UNIVERSITEIT GENT

FACULTEIT ECONOMIE EN BEDRIJFSKUNDE

ACADEMIEJAAR 2011 – 2012

The Influence of Health on Overeducation. An Econometric Evaluation.

Masterproef voorgedragen tot het bekomen van de graad van

Master of Science in de Algemene Economie

Robin Deman

onder leiding van

Prof. Dr. D. Verhaest

UNIVERSITEIT GENT

FACULTEIT ECONOMIE EN BEDRIJFSKUNDE

UNIVERSITEIT GENT

FACULTEIT ECONOMIE EN BEDRIJFSKUNDE

ACADEMIEJAAR 2011 – 2012

The Influence of Health on Overeducation.

An Econometric Evaluation.

Masterproef voorgedragen tot het bekomen van de graad van

Master of Science in de Algemene Economie

Robin Deman

onder leiding van

Prof. Dr. D. Verhaest

PERMISSION

Ondergetekende verklaart dat de inhoud van deze masterproef mag geraadpleegd en/of gereproduceerd worden, mits bronvermelding.

Robin Deman

FOREWORD

I would like to thank my promoter, prof. dr. D. Verhaest, for his assistance, advice, feedback and guide concerning the subject and the econometric model. The helpful comments from Stijn Baert are gratefully acknowledged.

Thank you Laura for your support. I know you always believed in me.

Table of Contents

List of Tables

1. Introduction

Investment in human capital is a key element of economic progress. Over the past few decades, investment in education has steadily increased in most industrialized countries. Parallel to this evolution have been growing concerns about overeducation – a situation in which an individual has acquired more education than his or her job requires to perform adequately. The notion that a substantial number of workers are overeducated challenges the relevance of more investments in the education of the European workforce. Overeducated workers do not put all their skills to full use, skills that have been attained following costly investments in education and at a great expense to the economy, firms and individuals (Büchel, de Grip, & Mertens, 2003; Budria & Moro-Edigo, 2009). However, it should be noted that it is often difficult to distinguish between overeducation and an upgrading of the skills or educational credentials demanded in the job, as in both cases higher-educated people obtain the jobs that were previously held by lower-educated workers (Borghans & de Grip, 2000). In case of overeducation, this is due to an excess supply of high-skilled workers. In case of upgrading, it is due to the increasing complexity of jobs.

Educational mismatch has been analysed from a variety of perspectives such as educational returns, career mobility, and labour market outcomes. In this dissertation, the influence of health on the overeducation probability is examined. Existing literature has mainly focussed on the effects of ill-health on earnings and labour market participation, concluding that unhealthy individuals are characterized by lower wages and lower participation rates (Stern, 1996; DeLeire, 2001; Madden, 2004; Jones, 2006; Mavromaras, Oguzoglu, Black, & Wilkins, 2006; Cai, 2010). However, poor health does not solely affect the supply of labour. It is also possible that it affects labour market outcomes via the demand of labour. Those with poor health may exhibit a greater chance of being overeducated, given the same initial human capital endowments as those with good health. A possible explanation is that a person's human capital is affected by poor health, i.e., the current or future productivity is reduced. Or, alternatively, there may be discrimination against the unhealthy.

Surprisingly, little attention has yet been paid to the relationship between health and overeducation. This deficit of research is remarkable, as health is a well-established determining factor for many labour market outcomes, e.g. unemployment. However, there is a substantial number of articles linking disability with educational mismatch. A common finding is that disabled workers are more prone to skill mismatch, compared to nondisabled workers (Blazquez & Malo, 2005; Jones, Latreille, & Sloane, 2006;

Jones, Mavromaras, Sloane, & Wei, 2011). Although the effects of being disabled do not fully represent health effects, and skill mismatch is definitely not a synonym for overeducation, the underlying causes are worth examining. Three explanations are often mentioned. First, a mismatch can be viewed as a compensation for a lack of human capital endowments, such as ability, on-the-job training, or experience. Second, unhealthy workers are characterized by a reduced spatial flexibility. Hence, these workers depend highly on the size of the labour market. Third, unhealthy workers may be the victim of discrimination. Employer discrimination reduces the probability of employment, so that unhealthy individuals may be more likely to accept employment which does not fully utilize their skills or qualifications.

Considering these explanations, we expect that unhealthy workers exhibit a higher probability of being overeducated. To test this hypothesis, a simultaneous equation ordered probit model is introduced, as causality between overeducation and health may run in both directions. Overeducation may have adverse impacts on individual's health, and at the same time, individuals may use health conditions to justify their labour force status. Valid inferences based on the reported effect of health on overeducation are then not possible. Hence, this dissertation estimates a model of overeducation in which overeducation and health are jointly endogenous. The data consists of individual observations, retrieved from a combination of the European Social Surveys of 2002, 2004, 2006, 2008 and 2010, as well as country-level data. Using a two-equation ordered probit model, both the reverse effect of overeducation on health and the endogeneity of health arising from unobserved heterogeneity can be estimated.

Within this dissertation, an individual i is defined as being overeducated (OV_i) if his or her educational level (e_i) exceeds the required level of education to be qualified for his or her job (r_i). This operationalisation method, known as the job analysis method, has two major shortcomings. First, formal education is an incomplete measure of human capital (Verhaest & Omey, 2010). Individuals also differ with respect to informal types of skill acquisition such as experience and on-the-job training. Moreover, individuals may be heterogeneous in their ability and the quality of their education. Hence, the concept of overeducation is solely used in the sense of an educational mismatch. The subjective experience of the employee or employer does not matter.

Second, individuals need to have a job before they can be labelled as overeducated. However, it is possible that individuals use health conditions, or the high probability of being overeducated, to justify unemployment or non-participation – a justification hypothesis known in the literature (Cai, 2010). In other words, the relationship between the availability of employment opportunities and overeducation is not straightforward (Büchel & van Ham, 2003). Those in employment comprise a selective group, and analyzing overeducation while restricting the sample to the employed could lead to biased results. Unfortunately, the attempts to estimate a three-equation model, introducing an employment equation in combination with the overeducation and health equations, were not successful. Nevertheless, we estimate a Heckman probit model with sample selection in order to take a closer look at how health may affect the probability of being employed.

The dissertation is organized as follows. Section 2 discusses the different theories concerning the relationship between health and overeducation. The data, methodology and a description of the (in)dependent variables are presented in section 3. Section 4 discusses the results of the empirical analysis. I end the Master dissertation with some general conclusions.

