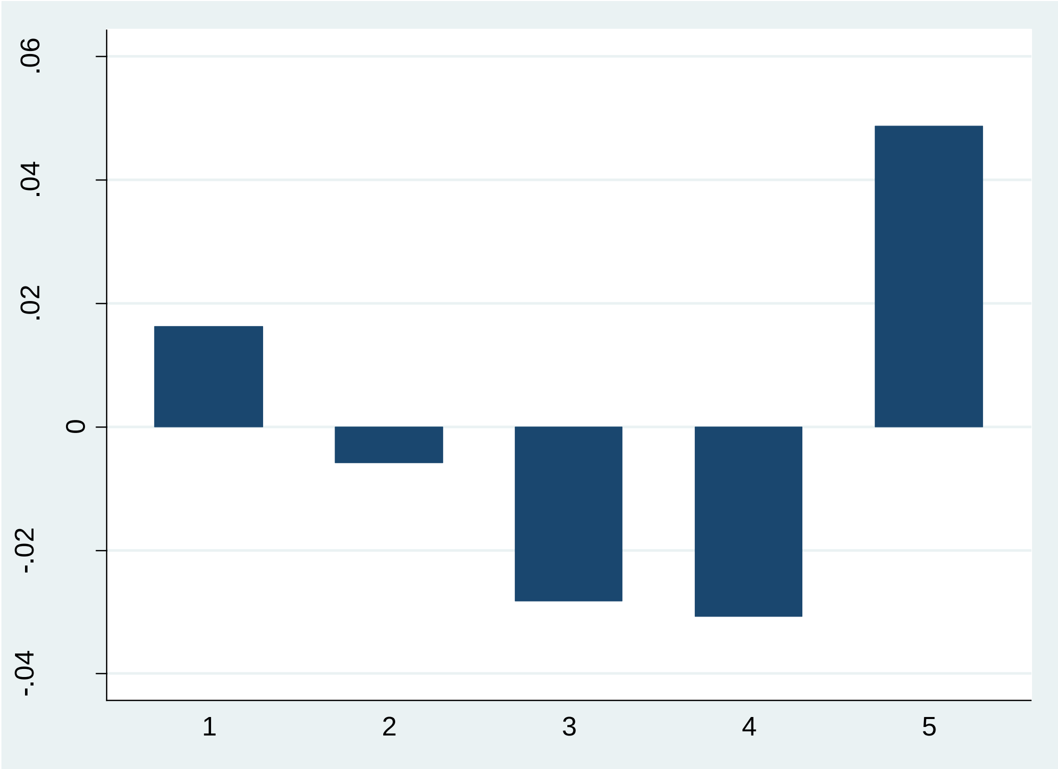


Figure 4.A3: Mean Change of the Employment Share when Jobs are Ranked According to the Mean Wage of Job Holders for the Period <u>1999-2010</u>

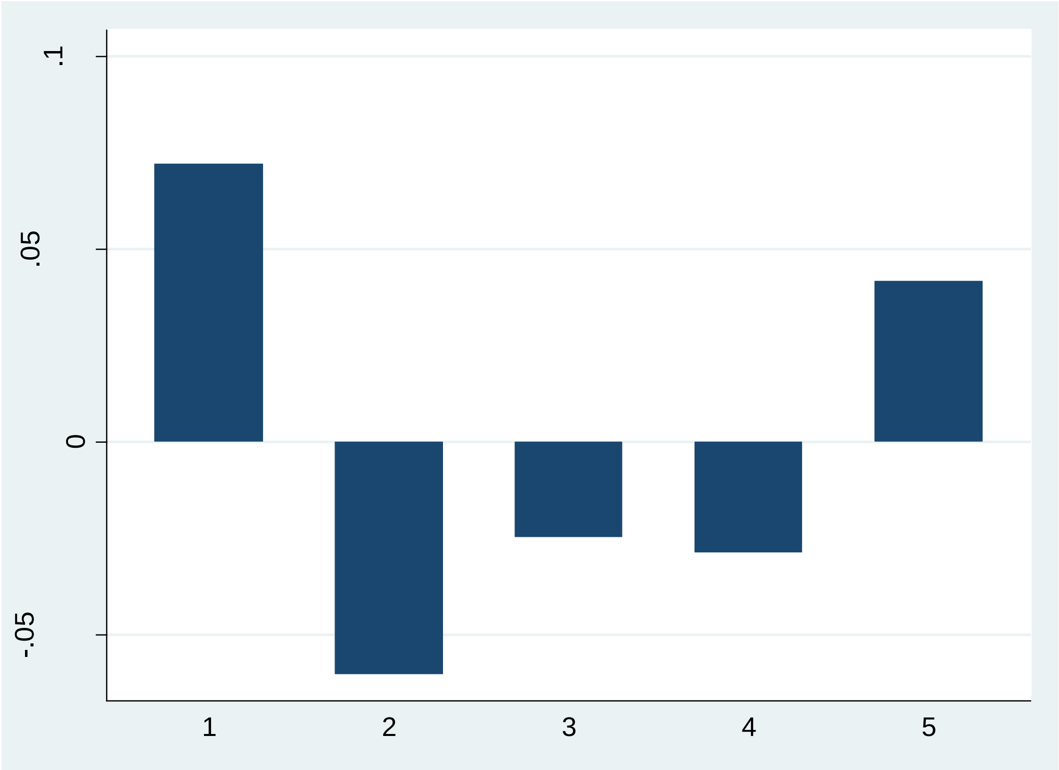
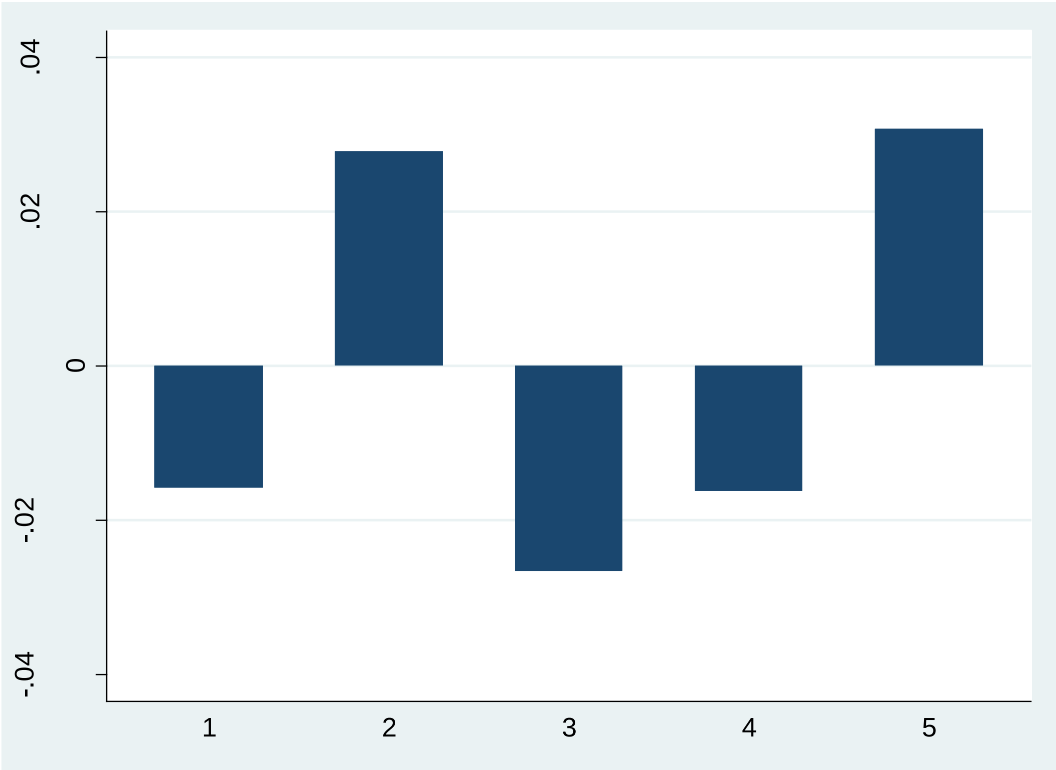
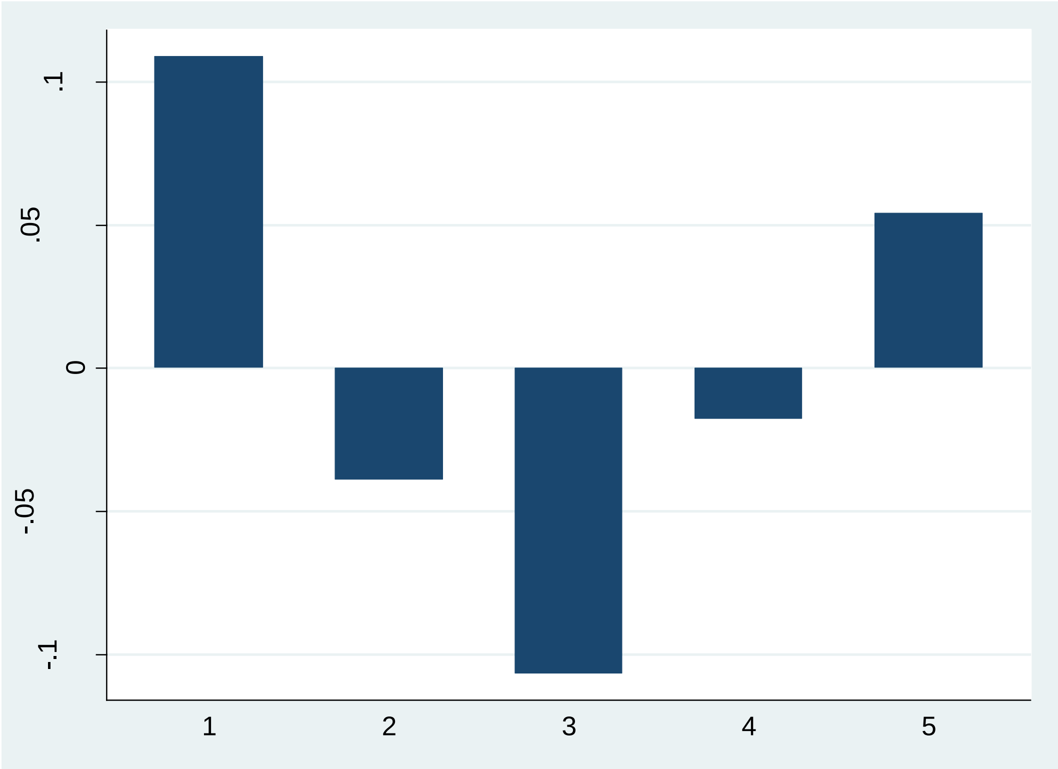