2. Theoretical Framework and Empirical Evidence

In this section, we will first briefly discuss the overeducation phenomenon and its well-known penalizing effects on earnings and job satisfaction. In a second part, we discuss how health may affect the overeducation probability. Although the positive association between education and health is well established, the relationship between overeducation and health is mainly neglected. Therefore, it is difficult to predict how they affect each other, and which underlying causes play a role if true. On the other hand, several studies have examined the influence of disability on the chance of being overeducated, concluding that disabled workers are more prone to mismatches (Blazquez & Malo, 2005; Jones et al., 2006; Jones et al., 2011). It is worthwhile to take a closer look at the underlying determining causes. However, extrapolating these study results is hazardous: unhealthy workers do not necessarily suffer from a disability, and vice versa.

From a macroeconomic point of view, overeducation is defined as a by-product of an inferior occupational structure of the labour market (Dolton & Silles, 2008). If the growth in the supply of higher-educated workers outpaces the growth in demand, overeducation of the workforce is the likely result. In other words, the labour market lacks the capacity to absorb the increased number of educated workers into traditional graduate occupations (Dolton & Silles, 2008). In microeconomic terms, overeducation refers to a situation in which an individual possesses a higher level of education than that which is putatively required or historically typical for a given job (Groot & Maassen van den Brink, 2000).

It has been clearly shown that a significant proportion of workers is overeducated, with estimates for different countries ranging from 10% to more than 40% of the working population (Groot & Maassen van den Brink, 2000; Budria & Moro-Edigo, 2009). This large estimation differential can be attributed to the different techniques employed for the measurement of overeducation (see Verhaest & Omey, 2010 for a discussion). The different techniques also make the consequences and determinants of overeducation difficult to validate (Battu, Belfield, & Sloane, 2000; Verhaest & van der Velden, 2010). For example, the study of Verhaest and Omey (2010), finds that women and school leavers with a highly educated father have a higher likelihood of being objectively overeducated, but not of being subjectively overeducated.

It is not always clear whether overeducation is a desirable or a disadvantageous situation. Theoretically, overeducation does not automatically indicate a problem. This is only the case if the returns to surplus schooling are found to be permanently lower than those to required education (Mehta, Filipe, Quising, & Camingue, 2011). From a public policy perspective, public resources are then wasted, i.e. the excess level of education received by the overeducated worker costs more than the productivity advantage it offers. Pragmatically, several studies have demonstrated that overeducation has negative consequences, i.e. a wage penalty and a reduction in job satisfaction (Dolton & Vignoles, 2000; Hartog, 2000; Allen & van der Velden, 2001; Bauer, 2002; Chevalier, 2003; Dolton & Silles, 2008; Fleming & Kler, 2008). The available income differential evidence shows that, within the same educational level, overeducated workers earn 12% to 27% less than their adequately allocated counterparts (Budria & Moro-Edigo, 2009). However, overeducated workers might still be earning higher wages than had they not acquired their excess education (Dolton & Silles, 2008). Hence, surplus schooling tends to be characterized by the law of diminishing returns.

Prudence is also called for interpreting the reduced level of workplace satisfaction, as the level among the overeducated remains high in absolute figures (Fleming & Kler, 2008). The finding of Fleming & Kler (2008) that overeducated workers are less satisfied compared to their nonovereducated counterparts, but equally satisfied compared to their immediate colleagues, suggests that the overeducated compare themselves to their similarly educated peers who are in non-overeducated jobs. In other words, satisfaction levels should be viewed from a relative perspective, rather than an absolute one.

Only few studies emphasize the positive effects of overeducation. The study of Büchel (2002) for example, finds that overqualified employees are healthier and more productive compared to their correctly allocated colleagues, as they tend to be more work- and career-minded, more likely to participate in on-the-job training, and report longer periods of tenure with the same firm.

In this dissertation, the overeducation incidence rate is calculated by comparing the individual's educational credentials with the educational level that is believed to be required for his or her job. The main consequence of this choice is that we assume that skills can only be obtained through formal education. However, most skills may be, and often are, acquired through experience and informal training. As a result, workers with similar levels of education may have different skill endowments. However, there is also a positive aspect attached to this narrowed interpretation of reality: a clear distinction can be made between educational and skill mismatch. The two concepts are often assumed to be closely related. However, overeducation is not a valid basis for inferring whether a worker's skills and abilities are underutilized on his or her job.

Halaby (1994) states that indicating the underutilization of skills by using measures of overeducation, rests on highly uncertain assumptions about the relation of schooling to skills. There are several analytical issues speaking against the use of overeducation as an indicator of skill mismatch. First, if workers who stand far above the mean of their occupation-specific schooling distribution are assumed to be in jobs that do not fully use their skills, then workers who stand far below the mean are assumed to be in jobs that require skills they do not possess. Second, operationalizing skill mismatch in terms of overeducation confounds causal analysis aimed at determining how schooling itself may be implicated in the structure of forces that shape the risk of a worker ending up in a job that is a poor match to his or her educational level (Halaby, 1994; Green, McIntosh, & Vignoles, 1999). Educational mismatches are thus neither a necessary nor a sufficient condition for skill mismatches. The study of Allen & van der Velden (2001) confirms this statement. They show that skill mismatch is negatively correlated with job satisfaction, while the link between educational mismatch and job satisfaction is found to be insignificant.

There are several factors explaining why some individuals exhibit a higher probability of being overeducated than do others. Individuals might face different informational, financial, time or spatial constraints in their search for a job. The question whether these constraints especially affect the unhealthy or disabled individuals in the labour market is difficult to confirm or falsify, as the relationship between overeducation and health is largely ignored in the literature, even to the extent of not always controlling for health in regression analyses (Jones & Sloane, 2010). Nevertheless, there are several explanations why a poor health status may lead to a higher chance of being overeducated.

A first explanation is that overeducation may function as a compensation for a lack of human capital endowments, such as ability, on-the-job training, or experience (Groot & Maassen van den Brink, 2000). It is plausible that unhealthy workers lack certain skills and abilities that are in fact necessary to obtain a job that matches their educational level. To test this consideration, following the basic assumption of the Human Capital Theory formulated by Becker (1964) – individuals are paid by their marginal product or productivity – a wage penalty for unhealthy, overeducated workers, would be a reflection of a certain lack of human capital endowments (Budria & Moro-Edigo, 2009). The assumption is that the presence of unhealthiness directly reduces a person's ability to work productively. Several studies have showed that there are indeed substantial wage and earnings gaps between people with and without disabilities. DeLeire (2001) and Jones (2006) suggest that the average wage of the disabled is approximately 85% of their nondisabled colleagues. The analysis by DeLeire (2001) further suggests that only a small fraction of this wage gap (5 to 8%) is due to discrimination.