Figure 4.A4: Mean Change of the Employment Share when Jobs are Ranked According to the Mean Wage of Job Holders for the Period <u>2011-2014</u>



Appendix 4E: Job Polarisation by Age Group[110]

Figure 4.A5: Mean change of the employment share when Jobs are ranked according to the mean wage of job holders Aged 15-24 for the period 1999-2014



[110] Graphs for Age groups 55-59 and 60-64 in 1999 are missing here as they have retired before the end of the observation period and hence it is not possible to calculate the change in the employment share between 1999 and 2014

Figure 4.A6: Mean Change of the Employment Share when Jobs are Ranked According to the Mean Wage of Job Holders Aged 25-34 for the Period 1999-2014

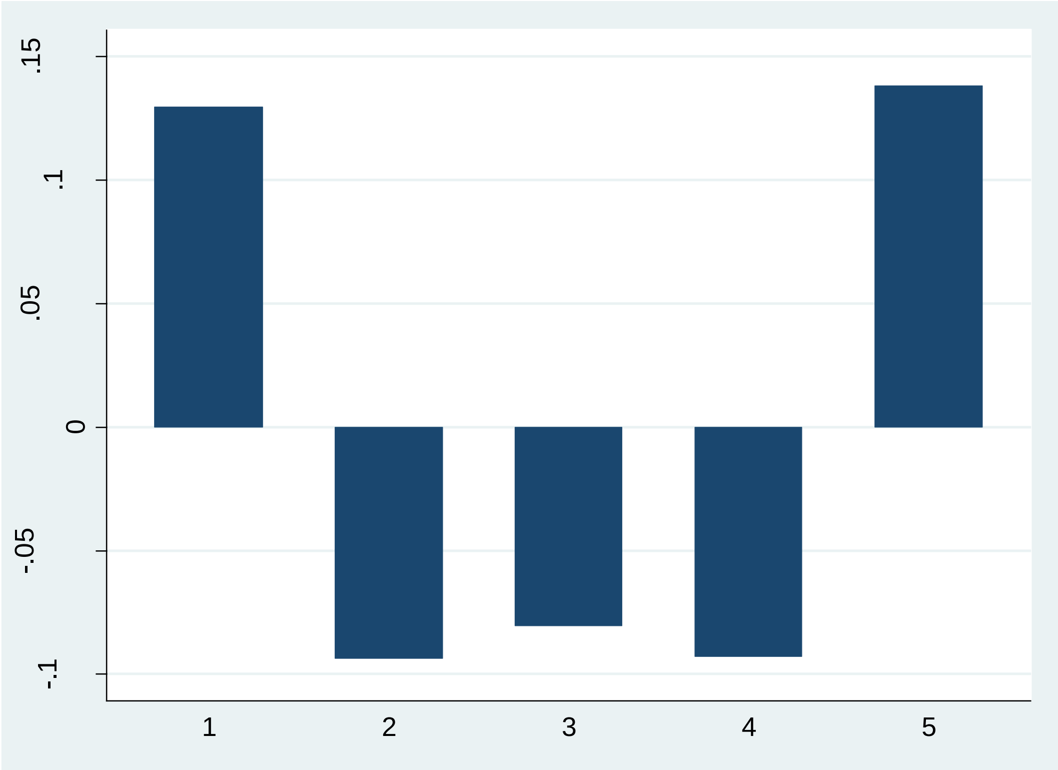


Figure 4.A7: Mean Change of the Employment Share when Jobs are Ranked According to the Mean Wage of Job Holders Aged 35-44 for the Period 1999-2014

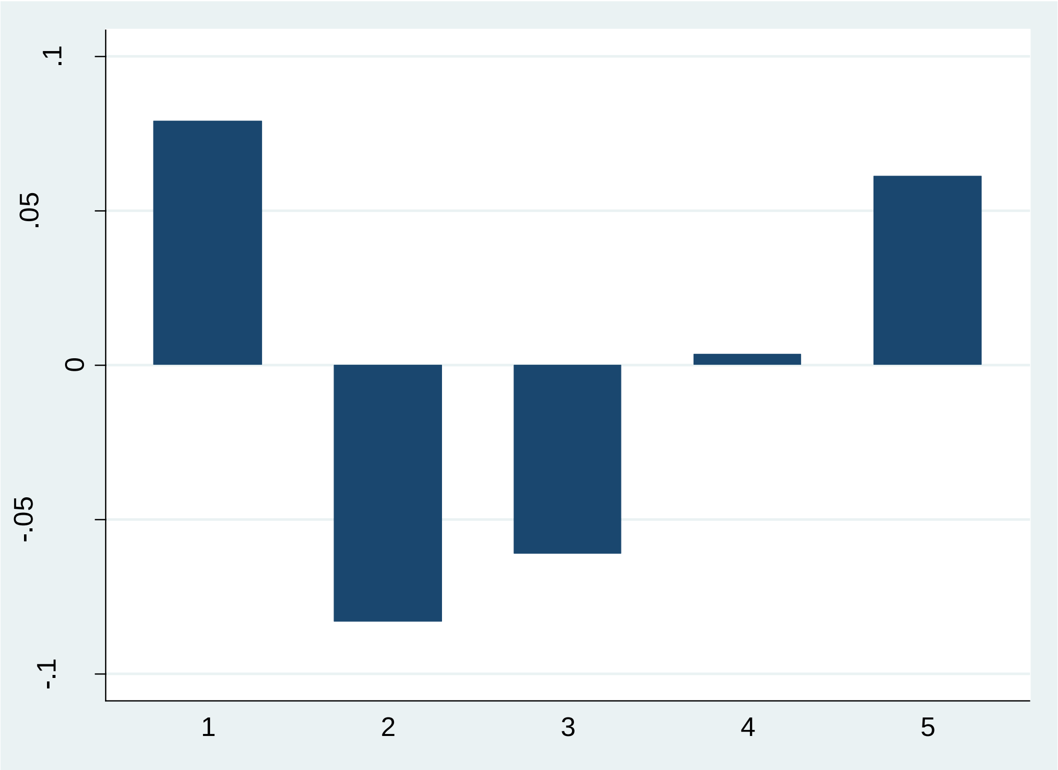
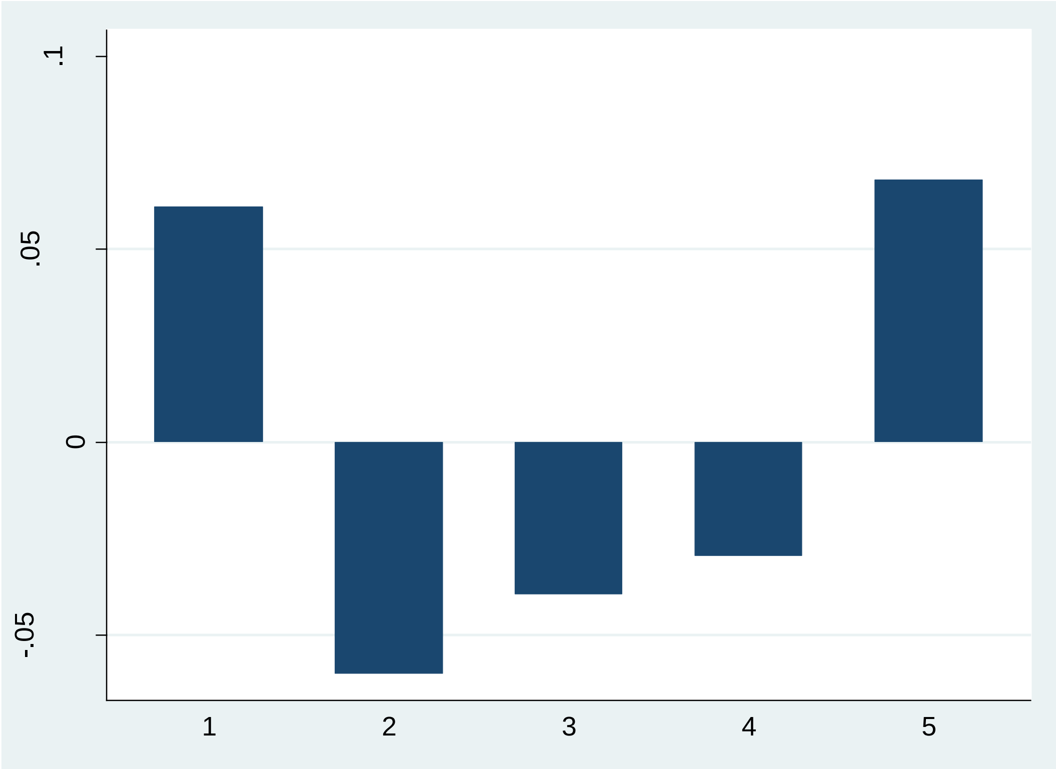


Figure 4.A8: Mean Change of the Employment Share when Jobs are Ranked According to the Mean Wage of Job Holders Aged 45-54 for the Period 1999-2014



Appendix 4F: Routine Task Importance Data used for the Derivation of the IV for Changes in the Proportion of a Cohort in Mid-Level Jobs

Occupations	ISCO Code	Routine task importance
Corporate Managers	12	-1.18
Managers of small enterprises	13	-1.18
Physical, mathematical and engineering professionals	21	-0.86
Life science and health professionals	22	-0.16
Other professionals	24	-1.63
Physical, mathematical and engineering associate professionals	31	0.20
Life science and health associate professionals	32	0.21
Other associate professionals	34	-1.37
Office clerks	41	-1.29
Customer service clerks	42	-0.82
Personal and protective service workers	51	-0.16
Models, salespersons and demonstrators	52	-0.94
Extraction and building trades workers	71	0.98
Metal, machinery and related trade work	72	1.16
Precision, handicraft, craft printing and related trade workers	73	0.81
Other craft and related trade workers	74	0.67
Stationary plant and related operators	81	1.33
Machine operators and assemblers	82	1.31
Drivers and mobile plant operators	83	1.33
Sales and service elementary occupations	91	-0.11
Labourers in mining, construction, manufacturing and transport	93	0.52

Source: Goos et al (2009), Table 4, pp 41

Chapter 5: Conclusions

5.1 Thesis Summary

The present thesis uses individual level microdata from the EU-SILC and EU-LFS as well as macro-level data from other sources and an array of econometric methods to empirically examine three important topics within the realm of education and labour economics for the country of Cyprus. The topics analysed in each of the three empirical investigations of this thesis are the following: over-education micro and macro determinants and over-education state dependence in Chapter 2, on-the-job search as a consequence of over-education and other determinants of on-the-job search in Chapter 3 and the polarisation of the jobs distribution and job mobility of workers displaced from mid-level jobs because of routinisation in Chapter 4.

5.1.1 Over-Education

The first empirical chapter of this thesis, Chapter 2, has provided a thorough investigation of the micro and macro determinants of over-education in Cyprus using panel data from the EU-SILC for the period 2005-2011. To do so, an array of binary probit models was employed, including the Wooldridge (2005) dynamic probit model with Mundlak (1978) corrections. This allowed the isolation of the main determinants of over-education within a methodological setting that copes with the initial conditions problem and unobserved heterogeneity. It has found that factors such as being a woman, married, working under a temporary contract arrangement, having had a job change since last year and recent entry into the labour market from inactivity have a positive impact on the likelihood of over-education. The dynamic econometric modelling used in examining the determinants of over-education and the panel nature of the data set used make this chapter, to my knowledge, the first of its kind for Cyprus.

In addition to contributing to the literature of the determinants of over-education, Chapter 2 has also answered two calls in the over-education literature. First, it has answered calls for further examination of the dynamic properties of over-education. Most importantly, it has identified the extent of over-education state dependence, i.e. the situation whereby the negative impact of those characteristics that were responsible for firstly becoming over-educated is heightened via the continued presence of over-education, thus reinforcing the labour market costs associated with over-education (Mavromaras and McGuinness, 2012). More specifically, individuals may view a job for which they are over-educated as a stepping stone to a matched job, by for example gaining work experience or a temporary income while engaging in on-the-job search in an attempt to achieve a better job match. Chapter 2 has shown that this could result in them being trapped in over-education. More specifically, it has shown that over-education is not only a long-run phenomenon for a significant part of the Cyprus labour force but that it is also highly state dependent or self-perpetuating, with past over-education experience making present over-education more likely. This translates into the possibility of a scarring effect for over-educated workers who fail to make a transition out of over-education, especially given that over-education also appears to be highly self-persistent for workers at all career stages, even after controlling for the employees' observed and unobserved individual characteristics.

Secondly, Chapter 2 has answered calls in the literature for more research into how the overall macroeconomic conditions affect the likelihood of over-education by analysing the effect of a number of labour market conditions on over-education. The inclusion of macro level independent variables to specifically control for both aggregate supply and aggregate demand labour market conditions is a clear contribution to the literature of the determinants of over-education as only a limited number of studies include macro level independent variables in their regressions. More specifically, initial unemployment at the start of paid employment as well as the unemployment rate at the time of the survey, two variables that have been scarcely examined in the past, do not appear to have a significant effect on the probability of over-education. On the other hand, the annualised change in the labour supply by educational category and sex, used to serve as an indication of the level of worker competition in the labour market, and the annualised change in the employment share by occupation and sex, serving as a proxy for labour market demand, two novel variables in the literature of the macro determinants of over-education, both have strongly significant effects with the expected sign. More specifically, as the percentage change in the supply of individuals with equivalent level of education increases, competition for jobs requiring this specific education level increases and as matched jobs are limited, their risk of accepting a job for which they are over-educated increases. Similarly, already over-educated workers may find it harder to escape over-education if the supply of equivalently educated people is high. On the other hand, as the employment share by occupation increases, workers have more opportunities to find a matched job within their preferred occupation or to escape from an over-educated job by finding a matched one either within or outside their firm. These macro variables hence appear to be more important than measures of labour market slack in explaining over-education.

Taken together, the results from Chapter 2 provide evidence that over-education is not random but instead depends on a number of micro as well macro level factors and that previous over-education makes present over-education more likely with individuals who pass through this state running the risk of developing a long-term labour market disadvantage. These findings are important because they can have direct implications both at the individual but also at the policy level. These are discussed in sub-section 5.2 below.

5.1.2 On-The-Job Search

Following on from Chapter 2, Chapter 3 sheds light into the phenomenon of on-the-job search i.e. looking for another job while in employment. It does so using EU-LFS pooled cross sectional data for Cyprus for the fifteen year period between 2000 and 2015. The special interest of this chapter is to disentangle the effect of being over-educated on on-the-job search. An empirical complication in doing so arises if over-education is endogenous i.e. if the factors that affect its probability, also affect the probability of on-the-job search. For example, while unobserved ability increases the chances of less able people being employed in jobs for which they are over-educated, it could also be lowering their likelihood of looking on-the-job. This may be so if such individuals do not consider themselves over-educated at all or if they are less driven to find a matched job. If this is the case, and given that unobserved ability remains uncontrolled for in cross sectional regressions, it could be correlated with the error term hence biasing the over-education regressor in the Probit and OLS specifications.

In this chapter, this possible endogeneity stemming from the fact that the variable of interest may not be random but rather caused by other factors such as individual heterogeneity that are unobserved and so uncontrolled for, is tackled via the use of

Instrumental Variables (IV), which is something new in the on-the-job literature. Via the use of IV regressions any possible (negative) endogeneity bias is eliminated and therefore an unbiased estimate of the base coefficient is revealed. This is done by instrumenting the over-education variable with the annualised change in labour supply by education and sex which was found to significantly affect the probability of being over-educated in Chapter 2. This provides an exogenous shock to over-education while at the same time it has no effect on the dependent variable i.e. on-the-job- search.

Both Probit/OLS regressions as well as the IV regressions confirm the strong and positive effect of over-education on on-the-job search. As expected, for the reasons discussed above, treating over-education as endogenous makes the coefficient of this variable larger than when treating it as exogenous. This means that those over-educated for their jobs are indeed more likely to be looking for another job while employed, a finding in line with the matching theory of over-education which suggests that over-education is sub-optimal from the worker's perspective. Given the scarcity of panel data that include questions related to on-the-job search, the proposed methodology used in this chapter offers a novel point of view into this relationship using pooled cross sectional data that are more readily available while controlling for the possibility of endogeneity bias.

Even though on-the-job search can be an important mechanism for correcting mismatches in the labour market, it is also expensive for the individual and the firm hence potentially creating inefficiencies at the country level. This is not only the case if on-the-job search behaviour is considered as a predecessor of voluntary turnover, in which case costs are related to foregone screening, hiring and training employees who then leave the firm. More specifically, on-the-job search as an employee withdrawal behaviour indicates lower employee commitment and subsequently lower productivity and lower output, even if it does not result in turnover. Hence, the above finding is important as it provides indirect evidence that over-education may not be voluntary or simply a result of lower ability. Moreover, when combined with the picture of over-education as a permanent and self-persistent phenomenon drawn from the results in Chapter 2, this finding interestingly suggests that on-the job search is not successful in freeing over-educated workers before they become trapped in mismatched jobs.

Other determinants of on-the-job search confirmed in this chapter are: age and tenure which are found to have a negative effect on on-the-job search with workers more likely to search on-the-job when they are younger and have short tenures; working part-time or under a temporary contract which are found to increase the likelihood of on-the-job search, and being married which also seems to hinder job search efforts. Married females are also found to elicit less job search compared to single females.

Lastly, Chapter 3 has replicated the Cyprus analysis for the UK and Germany, two countries that differ in terms of their labour market flexibility compared to Cyprus. Results reveal that the relationship between over-education and on-the-job search is also confirmed in other countries and that Cyprus is more similar to the more flexible UK labour market in terms of the studied relation than the more repressed German labour market.

5.1.3 Job Polarisation

Chapter 4 of this thesis has described, analysed and evaluated the phenomenon of job polarisation in Cyprus using cross sectional data from the EU-LFS for the period 1999-2014. Firstly, a jobs-based methodology that defines a job as an occupation within a sector was employed so as to identify how net employment shifts in Cyprus have been distributed across jobs when these are ranked according to two proxies of job quality, namely education and wage criteria. Results from plotting the quantitative evolution of jobs over the studied period have demonstrated that while employment change has followed an upgrading trend when jobs were ranked according to their initial average level of education, this was not the case when jobs were ranked according to the average wage they were paying between 2005 and 2006. In the latter case, the employment share in high and low-level jobs has increased while employment in mid-level jobs has declined, a trend that has been described in the literature as the phenomenon of job polarisation or the hollowing- out of the jobs distribution. Chapter 4 explains that the divergence in the evolution of jobs across the two job quality rankings is due to the fact that an important proportion of jobs in the middle of the wage distribution has a higher relative position in terms of wages than in terms of education. Hence, irrespective of the fact that they are found at different points of the education and wage distributions, the jobs whose employment share declines are the same in both cases, leading to an upgrading pattern in terms of attained education and to polarisation in terms of wages.