Now, if overeducated workers are only characterized by a lower average ability level compared to adequately allocated workers with a similar education background, the literature on overeducation would be 'much ado about nothing' (Sloane, 2003). In that case, overeducation does not represent market failure, nor can it be taken as an indicator of an overproduction of graduates (Sloane, 2003). Only a limited number of studies have already investigated this problem (Verhaest & Omey, 2012). Nevertheless, these studies provide only limited support for the hypothesis that overeducation is only a sign of lower ability (Budria, 2011). The study of Chevalier (2003) for example, which includes ability related indicators in the earnings equation, still finds substantial wage penalties for overeducated workers.

Another criticism regarding the 'lack of human capital' explanation is formulated by Blazquez & Malo (2005). They find that disabled workers in the Spanish labour market have instead a lower probability of being overeducated. According to the authors, one of the main explanatory reasons is that disabled people present lower educational attainments compared with the rest of society. The reduced overeducation probability is then a logical outcome, as higher educational degrees are presumably found to be important for the incidence rate of overeducation. Consequently, a lower investment in education is rational, as people with disabilities anticipate lower returns with respect to the rest of the population for the same educational levels (Blazquez & Malo, 2005). The relatively lower educational level of unhealthy and disabled people is indeed a 'stylized fact' at international level, especially in southern European countries (Zwinkels, 2001).

However, is a highly educated individual really more prone to overeducation, as Blazquez & Malo (2005) presume? From a macro-level point of view, it is not correct to imply that countries with a more highly educated work force also face more overeducation problems. The study of Verhaest & van der Velden (2010), examining cross-country differences in graduate overeducation and its persistence, did not find any evidence for a relationship between overeducation and the absolute share of highly educated individuals, i.e., it seems that supply might create its own demand.

From a micro-level point of view, an analogous contra-argument can be made. First, several studies have showed that the probability of overeducation decreases with increasing years of education (see, for example: Büchel & van Ham, 2003). Second, it can be expected that extra-educational factors are more important than the degree itself. For example, the study of Verhaest & Omey (2010) finds that academic achievement – graduating with distinction or great distinction – and starting the job search activity before graduation, leads to a significantly lower probability to be overeducated.

A second reason why unhealthy individuals may be more prone to overeducation enhances the spatial constraints that come along with a job search. Not only the geographical location, but also the physical (and emotional) demands of employment, hours of work and accessibility, induces unhealthy and disabled workers to search in a smaller pool of jobs and be more at risk of accepting mismatched employment (Jones & Sloane, 2010). A formal theory closely related to this basic thought is the 'theory of differential overqualification'. This theory, formulated by Frank (1978), states that searching a job in smaller labour markets is accompanied by a higher risk of working in jobs for which current qualifications exceed the educational requirements. Frank (1978) particularly focuses on married women, who face the problem of a dual job search, i.e. the job search of women is undertaken under the condition that the job search of their husbands is optimized. Consequently, married women are forced to search jobs in a smaller pool of jobs compared to their husbands, which leads to a higher overeducation risk.

The lack of empirical support for the differential overqualification hypothesis (see: McGoldrick & Robst, 1996; Büchel & Battu, 2002) does not mean that spatial determinants do not affect the over–education risk. Besides, Frank (1978) focuses exclusively on married women, i.e. his research sample is very limited. In the study of Büchel & van Ham (2003), a more general theoretic framework is presented. The authors argue that workers seek jobs located as closely as possible to their place of residence in order to economize on search and mobility costs. Second, workers seek jobs that require the skills they have acquired. If no suitable jobs are available, and mobility costs to 'escape' are too high, workers might 'bump down' and accept a job below their level of qualification (Büchel & van Ham, 2003). The overeducation risk is, in other words, the highest for those living far away from a large concentration of employment, and/or for those who are not spatially flexible. In sum, Büchel & van Ham (2003) postulate that the smaller the local labour market, the higher the risk of overeducation. Furthermore, they show that spatial flexibility helps workers to overcome a poor local market: access to certain means to travel longer distances helps to avoid overeducation.

For disabled and unhealthy workers, mobility costs are typically high and spatial flexibility is low (Jones & Sloane, 2010). Consequently, it seems rational that these workers depend highly on the size of the labour market and are more prone to overeducation than their nondisabled counterparts. However, we did not find any empirical evidence that endorses this theoretical consideration.

A third explanation why unhealthy workers may be more prone to overeducation is discrimination. Prejudices and lack of information regarding impairments and their consequences may persuade employers that unhealthy workers are less productive, even when they are not. Employer discrimination reduces the probability of employment, so that unhealthy individuals may be more likely to accept employment which does not fully utilize their skills or qualifications (Jones & Sloane, 2010).

The measurement of a possible discrimination effect in the labour market has become widespread and has principally been used to measure discrimination on the basis of gender and ethnic background (Madden, 2004). If an employer employs discrimination in his or her selection of employees, the health status of the latter is viewed as an endogenous characteristic. This is a well-known complication in the measurement of discrimination effects (Madden, 2004; Jones, 2006; Lindeboom & Kerkhofs, 2009). While certain diseases and disabilities are exogenous, ill-health may also be the result of lifestyle choices. The unobservable factors affecting these choices are also likely to affect the job search in the labour market. Thus what is measured as discrimination may in fact reflect an unobserved difference in productivity (Madden, 2004; Jones, 2006). The available evidence is rather surprising. When the unobserved effect of productivity is controlled for, there is no – or only modest – evidence of employ–ment discrimination against the disabled (Madden, 2004; Jones, 2006). According to Madden (2004, p.430), policy should therefore focus on increasing the productivity of the ill-health and disabled.

In sum, the paragraphs above reveal that is not always clear whether unhealthiness leads to a higher probability of being overeducated. The first explanation, which viewed overeducation as a form of compensation for a lack of human capital endowment, is not a sufficient explanation for the hypothesis that disabled or unhealthy workers are more prone to overeducation. Although it is definitely true for some individuals, generalizing it would be incorrect. Moreover, overeducation may be part of a career mobility strategy (Sicherman & Galor, 1990; Sicherman, 1991; Groot & Maassen van den Brink, 2000). If the returns to schooling are lower while one works in a specific occupation, the effect of schooling on the probability and the rapidity of being promoted within or across firms will be higher (Sicherman & Galor, 1990; Sicherman, 1991). In other words, workers may deliberately enter the labour market in jobs for which they are overeducated and later on move to jobs that more match their educational attainments. The second explanation, stating that unhealthy individuals are affected by several spatial constraints in order to find a job that matches their educational credentials, is merely a theoretical one. Although there are no real objections against the basic thought – unhealthy individuals are more prone to overeducation because they have to find a job in smaller labour market and are less spatially flexible – this point of view can neither be confirmed due to a lack of empirical studies. The third explanation, introducing a potential discrimination effect, has been (partly) falsified by the studies of Madden (2004) and Jones (2006). Nevertheless, taking into account these three constraints that unhealthy workers might have to overcome in order to find a suitable job, leads to the following hypothesis: unhealthy workers exhibit a higher probability of being overeducated (H1).

3. Data and Methodology

3.1. Database and description of the variables

The data used in this Master dissertation comprise a combination of the European Social Survey (ESS) waves of 2002, 2004, 2006, 2008 and 2010. Because overeducation is measured by comparing the acquired level of education with the level of education considered most appropriate for the job – the job analysis method – only working respondents are included in the analyses. As a result of this selection procedure, students, unemployed and retired people, and armed forces[1] are excluded from the analysis. Note that analyzing overeducation while restricting the sample to the employed could lead to biased results, as individuals may use health conditions to justify unemployment or non-participation. Unfortunately, the attempts to estimate a three-equation model, introducing an employment equation in combination with the overeducation and health equations, were not successful. Nevertheless, we also present a Heckman probit model with sample selection, using the ESS data wave of 2010, in order to take a closer look at how health may affect the probability of being employed.

[1] The occupation class 'armed forces' is excluded from the analysis, because other aspects of the type of work were considered more important as similarity criteria, i.e. military duties.

After data cleaning across all the relevant variables, the resultant dataset comprises 108 211 useable observations. These are spread over thirty countries: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Germany, Estonia, Finland, France, Greece, Hungary, Ireland, Israel, Italy, Luxembourg, the Netherlands, Norway, Poland, Portugal, Russian Federation, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, and United Kingdom.

(i) Overeducation

Two general types of overeducation definitions have been used regularly in the existing literature: 'objective' and 'subjective' definitions (see Groot & Maassen van den Brink, 2000, and Verhaest & Omey, 2010 for a discussion). The subjective methods are based on individual workers' self-reports on their skill utilization. Either workers are asked directly whether they are overeducated (or undereducated) for the work they do, or workers are asked what they think the minimum educational requirements are for their job. In the latter case, the self-report on the required educational level is compared with the actual level to assess the job match. The objective definitions can also be classified into two types. In the first, the statistical method, overeducation is defined by comparing the years of education attained with the average years of education within the current occupation of the worker. A worker is labelled as overeducated if he or she has more years of education than the average years of education within the worker's occupation plus one standard deviation (Groot & Maassen van den Brink, 2000). The second objective definition, the job analysis method, is based on a comparison between the actual education level and job level requirements. The latter is usually determined by labour market experts. All of the definitions have their drawbacks and limitations. Moreover, Verhaest & Omey (2010) have demonstrated that the choice of the overeducation measure clearly influences the outcomes of the analysis.

Because the ESS questionnaire does not contain any subjective questions regarding overeducation, and the statistical method tends to underestimate the incidence rate, the job analysis method is used as operationalisation technique. The ESS surveys make use of the International Classification of Occupations 88 (ISCO88), designed by the International Labour Office (ILO), to structuralize the heterogeneity of occupations (ILO, 2011). On the one-digit-level, nine occupational classes are distinguished, each linked with a certain skill level. These classes, presented in Table 1, are aggregated categories, consisting of occupations grouped together on the basis of the similarity of skills required to fulfil the tasks and duties of the jobs. The skill level, ranging from 1 (easy) to 4 (complex), reflects the range and the complexity of the tasks involved. Each skill level is linked with a certain education level, defined by the International

Standard Classification of Education 97 (ISCED97). This nomenclature, designed by UNESCO, distinguishes seven education levels: (1a) pre-primary education; (1b) primary education or first stage of basic education; (2) lower secondary or second stage of basic education; (3) upper secondary education; (4) post-secondary non-tertiary education; (5) first cycle of tertiary education; (6) second cycle of secondary education (UNESCO, 2006; Schneider, 2007). By comparing the acquired level of education (e_i) with the level of education considered most appropriate for the job (r_i), overeducation (OV_i) can be measured. Mathematically, this can be written as follows:

$$\mathrm{OV}_i = \begin{cases} 0 & \text{if } e_i \leq r_i \\ 1 & \text{if } e_i > r_i \end{cases}$$

Table 1 ISCO88 and ISCED97 classifications

ISCO88 occupation class	**ILO skill level**	**ISCED97 education level**
1. legislators, senior officials and managers	4	6
2. professionals	4	5
3. technicians and associate professionals	3	4
4. clerks	2	3
5. service workers and shop and market sales workers	2	3
6. skilled agricultural and fishery workers	2	3
7. craft and related trade workers	2	3
8. plant and machine operators and assemblers	2	2
9. elementary occupations	1	1

Sources: Hoffmann (1999); ILO (n.d.); Schneider (2007). How to interpret the Table? Each ISCED97 education level figure is a minimum level required to do the job, with 1=primary or first stage of basic education; 2=lower secondary education; 3=upper secondary education; 4=post-secondary non-tertiary education; 5=first stage of tertiary education; 6=second stage of tertiary education.

So, if the individual's educational level is higher than the level considered most appropriate for the job, he or she is defined as being overeducated; if not, he or she was classified as being correctly allocated (validating the phenomenon of undereducation lies beyond the scope of this dissertation). After weighting, 30% of the respondents in the research sample were identified as being overeducated. There is a small gender difference: 29% of the employed men and 31% of the employed women are overeducated. Due to the fact that legislators, senior officials and managers require the highest level of education to be qualified for their jobs, no one is overeducated in this occupation class. The highest overeducation percentage is found in the elementary occupation class: 80.2%.

(ii) Health

Health is measured by asking respondents directly how they evaluate their health status. Answer categories range from very good, good, fair, bad, to very bad. The key potential problem with self-rated health (SRH) is that individuals might have incentives to strategically answer this personal question. They might feel obligated to justify some of their observed actions, or might question the confidentially of the survey and therefore overrate their health status (Benitez-Silva et al., 2004). So despite the measure's popularity, debates continue as to what SRH exactly captures. While numerous studies show good agreement between SRH and more objective measures of health, other studies suggest a discrepancy between them. According to Layes, Asada, & Kephart (2012), SRH can be conceptualized as consisting of two components: latent health status and reporting behaviour. Latent health status is defined as "the value assigned to the duration of life as modified by the impairments, function states, perceptions, and social opportunities that are influenced by disease, injury, treatment, or policy" (Layes et al., 2012, p.1). Reporting behaviour is the systematic measurement error between SRH and the latent health status. The literature provides evidence that the individual's reporting of their health is especially influenced by the knowledge about one's own health, social norms and expectations for illness, and adaptability to and acceptability of illness (Layes et al., 2012).