The polarisation trend observed in Cyprus over the studied period when jobs are ranked according to wages may be explained by the Task Biased Technological Change (TBTC) theory which postulates that jobs in the high end of the jobs distribution are expected to increase their employment share as they involve tasks which are non-routine and are complementary to technology. This routinisation hypothesis also applies to low-level jobs which involve tasks that are not easily replaceable by technology as they involve interaction and physical presence by workers (Maselli, 2012). On the other hand, many jobs in the middle of the wage distribution incorporate tasks which are highly routine. Following the computer evolution these are the tasks and hence jobs most likely to be replaced by machines resulting in this hollowing-out in the middle of the jobs distribution.

Following the finding of job polarisation, the raw proportions in low, mid and high-level jobs by age and year, as well as broken down by education level, were presented in an attempt to show descriptively where people end up working. More specifically, this section has demonstrated how the workforce has changed its shares across the different job levels over the years of observation and how the shares of younger vs older age groups have changed their position along the different wage quintiles or job levels as well as out-of-employment.

Chapter 4 proceeds to shed light onto job mobility of workers displaced from mid-level jobs due to routinisation. This is an important question that needs to be answered following the earlier finding that the jobs distribution has polarised, as this group of workers is the one mostly at risk of being negatively affected as a consequence. Several calls in the literature for more research into whether workers from the middle category will compete for lower or higher-level jobs have also been made in the past (e.g. Maselli, 2012). Chapter 4 answers such calls in the literature for more research in relation to job mobility of displaced mid-level workers and it does so using repeated cross sections to form pseudo cohorts of workers who are then followed over time. To this end, eighteen different pseudo or artificially created cohorts based on education level and age were constructed and observed over four distinct periods of time so as to study their job mobility over time.

In order to be able to infer causality of the flows, in other words to pinpoint whether the observed flows from one job group to the other is because of polarisation due to routinisation and not due to job mobility for other reasons, such as career changes and promotion, an Instrumental Variable (IV) methodology was employed. This was done by instrumenting the change in the proportion of a cohort in mid-level jobs with the proportion of people working in routine occupations in the previous period, using routine task importance scores from Goos et al. (2009). As per the routinisation hypothesis, the lag of the proportion of people in a cohort employed in jobs that have a high routine task composition, provides an exogenous variation to the change in the proportion of people employed in mid-level jobs. In this way the true effect of the change in the proportion of people employed in mid-level jobs on the change in the proportion employed in each of the other job groups as well as on the change in the proportions out-of-employment that is specifically due to routinisation is revealed. To my knowledge, there do not exist other studies in the literature of job polarisation that use pseudo cohorts to analyse job mobility or that use IV regressions at the cohort level to establish the impact of routinisation on job mobility of workers displaced from mid-paid jobs.

IV regressions ran at the cohort level provided evidence that employees previously working in mid-level jobs move to low and high-level jobs rather than to low-to-mid or mid-to-high-level jobs as a result of routinisation. No evidence was found that they are forced out of the labour market due to routinisation. Nevertheless, the largest proportion of these displaced workers is found to be moving to low rather than to high-level jobs which means that there is a deterioration in the labour market position for the largest part of these employees. This is an important finding with implications at the policy level. Some of these are discussed in the following section.

5.2 Policy Implications and Future Research

Given that skills are a key part of the infrastructure of the economy, the path the economy takes is dependent on decisions relating to investment in education and skills made by policymakers, firms and individuals (Wilson and Zukersteinova, 2011). Studying developments in labour supply (e.g. education demand and education mismatch) and labour demand (e.g. trends in job change over the years) is indispensable not only to guide individuals in making informed career choices but also to provide valuable information to policy makers and other labour market participants about how the future skill demand and supply is likely to develop as the observed trends provide an indication of future changes.

As discussed in Chapters 1 and 2 of this thesis, further investment in education is utilised by governments as a tool in order to increase the country's productivity and competitiveness and hence achieve greater economic growth. However, the finding in Chapter 2 that over-education is self-perpetuating interestingly suggests that over-educated jobs could be acting as a trap for workers. This in turn could translate into a scarring effect for over-educated individuals who could develop a long-term labour market disadvantage. This can have a "dampening effect on the growth potential of the economy" (Mavromaras et al. 2012, 10) especially given the increasing number of people who pursue higher education studies and the large investments in educational spending incurred by governments.

This finding therefore signals that society may not be making the right educational investments and calls for a reconsideration of the current education policies and a diversion of policies towards preventing entry into over-education and discouraging people from accepting mismatched jobs as a career strategy rather than finding measures to correct it at a later stage or letting it correct for itself. This could be done, for example, if policies that facilitate entry of young educated people into jobs commensurate with their education by for example subsidising part of their salaries are enhanced. Such policies give motives to firms to employ young graduates directly into matched jobs hence preventing them from accepting jobs for which they are over-educated and being trapped as a consequence. Similarly, there is a need for programs specifically designed to offer employment and hence work experience and industry-specific knowledge alongside early career counselling and correct matching by government job centers to help individuals stay out of over-education.

Chapter 2 has also provided evidence of macro level labour supply and labour demand factors affecting the likelihood of over-education. This is an important finding because it can have direct implications not only at the individual but also at the policy level. More specifically, it suggests that timely labour market information is indispensable so as to inform individuals' career decisions as to which direction to follow in their studies and job search. In other words, there is a need for governments to create mechanisms that will provide students with information on changing labour market macro conditions to help them make informed decisions as to their further education and labour market choices. For example, information on the numbers of individuals by education level and on the employment share within occupations can give students an overall idea of what to expect when they enter the labour market and how to maximise their chances of finding a matched job. Education counsellors within the school system who are up to date with changing macro conditions are also important for helping students make the correct decisions bearing in mind the changing labour market situation so as to avoid being trapped in a job for which they are over-educated.

Similarly, information in terms of the number of people graduating from the different levels of education can assist policy makers towards creating opportunities of employment for those groups whose labour supply is in abundance so as to absorb them in occupations matching their education level, hence preventing over-education before it happens. In this way education investments are not wasted. Lastly, increasing the share of vocational education and training, as opposed to the more general education commonly provided in Cyprus, especially for entry into growing sectors or sectors with lower supply of educated labour, is another possible policy route.

The above findings also call for more research not only into the connection between macro conditions and vertical mismatch but also into horizontal mismatch, a situation in which workers are mismatched because they work in a different field to the one they studied. Such analyses will enable an understanding of how field of study choice affects the probability of over-education and whether more informed choices could lessen mismatch. A replication of the analysis undertaken in Chapter 2 using a longer panel duration so as to observe the longer-run effects of over-education would also be beneficial.

Chapter 3, the second empirical investigation of the present thesis, has confirmed the positive relation between over-education and on-the-job search. On-the-job search, as a signal of employee withdrawal behaviour, is linked to lower commitment and this is negatively correlated with employee and hence firm output and productivity. An interesting implication of this finding at the firm level is that firms will be opt to avoid hiring over-educated applicants (Wald, 2005).

Moreover, combining the findings of Chapters 2 and 3, it is to be expected that on-the-job search may not be able to free over-educated workers from sub-optimal matches and into matched employment, even if the data set used does not allow a direct examination of on-the-job search outcomes. In other words, given that being over-educated in one year is found to significantly cause being over-educated the next year and the fact that over-education is found to be a permanent state in Chapter 2, it is unlikely that on-the-job search by over-educated workers is to be expected to have a positive outcome in terms of restoring a good education-job match. All in all, the analysis in Chapter 3 offers evidence that, via its effect on on-the job search, over-education could have a real negative productivity penalty not only for the worker but also for the firm and the economy as a whole and therefore the creation of policies to prevent and/or reduce the level of over-education should be placed high on the political agenda.

As also pointed out by DeLoach and Kurt (2018), on-the-job search literature relying on cross sectional data sets, could be problematic given that only a small proportion of workers report such activities at one single point in time, i.e. the time of the survey, meaning that the extent of on-the-job search in the economy is expected to be much greater as workers that for example are not looking now have searched on-the-job at some point in their careers. In order to overcome the above issue, panel data is required. To my knowledge, there is no available panel data set with a direct measure of on-the-job search available for Cyprus at the moment and this is also the case in many other countries which explains the limited empirical evidence on on-the-job search in the literature. Even if a handful of studies utilise panel data to study on-the job search or related issues such as intention to quit and voluntary job mobility(eg. Mavromaras et al. 2013, McGuiness and Wooden 2007; Allen and van der Velden 2001; Congregado et al. 2016), the on-the-job search literature could benefit from further research, ideally using panel data. Such longitudinal analysis can also permit an observation of on-the-job search outcomes and could add useful evidence as to the degree to which on-the-job search is able to free mismatched workers from over-educated jobs.

Lastly, in terms of job polarisation, while this thesis provides an optimistic account in terms of the higher growth in high-level jobs, compared to the growth in low-level jobs, the fact that the middle of the job distribution has hollowed-out is a cause of concern for policy makers. What is more, the cohort level regression analysis in Chapter 4 has demonstrated that when the proportion of mid-level jobs drops due to routinisation, the largest proportion of displaced mid-level workers moves down the job distribution and into low-level jobs. This means that the largest part of workers previously working in mid-level jobs experience a worsening of their labour market position as the proportion of the different cohorts in low-level jobs increases to a larger extent than the increase in the proportion in high-level jobs. The implication of this finding is that policies should focus on improving career prospects of displaced workers by facilitating their moves to other mid-paid jobs that do not decline or offering development and (re-) training opportunities (Smith, 2013) to help them use their experience to progress from mid to high rather than to low-level jobs. As noted by McIntosh (2013), if policies enabling moves of displaced workers towards intermediate (and/or high) rather than low- level jobs are not put in place, then the labour market position of many workers will be weakened as mid-level jobs are usually perceived as 'good jobs' for non-graduates whereas lower-level jobs are associated with lower pay, less interesting work and less progression prospects and could hence be viewed as 'bad jobs' (McIntosh, 2013).

Another potential implication of the above finding, which is nevertheless not examined in the present thesis, is the claim that the slow growth in mid-level jobs, could impede career prospects of people working in low-level jobs to move up the employment structure into higher level jobs (Wright and Dwyer, 2003). Even if the outcomes of workers in low-level jobs are not observed in the present thesis, the polarising trends found when jobs are ranked according to wages as well as the regression results that provide evidence of job mobility from declining mid-level jobs mostly into low-level jobs, the above mentioned concern also seems warranted in the case of Cyprus. In other words, given that mid-level jobs are declining in numbers because of routinisation and displaced mid-level workers are found to increasingly move down the jobs distribution into low-level jobs, then the prospects for low-level job holders of progression into mid-level jobs, i.e. upward mobility, could be severely disrupted. What is more, the risk exists that such workers could also be bumped out of the labour market altogether. As pointed out by McIntosh (2013), no empirical evidence seems to exist in relation to the changing likelihood of progression from entry-level jobs due specifically to the polarisation of the labour market and this calls for more research towards this direction.

Moreover, Holmes (2010) calls for further investigation as to whether job polarisation has differentially affected new labour market entrants and younger generations compared to those already in the labour market, i.e. whether job polarisation has caused new entrants to enter the labour market at the lower and higher ends of the jobs distribution, while those already in the labour market have remained in mid-level jobs. This seems to be, in part supported by the descriptive analysis in this chapter, however further work, ideally using panel data to follow multiple cohorts over time and to specifically distinguish new entrants and observe actual transitions, is required to look into the outcomes of new entrants compared to those already employed in mid-level jobs in greater detail.

As mentioned earlier, one of the limitations of the polarisation literature in general, as well as of the analysis in Chapter 4 of this thesis, is that the polarisation results are conditional on the observed job quality at one single fixed point in time. Further research, possibly relaxing this common assumption of a fixed distribution of jobs defined by wages at one specific point in time, to examine for example whether new mid-level jobs are created over the years and what these jobs are would be useful. Similarly, and as pointed out by McIntosh (2013), jobs can change their relative position along the jobs distribution as a result of changes in relative wages over time for example, driven by changes in the structure of labour demand, labour supply or due to labour market institutions (McIntosh, 2013). Such research could also attempt to answer questions as to whether workers now working in these new jobs are those displaced from previously mid-level jobs who have managed to find new employment in the middle of the jobs distribution, or whether they consist of individuals who regressed (progressed) from high-level (low-level) jobs that have deteriorated (improved) in terms of the wage they pay and have hence moved down (up) into the middle of the jobs continuum (Holmes, 2011).

Lastly, as pointed out by Tüzemen and Willis (2013), the observed polarisation with employment growth at the higher-end of the job distribution and a shrinking of the middle, could act as an incentive to workers to attain higher levels of education, yet this is conditional on a number of personal characteristics. Worker responses to job polarisation is another area that can benefit from additional research, especially given the fact that it can have potential consequences for issues relating to labour supply and possible imbalances such as over-education examined in the previous chapters of this thesis with important policy implications.

To summarise, each of the chapters in this thesis individually contributes to the literature of the micro and macro-level determinants of over-education, its dynamic properties and state dependence, on-the-job search as a result of over-education and job polarisation and the resulting job mobility of displaced mid-level workers. The increasing number of highly educated workers over the years paired with the polarisation of the jobs distribution are two of the most important labour market phenomena of the previous decades that shape both the supply and demand sides of the labour market. The results of the present thesis highlight both of these trends in the case of Cyprus.

Over-education and job polarisation, the two main studied phenomena in this thesis could potentially be linked if for example, as the job distribution polarises, with jobs growing in the extremes while hollowing-out in the middle, this acts as a signal for individuals to gain more education so as to compete for the more abundant high-level jobs. This in turn increases the supply of labour for those high-level jobs and could lead some of those highly educated workers to be employed for jobs in which they are over-educated as the empirical findings in Chapter 2 demonstrate. This is similar for those previously employed in mid-level jobs who may find themselves over-educated if they move down to low-level jobs due to polarisation. Similarly, if workers displaced from mid-level jobs move up to high-level jobs, they could be causing those with high-education to become over-educated if the demand cannot meet the supply for such jobs. On the other hand, those over-educated may manage to escape over-education if jobs at the high end of the job distribution increase and are so able to absorb over-educated workers who for example were previously working in mid-level jobs and are displaced as a result of polarisation. Similarly, another channel via which changes in educated supply could be adding up to polarisation is if younger cohorts decide to stay longer in education and to provide their labour to high-level rather than mid-level jobs hence reducing the inflow into mid-level jobs. This combined with the increase in the outflow of mid-level workers due to routinisation, as documented in Chapter 4, adds to the picture of polarisation. At the firm level, employers may also respond to the increased educational attainment of the population and the shrinking of mid-level jobs by raising the minimum educational hiring standards of certain jobs-known as credentialism (Goos and Manning, 2007). Therefore, it is evident that even though over-education and job polarisation are two distinct phenomena, they can be highly interrelated via a number of different pathways.

Bibliography

Acemoglu D. (1999): Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence, *American Economic Review* 89 (5), 1259-1278

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of Labor Economics*, 4, 1043-1171.

Alba-Ramírez, A. (1993). Mismatch in the Spanish Labor Market: Overeducation? *Journal of Human Resources*, 28(2), 259-278.

Allen, J., & Van der Velden, R. (2001). Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search. *Oxford Economic Papers*, 53(3), 434-452.

Antràs, P., Garicano, L., & Rossi-Hansberg, E. (2006). Offshoring in a knowledge economy. *The Quarterly Journal of Economics*, 121(1), 31-77.

Arellano, M., & Carrasco, R. (2003). Binary choice panel data models with predetermined variables. *Journal of Econometrics*, 115(1), 125-157.

Arulampalam, W., & Stewart, M. B. (2009). Simplified Implementation of the Heckman Estimator of the Dynamic Probit Model and a Comparison with Alternative Estimators. *Oxford Bulletin of Economics and Statistics*, 71(5), 659-681.

Autor, D. H., & Dorn, D. (2013). How technology wrecks the middle class. *The New York Times*, 24.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4), 1279-1333.

Autor, D., & Dorn, D. (2009). *Inequality and specialization: the growth of low-skill service jobs in the United States.* NBER Working Paper Series, (15150).

Autor, D., Katz, L. F., & Kearney, M. (2006). Measuring and interpreting trends in economic inequality. In *AEA Papers and Proceedings,* 96(2), 189-194.

Bacchus, M. K. (2008). The education challenges facing small nation states in the increasingly competitive global economy of the twenty-first century. *Comparative Education*, 44(2), 127-145.

Bachmann, R. (2005). *Labour market dynamics in Germany: Hirings, separations, and job-to-job transitions over the business cycle* (No. 2005, 045). SFB 649 Discussion Paper.

Baert, S., & Verhaest, D. (2014). *Unemployment or Overeducation: Which is a Worse Signal to Employers?* (No. 8312). Institute for the Study of Labor (IZA).

Baert, S., Cockx, B., & Verhaest, D. (2013). Overeducation at the start of the career: stepping stone or trap? *Labour Economics,* 25, 123-140.

Battu, H., & Sloane, P. (2000). Overeducation and crowding out in Britain, pp.157-174 in Borghans, L., & De Grip, A. (Eds) *The Overeducated Worker?* Cheltenham: Edward Elgar Publishing.

Battu, H., & Sloane, P. J. (2004). Over-education and ethnic minorities in Britain. *The Manchester School*, 72(4), 535-559.

Battu, H., Belfield, C. R., & Sloane, P. J. (2000). How well can we measure graduate over-education and its effects? *National Institute Economic Review*, 171(1), 82-93.

Bauer, T. K. (2002). Educational mismatch and wages: a panel analysis. *Economics of Education Review*, 21(3), 221-229.

Baum, C. F., Schaffer, M. E., & Stillman, S. (2010). *ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression. Statistical Software Components S425401*, Boston College Department of Economics, Boston College, Boston.

Becker, G. (1964). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education.* New York: Columbia University Press.

Bell, D. N., & Blanchflower, D. G. (2011). Young people and the Great Recession. *Oxford Review of Economic Policy*, 27(2), 241-267.

Berrington, A., Smith, P. W. F., & Sturgis, P. (2006). *An overview of methods for the analysis of panel data.* Southampton: ESRC National Centre for Research Methods

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates?. *The Quarterly Journal of Economics*, 119(1), 249-275.

Bjelland, M., Fallick, B., Haltiwanger, J., & McEntarfer, E. (2011). Employer-to-employer flows in the United States: estimates using linked employer-employee data. *Journal of Business & Economic Statistics*, 29(4), 493-505.

Blau, F. D., & Kahn, L. M. (1981). Race and sex differences in quits by young workers. *Industrial and Labor Relations Review*, 34(4), 563-577.