Unfortunately, the merged ESS dataset we use does not contain any variables that could bring this reporting error into account. Therefore, it should be clear that the health variable is possibly biased. However, the incorporation of a subjective experience component can also be viewed as an enhancement: health is measured as a state of well-being, not simply as the absence of disease. This endorses the World Health Organization's definition of health: "health is a state of complete physical, mental, and social well-being and not merely the absence of disease of infirmity" (WHO, 2011a).

Due to the fact that the category 'very bad health' contains only a few respondents (N=300, or 0.3%), it has been merged with the category bad health status. This modification leads to a health variable consisting of four groups: those with a bad or very bad health status (2.8%), a fair health status (21.4%), a good health status (48.5%), and a very good health status (27.4%). The latter is used as reference category. Table 2 shows the distribution of these categories over the overeducation variable.

Table 2 Percentage overeducated individuals per health category

Health	Overeducation	
	yes	no
bad or very bad health (N = 3 054)	32,3%	67,7%
fair health (N = 23 682)	32,4%	67,6%
good health (N = 53 874)	28,6%	71,4%
very good health (N = 27 601)	30,3%	69,7%

Source: own calculations based on working respondents in ESS2002, 2004, 2006, 2008 and 2010. N = 108 211.

Taking into account the nested nature of the data, several country characteristics are also included in the analysis. One of these characteristics is the country-specific percentages of the following education categories: upper secondary educated workers, post-secondary non-tertiary educated workers, first stage of tertiary educated workers, and second stage of tertiary educated workers. The country percentages of primary and lower secondary education are omitted in order to avoid perfect collinearity. Another indicator of the national educational system we include is the number of students per teacher. Ratios for both primary and secondary education are introduced (UNESCO, 2010). A third country variable is GDP per capita, measuring the gross domestic product or the value of all final goods and services produced within a country in a given year divided by the national population for the same year. Given national differences in prices, the national GDP per capita figure is measured in purchase power parity (PPP), using the dollar exchange rate of 2005 (UN, 2011). A fourth country variable measures the number of physicians per 1000 habitants. Figures are obtained from the World Health Organisation (WHO, 2011b). Under the term physician, WHO understands medical doctors, including generalist and specialized medical practitioners. A fifth and last country variable takes into account the amount of air pollution, measured as ton CO_2 per capita (The World Bank, 2011). Summary statistics and definitions of all these variables are presented in Table 3.

3.2. Model

Theoretical considerations have two important implications for the methodology used in this dissertation. First, as mentioned in the previous paragraph, the cases are clustered by country. Therefore, the data includes both individual level and country characteristics, and there is more than one

Table 3 Variable summary statistics and operationalisation/definition

Variable	Operationalisation/definition	Mean	St. dev.
Dependent: overeducation	1 if overeducated	0.300	0.458
Dependent: health status	1 = (very) bad health, 2 = fair health, 3 = good health, 4 = very good health (ref.cat.[a])	3.004	0.773
Educational level	1 = primary or first stage basic education, 2 = lower secondary education, 3 = upper secondary education, 4 = post-secondary, non-tertiary education, 5 = first stage tertiary education (ref.cat.[a]), 6 = second stage tertiary education	3.525	1.339
Age	years	41.758	11.860
Age squared	years	1884.35	1020.49
Foreigner	1 if foreigner	0.083	0.275
Work experience	age – years of schooling	28.510	12.726
Work experience squared	(age – years of schooling)²	974.791	784.979
Intimate contact	1 if yes	0.921	0.270
Type of residence	1 = farm or home in countryside, 2 = village, 3 = town or small city, 4 = suburbs, 5 = big city (ref.cat.[a])	3.153	1.235
Marital status	1 = never married, 2 = widowed, 3 = divorced, 4 = separated, 5 = in civil partnership, 6 = married (ref.cat.[a])	4.282	2.184
Gender	1 if female	0.466	0.499
Cohabitation	1 if yes	0.691	0.462
Children	1 if children living at home	0.510	0.500
Female*Cohabitation	interaction between female and cohabitation	0.308	0.461
Female*Children	interaction between female and children	0.250	0.433
ESS wave	1 = ESS2002, 2 = ESS2004, 3 = ESS2006, 4 = ESS2008, 5 = ESS2010 (ref.cat.[a])	3.070	1.403
Country percentage per education category	lower secondary education	0.138	0.072
	upper secondary education	0.410	0.157
	post-secondary non-tertiary education	0.035	0.034
	first cycle tertiary education	0.307	0.115
	second cycle tertiary education	0.047	0.025
Student-teacher ratio primary	country-specific ratio of number of students per teacher in primary school	13.10	3.160
Student-teacher ratio secondary	country-specific ratio of number of students per teacher in secondary school	10.46	1.200
GDP per capita	GDP per capita (PPP dollars) (divided by 1000)	28.236	10.023
GDP*age	interaction between GDP per capita and age	1275.66	5.661
Physician density	number of physicians per 1000 national population	3.503	0.776
Physician density*age	interaction between physician density and age	146.535	53.402
Air pollution	ton CO_2 per capita	8.580	2.698
Air pollution*age	interaction between air pollution and age	358.191	153.642

Source: own calculations based on the working respondents in ESS2002, 2004, 2006, 2008 and 2010. N = 108 211.
[a] used as reference category.

respondent per country in the sample. Ignoring the nested nature of the data would violate the standard assumption of independence of observations that underlies traditional regression models. Taking this clustering phenomenon into account, by letting the STATA command control for it, solves this problem.