Blinder, A. S. (2007). Free trade's great, but offshoring rattles me. *Washington Post*, *6*, B04.

Blinder, A. S. (2009). How many US jobs might be offshorable? *World Economics*, 10(2), 41.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115-143.

Boll, C., Leppin, J. S., & Schömann, K. (2016). Who is overeducated and why? Probit and dynamic mixed multinomial logit analyses of vertical mismatch in East and West Germany. *Education Economics*, 24(6), 639-662.

Bond, S. R. (2002). Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal*, 1(2), 141-162.

Borghans, L., & De Grip, A. (2000). *The overeducated worker?* Cheltenham: Edward Elgar Publishing.

Brynin, M. (2002). Overqualification in employment. *Work, Employment & Society*, 16(4), 637-654.

Büchel, F. (2002). The effects of overeducation on productivity in Germany—the firms' viewpoint. *Economics of Education Review*, 21(3), 263-275.

Büchel, F., & Battu, H. (2003). The theory of differential overqualification: Does it work? *Scottish Journal of Political Economy*, 50(1), 1-16.

Burdett, K. (1978). A theory of employee job search and quit rates. *The American Economic Review*, 68, 212-220.

Burdett, K., & Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, 39(2), 257-273.

Campbell, I. (1997). Beyond unemployment: the challenge of increased precarious employment. *Just Policy: A journal of Australian Social Policy*, 11, 4.

Carrillo-Tudela, C., & Kaas, L. (2015). Worker mobility in a search model with adverse selection. *Journal of Economic Theory*, 160, 340-386.

Carrillo-Tudela, C., Hobijn, B., Perkowski, P., & Visschers, L. (2015). *Majority of hires never report looking for a job.* FRBSF Economic Letter, 10.

Cedefop (2010). *Skills supply and demand in Europe. Medium-term forecast up to 2020.* Luxembourg: Publication Office of the European Union

Cedefop (2012) *Future skills supply and demand in Europe. Forecasts 2012* Luxemburg: Publication Office of the European Union

Chevalier, A. (2003). Measuring over-education. *Economica*, *70*(279), 509-531.

Chevalier, A., & Lindley, J. (2009). Overeducation and the skills of UK graduates. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(2), 307-337.

Christofides, Kourtellos and Vrachimis (2007) Unemployment Indices for Cyprus: A Comparative Study. *Economic Policy Papers*, No. 01-07

Collard, F., Fonseca, R., & Munoz, R. (2002). *Spanish unemployment persistence and the ladder effect*. Centre for Economic Performance, London School of Economics and Political Science.

Congregado, E., Iglesias, J., Millán, J. M., & Román, C. (2016). Incidence, effects, dynamics and routes out of overqualification in Europe: A comprehensive analysis distinguishing by employment status. *Applied Economics*, 48(5), 411-445.

Cuesta, M. B., & Budría, S. (2011). *Overeducation Dynamics and Personality* (No. 2011/05). Universidad Autónoma de Madrid (Spain), Department of Economic Analysis (Economic Theory and Economic History).

Cuesta, M. B., & Ocaña, M. Á. M. (2005). Educational mismatch and labour mobility of people with disabilities: the Spanish case. *Revista de Economía Laboral*, 2(1), 31-55.

Cyprusprofile.com [online] Available at: https://www.cyprusprofile.com/en/economy [Accessed on 30 October 2018].

Cyprus Statistical Service, *Cyprus Statistic Service Official Website* [online] Available at: http://www.cystat.gov.cy [Accessed on 31 Oct. 2014]

David, H. (2015). Why are there still so many jobs? The history and future of workplace automation. *The Journal of Economic Perspectives*, 29(3), 3-30.

De Grip, A., Bosma, H., Willems, D., & Van Boxtel, M. (2008). Job-worker mismatch and cognitive decline. *Oxford Economic Papers*, 60(2), 237-253.

De Oliveira, M. M., Santos, M. C., & Kiker, B. F. (2000). The role of human capital and technological change in overeducation. *Economics of Education Review*, 19(2), 199-206.

Deaton, A. (1985). Panel data from time series of cross-sections. *Journal of Econometrics*, 30(1-2), 109-126.

DeLoach, S. B., & Kurt, M. (2018). On-the-Job Search, Mismatch and Worker Heterogeneity. *Journal of Labor Research*, 39(2), 219-233

Di Pietro, G., & Urwin, P. (2006). Education and skills mismatch in the Italian graduate labour market. *Applied Economics*, 38(1), 79-93.

Dieter, V., & Omey, E. (2009). *The Determinants of Overeducation: Different Measures, Different Outcomes?* Ghent University (No. 2009/592).

Dolado, J. J., Jansen, M., & Jimeno, J. F. (2003). *On-the-Job Search in a Matching Model with Heterogenous Jobs and Workers*. Discussion Papers No.886. Institute for the Study of Labor (IZA).

Dolton, P., & Silles, M. (2001). *Over education in the graduate labour market: some evidence from alumni data*. Centre for the Economics of Education, London School of Economics and Political Science.

Dolton, P., & Silles, M. (2003). The Determinants and Consequences of Graduate Overeducation, pp.189-216 in De Grip, A., & Mertens, A. (Eds.) *Overeducation in Europe: Current issues in theory and policy*. Cheltenham: Edward Elgar Publishing.

Dolton, P., & Vignoles, A. (2000). The incidence and effects of overeducation in the UK graduate labour market. *Economics of Education Review*, 19(2), 179-198.

Drukker, D. M. (2003). Testing for serial correlation in linear panel-data models. *Stata Journal*, 3(2), 168-177.

Duncan, G. J., & Hoffman, S. D. (1981). The incidence and wage effects of overeducation. *Economics of Education Review*, 1(1), 75-86.

Dustmann, C., Ludsteck, J., & Schönberg, U. (2009). Revisiting the German wage structure. *The Quarterly Journal of Economics*, 124(2), 843-881.

Eurofound (2013). *Employment polarisation and job quality in the crisis: European Jobs Monitor 2013*, Eurofound, Dublin.

Eurofound (2015). *Upgrading or polarisation? Long-term and global shifts in the employment structure: European Jobs Monitor 2015*, Publications Office of the European Union, Luxembourg.

Eurostat (2008).NACE Rev 2 Statistical classification of economic activities in the European Community. Retrieved 7th March 2015 from http://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF/dd5443f5-b886-40e4-920d-9df03590ff91?version=1.0

Eurostat (2017). Methodological guidelines and description of EU-SILC target variables *Statistical Office of the European Union's Official Website* [Online] Available at: http://ec.europa.eu/eurostat/documents/1012329/8655367/PERSONAL+DATA_Basic_data.pdf/2b3673a5-5e80-438c-890f-72671cbcf2ae [Accessed 02 May 2018]

Eurostat (2018). EU Labour Force Survey Database User Guide. *Statistical Office of the European Union's Official Website* [online] Available at: http://ec.europa.eu/eurostat/documents/1978984/6037342/EULFS-Database-UserGuide.pdf [Accessed 3 May 2018].

Fallick, B., & Fleischman, C. (2001). *The importance of employer-to-employer flows in the US labor market* (No. 2001-18). Board of Governors of the Federal Reserve System (US).

Fernandez-Macias, E. (2012). Job polarization in Europe? Changes in the employment structure and job quality, 1995-2007. *Work and Occupations*, 39(2), 157-182.

Firpo, S., Fortin, N. M., & Lemieux, T. (2011). *Occupational tasks and changes in the wage structure* (No. 5542). Discussion paper series//Forschungsinstitut zur Zukunft der Arbeit.

Flisi, S., Goglio, V., Meroni, E. C., Rodrigues, M., & Vera-Toscano, E. (2017). Measuring occupational mismatch: overeducation and overskilling in Europe-evidence from PIAAC. *Social Indicators Research*, 131(3), 1211-1249.

Frank, R. H. (1978). Why women earn less: the theory and estimation of differential overqualification. *The American Economic Review*, 68(3), 360-373.

Freeman, R. (1976). The overeducated American. Cleveland: Academic Press

Frenette, M. (2004). The overqualified Canadian graduate: the role of the academic program in the incidence, persistence, and economic returns to overqualification. *Economics of Education Review*, 23(1), 29-45.

Gadrey, J. (1996). *Services, la productivité en question*. Desclée de Brouwer, Paris

Gibbons, R. D., & Hedeker, D. (1996). MIXOR: a computer program for mixed-effects ordinal regression analysis. *Computer Methods and Programs in Biomedicine*, 49(2), 157-176.

Gibbons, R. D., & Hedeker, D. (1994). Application of random-effects probit regression models. *Journal of Consulting and Clinical Psychology*, 62(2), 285.

Glass, A. J., & Saggi, K. (2001). Innovation and wage effects of international outsourcing. *European Economic Review*, 45(1), 67-86.

Goos, M, Manning, A. and Salomons, A. (2009). Job polarization in Europe. *American Economic Review: Papers and Proceedings*, 99(2), 58-63.

Goos, M, Manning, A. and Salomons, A. (2011). *Explaining Job Polarization: the Roles of Technology, Offshoring and Institutions.* (No. ces11. 34). KU Leuven, Faculty of Economics and Business, Department of Economics.

Goos, M. and Manning, A. (2007). Lousy and lovely jobs. The rising polarization of work in Britain. *The Review of Economics and Statistics,* 89(1), 118-133.

Goos, M., Manning, A., & Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *The American Economic Review*, 104(8), 2509-2526.

Green, C., Kler, P., & Leeves, G. (2007). Immigrant overeducation: Evidence from recent arrivals to Australia. *Economics of Education Review*, 26(4), 420-432.

Green, F., & McIntosh, S. (2007). Is there a genuine under-utilization of skills amongst the over-qualified? *Applied Economics*, 39(4), 427-439.

Green, F., & Zhu, Y. (2008):*Overqualification, job dissatisfaction, and increasing dispersion in the returns to graduate education.* No. 08,03 Department of Economics Discussion Paper, University of Kent.

Greene, W. H. (2008). *Econometric Analysis*. 6th Edition. New Jersey: Prentice Hall

Groot, W. (1993). Overeducation and the Returns to Enterprise-related Schooling. *Economics of Education Review*, 12(4), 299-309.

Groot, W. (1996). The incidence of, and returns to overeducation in the UK. *Applied Economics*, 28(10), 1345-1350.

Groot, W., & Maasen Van Den Brink, H. (1997). Allocation and the returns to over-education in the UK. *Education Economics*, 5(2), 169-183.

Groot, W., & Maassen van den Brink, H. (2003). The dynamics of skill mismatches in the Dutch labour market, pp. 49-64 in De Grip, A., & Mertens, A. (Eds.) *Overeducation in Europe: Current issues in theory and policy*. Cheltenham: Edward Elgar Publishing.

Groot, W., & Van Den Brink, H. M. (1996). Glass ceilings or dead ends: Job promotion of men and women compared. *Economics Letters*, 53(2), 221-226.

Groot, W., & Van den Brink, H. M. (2000). Overeducation in the labor market: a meta-analysis. *Economics of Education Review*, 19(2), 149-158.

Grossman, G. M., & Rossi-Hansberg, E. (2008). Trading tasks: A simple theory of offshoring. *American Economic Review*, 98(5), 1978-97.

Hartog, J. (1985). Earnings functions: testing for the demand side. *Economics Letters*, 19(3), 281-285.

Hartog, J. (2000). Over-education and earnings: where are we, where should we go?. *Economics of Education Review*, 19(2), 131-147.

Hartog, J., & Oosterbeek, H. (1988). Education, allocation and earnings in the Netherlands: 0verschooling? *Economics of Education Review*, 7(2), 185-194.

Heckman, J. J. (1981). Heterogeneity and state dependence, pp. 91-140 in Sherwin Rosen, (ed.) Studies *in Labor Markets*. Chicago: University of Chicago Press.

Heckman, J. J. (1991). Identifying the hand of past: Distinguishing state dependence from heterogeneity. *The American Economic Review*, 81(2), 75-79.

Heritage.org (2018). *The Heritage Foundation's official website* [online] Available at: https://www.heritage.org/index/labor-freedom [Accessed on 05 Apr. 2018].

Hersch, J. (1991). Education match and job match. *The Review of Economics and Statistics*, 73(1), 140-144.

Hersch, J. (1995). Optimal 'mismatch'and promotions. *Economic inquiry*, 33(4), 611-624.

Highereducation.ac.cy, (2018). *Cyprus Ministry of Education and Culture Official Website* [Online] Available at: http://www.highereducation.ac.cy/en/che-introduction.html [Accessed 10 May 2018].

Hijzen, A. (2007). International outsourcing, technological change, and wage inequality. *Review of International Economics*, 15(1), 188-205.

Holmes, C. (2010). *Job Polarisation in the UK: An Assessment Using Longitudinal Data.* SKOPE Research Paper No. 90.

Holmes, C. (2011). *The Route out of the Routine: Where do the Displaced Routine Workers Go?* SKOPE Research Paper No. 100.

Holmes, C. and Mayhew, K. (2012). *The Changing Shape of the UK Job Market and its Implications for the Bottom Half of Earners.* London: Resolution Foundation Report.

Holmes, C., Mayhew, K. and Chow, F. (2011). *Mobility and the Changing Structure of Occupations: a Cross-Cohort Comparison*. SKOPE mimeo.

Ilo.org, (2014) International Labour Organisation's Official Website. [online] Available at: http://www.ilo.org/public/english/bureau/stat/isco/isco08/ [Accessed 20 Apr. 2014]

Jensen, J. B., & Kletzer, L. G. (2010). Measuring tradable services and the task content of offshorable services jobs, pp. 309-335 in Katharine G. Abraham, James R. Spletzer, and Michael Harper (Eds) *Labor in the new economy*. Chicago: University of Chicago Press.

Jones, M. K., & Sloane, P. J. (2010). Disability and Skill Mismatch. *Economic Record*, 86(s1), 101-114.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy*, 87(5), 972-990.

Kahn, L. B. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labour Economics*, 17(2), 303-316.

Kampelmann, Stephan; Rycx, François (2011): *Task-biased changes of employment and remuneration: The case of occupations*, Discussion paper series

Keith, K., & McWilliams, A. (1999). The returns to mobility and job search by gender. *ILR Review*, 52(3), 460-477.

Kiersztyn, A. (2013). Stuck in a mismatch? The persistence of overeducation during twenty years of the post-communist transition in Poland. *Economics of Education Review*, 32, 78-91.

Kiker, B. F., Santos, M. C., & De Oliveira, M. M. (1997). Overeducation and undereducation: Evidence for Portugal. *Economics of Education Review*, 16(2), 111-125.

King, G., & Roberts, M. E. (2015). How robust standard errors expose methodological problems they do not fix, and what to do about it. *Political Analysis*, 23(2), 159-179.

Korpi, T., & Tåhlin, M. (2009). Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000. *Labour Economics*, 16(2), 183-193.

Koutsampelas and Polycarpou (2013): The Distributional Consequences of the Cypriot Crisis; *Cyprus Economic Policy Review*, 7(1), 47-61

Kucel, A., & Byrne, D. (2008). *Are over-educated people insiders or outsiders? A case of job search methods and over-education in UK* (No. 258). ESRI working paper.

Kucel, A., Molina, I. F., & Raya, J. M. (2016). Over-education and its opportunity cost in Japan. *Asia Pacific Education Review*, *17*(2), 299-312.

Kyrizi, A. (2014). *Essays on overeducation* (Doctoral dissertation, University of Leicester).

Lancaster, T. (1979). Econometric methods for the duration of unemployment. *Econometrica: Journal of the Econometric Society*, 47(4), 939-956.

Leuven, E., & Oosterbeek, H. (2011). Overeducation and mismatch in the labor market. *Handbook of the Economics of Education*, *4*, 283-326.

Levels, M., Van der Velden, R., & Allen, J. (2014). Educational mismatches and skills: new empirical tests of old hypotheses. *Oxford Economic Papers*, 66(4), 959-982.

Lindley, J., & McIntosh, S. (2008). *A Panel Data Analysis of the Incidence and Impact of Over-education* (No. 2008009). The University of Sheffield, Department of Economics.

Lindley, J., & McIntosh, S. (2010*). Is the Over-Education Wage Penalty Permanent?* (No. 2010004). The University of Sheffield, Department of Economics.

Linsley, I. (2005). Causes of overeducation in the Australian labour market. *Australian Journal of Labour Economics*, 8(2), 121.

Liu, K., Salvanes, K. G., & Sørensen, E. Ø. (2016). Good skills in bad times: Cyclical skill mismatch and the long-term effects of graduating in a recession. *European Economic Review*, 84, 3-17.

Liu, K., Salvanes, K. G., & Sørensen, E. Ø. (2016). Good skills in bad times: Cyclical skill mismatch and the long-term effects of graduating in a recession. *European Economic Review*, 84, 3-17.

Locke, E. A., & Dunnette, M. D. (1976). *The nature and causes of job satisfaction,* 1297-1349 in Handbook of industrial and organizational psychology. Skokkie: Rand McNally and Co.

Manning, A. (2004). We Can Work It Out: The Impact of Technological Change on the Demand for Low-Skill Workers. *Scottish Journal of Political Economy*, 51(5), 581-608.

March, J. G., & Simon, H. A. (1958). Organizations. Wiley, New York.

Maselli, I. (2012). The evolving supply and demand of skills in the labour market. *Intereconomics*, 47(1), 22-30.

Mavromaras, K., McGuinness, S., Sloane, P., & Wei, Z. (2010). *Job Mismatches and Labour Market Outcomes: Panel Evidence on Australian University Graduates* (No. 5083). Institute for the Study of Labor (IZA).

Mavromaras, K., & McGuinness, S. (2012). Overskilling dynamics and education pathways. *Economics of Education Review*, 31(5), 619-628.

Mavromaras, K., Mahuteau, S., Sloane, P., & Wei, Z. (2012). *The Persistence of Overskilling and Its Effects on Wages.* Research Report. National Centre for Vocational Education Research Ltd. Adelaide: Australia.

Mavromaras, K., Mahuteau, S., Sloane, P., & Wei, Z. (2013). The effect of overskilling dynamics on wages. *Education Economics*, 21(3), 281-303.

Mavromaras, K., McGuinness, S., & Fok, Y. K. (2009). Assessing the incidence and wage effects of overskilling in the Australian labour market. *Economic Record*, 85(268), 60-72.

Mavromaras, K., McGuinness, S., & Wooden, M. (2007). Overskilling in the Australian labour market. *Australian Economic Review*, 40(3), 307-312.

Mavromaras, K., McGuinness, S., O'Leary, N., Sloane, P., & Wei, Z. (2013). Job mismatches and labour market outcomes: panel evidence on university graduates. *Economic Record*, 89(286), 382-395.

McCormick, B. (1990), 'A Theory of Signalling During Job Search, Employment Efficiency, and "Stigmatised" Jobs', *Review of Economic Studies*, 57, 299-313

McGoldrick, K., & Robst, J. (1996). Gender differences in overeducation: A test of the theory of differential overqualification. *The American Economic Review*, 86(2), 280-284.

McGuinness, S. (2006). Overeducation in the labour market. *Journal of Economic Surveys*, 20(3), 387-418.

McGuinness, S., & Wooden, M. (2009). Overskilling, job insecurity, and career mobility. *Industrial Relations: A Journal of Economy and Society*, 48(2), 265-286.