Second, the relationship between overeducation and health is not straightforward. In this Master dissertation, it is hypothesized that overeducation is a function of health and simultaneously, health is a function of overeducation. A generic two-equation ordered probit model is used to control for this reversed causality. Estimations of joint probability distributions of a bivariate and an ordered categorical variable are less common in the literature. Similar to univariate ordered probability models, multivariate ordered probability models can derived from a latent variable model (Verbeek, 2008; Roodman, 2009). Consider the following equations (1) and (2), where y_{1i}^* represents overeducation and y_{2i}^* the individual's health status[2]:

$$y_{1i}^* = \gamma_1 y_{2i}^* + x_{1i}' \beta_1 + \epsilon_{1i} \qquad (1)$$

$$y_{2i}^* = \gamma_2 y_{1i}^* + x_{2i}' \beta_2 + \epsilon_{2i} \qquad (2)$$

where γ_1 and γ_2 are unknown parameters, β_1 and β_2 are vectors of unknown parameters, ϵ_{1i} and ϵ_{2i} are the error terms, and subscript *i* denotes an individual observation. As can be seen from these equations, y_{1i}^* and y_{2i}^* simultaneously determine each other. Changes in y_{2i}^* will lead to changes in y_{1i}^* via (1). The resulting changes in (1) will immediately lead to changes in y_{2i}^* via (2). Variables that display such relationships are termed endogenous variables. The remaining variables are termed exogenous. By itself, endogeneity is not an obstacle. The problem is that models containing such variables cannot be estimated by typical estimation procedures such as ordinary least squares (OLS), because the relationship specified by the equations violates the OLS assumption of zero covariance between the disturbance term and the independent variables. Thus, y_{2i}^* is correlated with ϵ_{1i}, indirectly. The same is true for y_{1i}^* and ϵ_{2i}. Estimation of such models via OLS, will lead to biased, inconsistent, and overestimated coefficients. The inconsistency is the most important problem, since, no matter the sample size, the coefficients will never converge to the true population coefficients.

The parameters in the systems of equations (1) and (2) are identified only by imposing an exclusion restriction on vectors x_1 and x_2. In other words, at least one element of x_2 should not be present in x_1 in order to avoid perfect collinearity. If there are some variables that are believed to be correlated with y_{2i}^*

[2] The following discussion borrows from Keshk (2003) and Sajaia (n.d.).

(latent variable measuring health) but independent of ϵ_1, these variables could be included in x_2 to obtain consistent estimates of γ_2, β_2 and ρ. Rho (ρ) stands for the correlation between the error terms ϵ_{1i} and ϵ_{2i}. It is assumed that these error terms are distributed as standard normal.

So, the key element of this strategy is to select at least one variable that affects the individual's health status, but not his or her overeducation risk. The difficulty here is that we can expect very similar factors to influence both the health status and the overeducation risk. There is no real formal econometric test that could indicate the correct specification of the model. In this dissertation, two strategies are implemented to justify exclusion restrictions: (a) substantial theoretical arguments; and (b) by using identical sets of regressors in both equations and identify which ones are only significant for the health equation.

The estimation of model 1 (Table 4) is executed by implementing strategy (a). Intimate contact, GDP per capita, physician density, air pollution, and the interaction terms of the last three country-specific variables with age, are selected to function as exclusion restrictions. The relationships between these variables and health are discussed in the paragraphs below. We also motivate why overeducation shouldn't be affected by them. To get an idea of how the model would be without any exclusion restrictions, model 2 (Table A1 in Appendix A) includes all the independent variables in both equations. In the third model (Table A2), exclusion restrictions are re-introduced by applying method (b) to model 2. Four exclusion restrictions were derived: physician density, the interaction effect between physician density and age, and the marital status dummies 'never married' and 'divorced'. Note that only physician density and its interaction effect with age are justified by both methods.

All models were estimated using maximum likelihood estimation (MLE). MLE is a method commonly employed in simultaneous equation models. As an alternative to ordinary least squares used in multiple regression, MLE is a procedure that iteratively improves parameter estimates to minimize a specified fit function (Keshk, 2003).

Now, in which way are the exclusion restrictions employed in model 1 associated with health? With respect to physician density, it is well established that an easily accessible physician has a positive impact on the health status (Sarma & Peddigrew, 2008; Cooper, 2009; Subramanian & Canning, 2009). On a short-term basis, physicians have a crucial role in the implementation of new technologies, in the form of new vaccines, drugs and medical procedures. These are, in turn, a major source of health improvement (Cooper, 2009; Subramanian & Canning, 2009). Physicians can also act as catalysts in motivating change in a patient's lifestyle or by telling people how to take care of themselves, which most noticeably affects morbidity and mortality in the long run (Galuska, Will, Serdula, & Ford, 1999; Fogel, 2004).

With respect to air pollution, measured as ton carbon dioxide (CO_2) per capita, unfavourable effects on the respiratory and the cardiovascular systems from short-term and long-term inhalation are well documented (for a more extensive overview of possible health effects, see: Robertson, 2006; Brunner & Maruyama, 2011; Dhondt et al., 2012; Rundell, 2012).

More important, we argue that physician density and air pollution only affect the health system, i.e. it is unlikely that they affect the probability of being overeducated. This statement is verified by invalidating the following contra-argument. It is possible that modern economies, with a high GDP per capita, a highly educated work force – and possible higher overeducation levels – can invest in a high-quality health system and green, low-emission technologies. In other words, it is imaginable that physician density, air pollution and overeducation are interrelated. However, we argue that any interconnectedness is rather improbable, as a large supply of educated workers does not necessarily imply more overeducation problems. Supply might create its own demand with some lags (Verhaest & van der Velden, 2012). Moreover, the study of Verhaest & van der Velden (2012), examining cross country differences in overeducation and its persistence, finds that the structural imbalance between the overall demand and supply of skilled workers and between the demanded and supplied fields of study, are more important predictors of overeducation figures than the supply of educated workers itself.

With respect to GDP per capita, it is generally well accepted that the population in countries with higher levels of GDP will have better health and longer life expectancy, as higher living standards lead to enhanced prevention and treatment of disease (Smith, 1999). This has been empirically verified by Swift (2011), who examines the relationship between health and GDP per capita for thirteen OECD countries over the last two centuries. She proves that GDP per capita especially affects life expectancy, and that there is also an important reverse effect: better health can lead to a growing GDP per capita (economic growth) through long-term gains in human and physical capital that raise productivity.

Although we did not find any studies validating a possible effect of GDP per capita on the over–education probability, it is again plausible that both concepts are interrelated. For example, it is likely that countries with high GDP levels invest proportionally more in the educational system, compared to countries with lower GDP levels. Consequently, high GDP countries are characterized by a larger supply of educated workers, which in turn may lead to an increased overeducation risk. However, we refer again to the study results of Verhaest & van der Velden (2012). If a growing supply of educated worker is accompanied by a growth in demand, higher overeducation figures are unlikely. Hence, labour market characteristics are far more important predictors of overeducation than GDP per capita, physician density and air pollution.

With respect to intimate contact, there is substantial proof that social support, as a sense of being cared for and loved, being esteemed and valued as a person, and being part of a network of communication and mutual obligation in which other can be counted on, positively affects health (Cobb, 1976; Ross & Wu, 1995; White, Philogene, Fine, & Sinha, 2009). A broad social network might reduce the chance of being overeducated, as there is more information about labour market opportunities, which in turn lead to making better decisions about careers (Kucel & Byrne, 2008). However, we argue that intimate contact only partly reflects the social network that one disposes, and therefore does not affect overeducation. Moreover, the study of Kucel & Byrne (2008, p.16) demonstrates that by applying for jobs through specialized private employment agencies, as well as replying to employers' advertisements, one can secure a better job match than by relying on personal contacts.

4. Results

Three models were estimated. In the first model (Table 4), coefficients are estimated with exclusion restrictions, i.e., the overeducation equation is estimated without the variables discussed above. To get an idea of how the model would be without any exclusion restrictions, model 2 (Table A1 in Appendix A) includes all the independent variables in both equations. In the third model (Table A2), exclusion restrictions are re-introduced, but they are only based by applying method (b) to model 2, i.e., the regressors who are statistically significant for the health equation but insignificant for the overeducation equation are not included in the latter.

Table 4 presents the results of model 1. The significant correlation coefficient rho (ρ = 0.287) indicates that there is a positive association between the error terms of the overeducation and the health equations. A significant ρ means that the use of a simultaneous equation ordered probit model is justified. The positive sign indicates that workers with a high probability of being overeducated are likely to report better health, and vice versa. In other words, for some people, working in a job for which they are overeducated may be a strategy to improve their health status.

With respect to the independent variables, most variables are found to be significant at the $p < 0.01$ level. Most importantly, working individuals with a bad or very bad health, a fair health or a good health have a higher chance of being overeducated, compared to individuals with a very good health. Moreover, this chance increases as the health status deteriorates. This means that the central hypothesis of this dissertation – unhealthy workers exhibit a higher probability of being overeducated – is confirmed.

The probability of being overeducated increases with the level of education: compared to individuals holding a bachelor degree (first cycle tertiary education), individuals with a master degree (second cycle tertiary education) report a higher probability of being overeducated. The opposite is true for individuals holding a lower or upper secondary degree. This finding may indicate that the European labour markets are not yet adjusted to the increasing number of highly educated workers.

The probability of being overeducated is non-linearly related to age: with increasing age, the probability of being overeducated first decreases and then increases. This means that, compared to people of intermediate age, younger and older workers run an above-average risk of overeducation. The career mobility theory by Sicherman & Galor (1990) states that graduates often start their occupational career with a job for which they are overqualified, whereas older employees face the problem of continuing obsolescence of the skills that they acquired in initial schooling and training (Blechinger & Pfeiffer, 2000). The significant effect of work experience seems to endorse these explanations: with increasing work experience, the probability of being overeducated decreases. This can be linked to the finding that overeducation sometimes functions as a compensation for a lack of work experience and other human capital elements such as on-the-job training (Groot & Maassen van den Brink, 2000).

Foreigners – individuals who are not officially a citizen of the country they live in – have a significant higher chance of being overeducated. Although a profound analysis of the relative position of ethnic minorities[3] lies beyond the scope of this paper, one study in particular is worth mentioning. Battu & Sloane (2004), analyzing the case of ethnic minorities in Britain, find that non-whites have a higher probability of being overeducated than whites. They argue that this difference can be explained by discrimination effects and spatial constraints: some employers may only hire members of ethnic minorities who possess higher educational levels than whites for the same job, or it may be possible that commuting is harder for isolated ethnic communities, which reduces the chance of a correct match.

[3] Note that individuals who don't possess the nationality of the country they live in *cannot* be automatically viewed as member of an ethnic minority.

Compared to city dwellers, people living in rural areas (farm, village, town or small city) have a higher chance of being overeducated. This finding can be linked to the study results of Büchel & van Ham (2003), which are worth repeating here (for a more extensive overview, see section 2). The authors argue that the individual spatial flexibility and the spatial distribution of suitable job opportunities largely determine the risk of overeducation. The theoretical framework is confirmed by their empirical results. For example, they find that the longer the travelling time to a large agglomeration, the higher the probability of overeducation.

There are no remarkable patterns regarding marital status, although there is one exception: widowed people have a higher chance of being overeducated compared to married people. More interesting is the effect of gender: women are characterized by a significant lower chance of being overeducated, and this finding remains consistent over all the models. Although mixed empirical support for the differential overqualification theory formulated by Frank (1978), some studies have effectively reported a gender effect regarding overeducation. In their meta-analysis, Groot & Maassen van den Brink (2000) suggest that overeducation is more frequent among female workers than among male workers, while the opposite holds true for undereducation. On the other hand, there are also studies in literature that did not note a gender effect (see: Dolton & Silles, 2002), while others find a similar negative effect of being female on the chance of being overeducated (see: Blazquez & Malo, 2005). Nevertheless, as McGoldrick & Robst (1996) state, this finding might be the result of requirement heterogeneity within occupations and simply indicates that, even within occupations, men take the best positions.

Another rather unexpected outcome is the negative effect of cohabitation: cohabitating has a positive effect on the chance of being correctly allocated. A possible explanation is that cohabitating indicates that one is settled, i.e. lives near a job he or she is likely to maintain. The income maximizing strategy of a job-seeking individual – find a job where the extent of overqualification is minimized – is successfully completed (Battu & Sloane, 2004). The negative effect of having children living at home, although it's almost not statistically significant at $p < 0.10$, can also be interpreted as such. The positive coefficient of the interaction effect between female and cohabitation indicates that the effect of gender for people living together with someone is 0.047 times higher than for people who are not cohabitating. This line of thought can be reversed: the effect of cohabitating is 0.047 times higher for women than for men. Since the main effects of gender and cohabitating are negative, 'higher' means they become smaller. Consequently, the difference between men and women regarding the probability of being overeducated is smaller when one cohabitates. In other words, cohabitating has a more penalizing effect for men than