McIntosh, S. (2005). *Using pseudo cohorts to track changes in the qualifications of national populations*. Department for Education and Skills.

McIntosh, S. (2013) *'Hollowing Out and The Future of the Labour Market*,' BIS Research Paper 134.

Mendes de Oliveira, M., Santos, M. C., & Kiker, B. F. (2000). The role of human capital and technological change in overeducation. *Economics of Education Review*, 19(2), 199-206.

Menon, M. E. (1998). Factors influencing the demand for higher education: The case of Cyprus. *Higher Education*, 35(3), 251-266.

Menon, M. E. (2008a). Perceived rates of return to higher education: Further evidence from Cyprus. *Economics of Education Review*, 27(1), 39-47.

Menon, M. E. (2008b). The economic benefits of higher education in Cyprus: The expectations of prospective students. *International Journal of Educational Development*, 28(3), 259-267.

Menon, M. E., Pashourtidou, N., Polycarpou, A., & Pashardes, P. (2012). Students' expectations about earnings and employment and the experience of recent university graduates: Evidence from Cyprus. *International Journal of Educational Development*, 32(6), 805-813.

Mincer, J. (1978). Family migration decisions. *Journal of Political Economy*, 86(5), 749-773.

Morano, C. P. (2014). The determinants of overeducation: Evidence from the Italian labour market. *Investigaciones de Economía de la Educación*, 9(9), 681-698.

Mortensen, D. (1975). *The turnover implications of learning about attributes on the job.* Northwestern University, processed.

Mortensen, D. T., & Pissarides, C. A. (1999). New developments in models of search in the labor market. *Handbook of Labor Economics*, 3, 2567-2627.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica: Journal of the Econometric Society*, 46(1), 69-85.

Nellas, V., & Olivieri, E. (2012). *The Change of Job Opportunities: the Role of Computerization and Institutions* (No. 804). Quaderni-Working Paper DSE.

Nordhaus, W. D. (2007). Two centuries of productivity growth in computing. *The Journal of Economic History*, *67*(1), 128-159.

Oesch, D. (2013). *Occupational change in Europe: how technology and education transform the job structure.* Oxford University Press.

Oesch, D. and J. Rodriguez Menés (2011). Upgrading or polarization? Occupational change in Britain, Germany, Spain and Switzerland, 1990–2008. *Socio-Economic Review 9*, 503–531.

Oldenski, L. (2012). The task composition of offshoring by US multinationals. *Economie Internationale*, (3), 5-21.

Oreopoulos, P., Von Wachter, T., & Heisz, A. (2012). The short-and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics*, 4(1), 1-29.

Orme, C. D. (1997) *The initial conditions problem and two-step estimation in discrete panel data models*, Mimeo, University of Manchester, October.

Parsons, D. O. (1972). Specific human capital: An application to quit rates and layoff rates. *Journal of Political Economy*, 80(6), 1120-1143.

Parsons, D. O. (1991). The job search behavior of employed youth. *The Review of Economics and Statistics*, 73(4) 597-604.

Piper, A. (2013). *A Note on Modelling Dynamics in Happiness Estimations* (No. 49364). University Library of Munich, Germany.

Piper, A. (2015). Heaven knows I'm miserable now: overeducation and reduced life satisfaction. *Education Economics*, 23(6), 677-692.

Pissarides, C. A. (1994). Search unemployment with on-the-job search. *The Review of Economic Studies*, 61(3), 457-475.

Pissarides, C. A. (1999). Policy influences on unemployment: the European experience. *Scottish Journal of Political Economy*, 46(4), 389-418.

Pissarides, C. A., & Wadsworth, J. (1994). On-the-job search: some empirical evidence from Britain. *European Economic Review*, 38(2), 385-401.

Ponzo(2012), On-the-job Search in Italian Labor Markets: An Empirical Analysis, *Int. J. of the Economics of Business*, 19(2), 213–232

Rosen, S. (1972). Learning and experience in the labor market. *Journal of Human Resources*, 7(3), 326-34

Royalty, A. B. (1998). Job-to-job and job-to-non employment turnover by gender and education level. *Journal of Labor Economics*, 16(2), 392-433.

Rubb, S. (2003). Overeducation in the labor market: a comment and re-analysis of a meta-analysis. *Economics of Education Review*, 22(6), 621-629.

Sala, G. (2011). Approaches to skills mismatch in the labour market: a literature review. *Papers: Revista de Sociologia*, (96), 1025-1045.

Sattinger, M. (1993). Assignment models of the distribution of earnings. *Journal of Economic Literature*, 31(2), 831-880.

Scherer, S. (2004). Stepping-stones or traps? The consequences of labour market entry positions on future careers in West Germany, Great Britain and Italy. *Work, Employment and Society*, 18(2), 369-394.

Sengupta, S., Edwards, P. K., & Tsai, C. J. (2009). The good, the bad, and the ordinary: Work identities in “good” and “bad” jobs in the United Kingdom. *Work and Occupations*, 36(1), 26-55.

Sicherman, N. (1991). "Overeducation" in the Labor Market. *Journal of Labor Economics*, 9(2), 101-122.

Sicherman, N., & Galor, O. (1990). A theory of career mobility. *Journal of Political Economy*, 98(1), 169-192.

Sloane, P. J. (2009). *Disability and Skill Mismatch* (No. 4430). Institute for the Study of Labor (IZA).

Sloane, P. J., Battu, H., & Seaman, P. T. (1999). Overeducation, undereducation and the British labour market. *Applied Economics*, 31(11), 1437-1453.

Smith, C. L. (2013). *The dynamics of labor market polarization* (No. 2013-57). Board of Governors of the Federal Reserve System (US).

Smith, M., Burchell, B., Fagan, C., & O'Brien, C. (2008). Job quality in Europe. *Industrial Relations Journal*, 39(6), 586-603.

Soderbom, M. (2009). *Applied Econometrics, Lecture 10: Binary Choice Models.* University of Gothenburg. [Accessed 05 May 2018] Available from http://www.soderbom.net/lecture10notes.pdf

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, 87(3), 355-374.

Spitz-Oener, A. (2006). Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. *Journal of Labor Economics*, 24(2), 235-270.

Thurow, L. C. (1975). *Generating inequality: Mechanisms of Distribution in the U.S. Economy*. New York: Basic Books, Inc.

Tsang, M. C. (1987). The impact of underutilization of education on productivity: A case study of the US Bell companies. *Economics of Education Review*, 6(3), 239-254.

Tsang, M. C., & Levin, H. M. (1985). The economics of overeducation. *Economics of Education Review*, 4(2), 93-104.

Tsang, M. C., Rumberger, R. W., & Levin, H. M. (1991). The impact of surplus schooling on worker productivity. *Industrial Relations: A Journal of Economy and Society*, 30(2), 209-228.

Tuma N B. (1985): 'Effects of Labor Market Structure on Job Shift Patterns', in Heckman J. Singer B (eds), *Longidutinal Analysis of the Labor Market Data*, Cambridge: Cambridge University Press.

Tüzemen, D., & Willis, J. (2013). *The vanishing middle: Job polarization and workers' response to the decline in middle-skill jobs*. Economic Review-Federal Reserve Bank of Kansas City, 5.

Unesco Institute for Statistics (2012) *International Standard of Education Classification ISCED 2011*. Retrieved 7th March 2015 from http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf

Valletta, Robert G. (2016): *Recent Flattening in the Higher Education Wage Premium: Polarization, Skill Downgrading, or Both?* IZA Discussion Papers, No. 101

Van Ophem, H. (1991). Wages, nonwage job characteristics and the search behavior of employees. *The Review of Economics and Statistics*, 73(1), 145-151.

Verbeek, M. (2008). *A guide to modern econometrics*. Chichester: John Wiley & Sons.

Verhaest, D., & Omey, E. (2006). The impact of overeducation and its measurement. *Social Indicators Research*, 77(3), 419-448.

Verhaest, D., & Omey, E. (2010). The determinants of overeducation: different measures, different outcomes?. *International Journal of Manpower*, 31(6), 608-625.

Verhaest, D., & van der Velden, R. K. (2010). *Cross-country differences in graduate overeducation and its persistence.* ROA Research memorandum. Research Centre for Education and the Labour Market| ROA. ROA-RM-2010/7 ROA.

Violante G.L. (2008) Skill-Biased Technical Change. In: Palgrave Macmillan (eds) The New Palgrave Dictionary of Economics. Palgrave Macmillan, London

Vishwanath, T. (1989): Job Search, Stigma Effect, and Escape Rate from Unemployment. *Journal of Labor Economics*, 7, 487–502

Wald, S. (2005). The impact of overqualification on job search. *International Journal of Manpower*, 26(2), 140-156.

Walker, I., and Zhu, Y. (2008) The College Wage Premium and the Expansion of Higher Education in the UK, *The Scandinavian Journal of Economics* 110 (4), 695-709.

Wilson, R., & Zukersteinova, A. (2011). *Anticipating changing skill needs: A Master Class*. Institute for Employment Research. [Accessed 05 Apr. 2016] Available from http://www.opetushallitus.fi/download/145693_868.pdf

Wolbers, M. H. (2003). Job Mismatches and their Labour-Market Effects among School-Leavers in Europe. *European Sociological Review*, 19(3), 249-266.

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, 20(1), 39-54.

Wright, E. O., & Dwyer, R. E. (2003). The patterns of job expansions in the USA: a comparison of the 1960s and 1990s. *Socio-Economic Review*, 1(3), 289-325.

Zakariya, Z. (2017). Job Mismatch and On-the-job Search Behavior Among University Graduates in Malaysia. *Asian Economic Journal*, *31*(4), 355-379.