Table 4 Model 1: ordered probit model with exclusion restrictions

	Overeducation equation		Health equation	
	Coefficient	Robust std. error[a]	Coefficient	Robust std. error[a]
Individual characteristics				
Bad or very bad health	1.179***	0.086		
Fair health	0.802***	0.057		
Good health	0.380***	0.031		
Low secondary education	–0.568***	0.015	–0.069***	0.012
Upper secondary education	–0.695***	0.011	–0.054***	0.009
Post-sec. non-tert. education	0.292***	0.022	–0.047**	0.019
Second cycle tert. education	1.073***	0.025	–0.035*	0.018
Age	–0.019***	0.005	–0.007	0.004
Age squared	0.0002***	0.000	–0.0002***	0.000
Foreign origin	0.137***	0.016	–0.027**	0.013
Work experience	0.010***	0.003	–0.029***	0.003
Work experience squared	–0.0004***	0.000	–0.0001	0.000
Intimate contact			0.198***	0.013
Farm or countryside	0.078***	0.020	0.034**	0.016
Village	0.059***	0.012	–0.035***	0.010
Town or small city	0.063***	0.012	–0.058***	0.010
Suburbs	0.024	0.015	–0.006	0.013
Never married	0.014	0.017	–0.032**	0.014
Widowed	0.065**	0.033	–0.039	0.027
Divorced	0.010	0.020	–0.057***	0.016
Separated	–0.035	0.038	–0.022	0.031
Partnership	0.039*	0.023	–0.071***	0.019
Female	–0.078***	0.016	–0.144***	0.013
Cohabitation	–0.063***	0.019	–0.021	0.015
Children	–0.025*	0.015	0.017	0.011
Female*cohabitation	0.047**	0.021	0.060***	0.017
Female*children	0.013	0.019	–0.003	0.015
ESS2002	–0.248***	0.015	–0.111***	0.012
ESS2004	–0.186***	0.014	–0.108***	0.011
ESS2006	–0.164***	0.014	–0.129***	0.011
ESS2008	–0.106***	0.013	–0.031***	0.011
Country characteristics				
% upper secondary educ.	0.981***	0.037	0.152***	0.028
% post-sec. non-tert. educ.	0.941***	0.189	–4.274***	0.144
% first cycle tertiary educ.	1.676***	0.051	0.222***	0.042
% second cycle tertiary educ.	1.491***	0.237	–2.425***	0.192
Student-teacher ratio prim.	0.007***	0.002	–0.024***	0.002
Student-teacher ratio sec.	–0.019***	0.003	0.020***	0.002
GDP per capita			–0.002	0.001
GDP per capita*age			0.0006***	0.000
Physician density (per 1000)			0.213***	0.016
Physician density*age			–0.003***	0.000
Air pollution			–0.005	0.005
Air pollution*age			–0.0005***	0.000
Constant	–0.927***	0.082		
Cut value_2_1			–2.418***	0.097
Cut value_2_2			–1.092***	0.096
Cut value_2_3			0.341***	0.096
***Log pseudolikelihood* = -170361.09**				
Correlation coefficient (ρ^b) =	0.287***	0.023		

Source: own calculations based on the working respondents in ESS2002, 2004, 2006, 2008 and 2010. N = 108 211. Note: * p < 0.10; ** p < 0.05; *** p < 0.01.

[a] standard errors adjusted for clustering on country level.

[b] Wald test for both equations: chi-square = 16 261.40; d*f* = 36; p = 0.000.

for women. This finding sharply contrasts with the baseline assumptions of the theory of differential overqualification, formulated by Frank (1978).

It seems that the probability of being overeducated has increased during the last ten years: ESS data of 2002, 2004, 2006 and 2008 show significantly lower probabilities of overeducation compared to the ESS data of 2010. This increasing incidence might be related to different growth rates in the supply and in the demand of educated workers. Looking at the country-specific percentage of each education category, it seems that the higher the percentage of educated workers, the bigger the proportion overeducated workers. However, as we already stated, assuming that countries with a more highly educated work force also face more overeducation problems is not correct. As Verhaest & van der Velden (2012) state, supply might create its own demand with some lags. At last, the country-specific student-teacher ratios for primary and secondary education indicate that a low ratio (few students per teacher) in primary school and a high ratio in secondary school result in a lower chance of being overeducated when entering the labour market. The negative effect of the student-teacher ratio in secondary school is rather surprising, since low student-teacher ratios can be viewed as an indicator of high-quality education, and a high-quality educational system presumably reduces the chance of overeducation. Nevertheless, it seems that educational institutions also play a role with respect to overeducation.

The coefficients of the health status determinants are only briefly discussed, as this equation merely functions as an instrument to purify the coefficients of the independent variables in the overeducation equation. Compared to individuals holding a bachelor degree, individuals with lower educational credentials exhibit a higher probability of reporting a poor health. According to Ross & Wu (1995), education improves health indirectly through work and economic conditions, social-psychological resources and health lifestyle. Compared to poorly educated individuals, the well-educated are more likely to be full-time employed, to have fulfilling, subjectively rewarding jobs, earn high incomes and experience low economic hardship. They also report a greater sense of control over their lives and their health, and they have higher levels of social support. The well-educated are also more likely to exercise and to drink moderately.

The possibility of having intimate contact, defined as the opportunity to privately discuss personal and intimate matters with someone, is associated with a better health. The finding is in line with the existing literature (Cobb, 1976; Ross & Wu, 1995; White et al., 2009). In all three models, the coefficient of intimate contact is significant at the $p < 0.01$ level.

Concerning the exclusion restrictions at country level, the analysis shows that GDP per capita and physician density are positively associated with health, while air pollution has a negative impact. The effects of physician density and air pollution are more or less self-evident: the more doctors, the less distance one has to travel when a physician is needed, i.e. the higher the probability of being cured or helped. The negative coefficient of the interaction term between age and physician density indicates that the negative health effect of becoming older is aggravated when physicians are easily accessible. This rather odd finding perhaps indicates a reverse causality: it is possible that less healthy individuals move to areas where there are enough physicians in order to get treated on a short-term basis. Analogous to the negative coefficient of the interaction term between physician density and age, the negative interaction effect of age and air pollution indicates that the obsolescence process and its negative health consequences are aggravated. The positive influence of GDP per capita on health confirms with the literature, as it is generally well accepted that individuals in countries with higher levels of GDP will have a better health and longer life expectancy, as higher living standards lead to enhanced prevention and treatment of disease (Smith, 1999; Swift, 2011).

The cut-off or threshold values of the health equation in model 1 imply that a value of the latent variable less than -2.418 corresponds to a bad or very bad health, a value between -2.418 and -1.092 to a fair health, a value between -1.092 and 0.341 to a good health, and a value above 0.341 to a very good health. Notice that the predicted value of y_{2i}^* for the reference individual, where all of the explanatory variables equal zero, is zero. This value lies between -1.092 and 0.341. Hence, the reference individual would be predicted to report a good health. These cut-off values are consistent over the three models.

The interpretation and the magnitude of the effects of the explanatory variables are also consistent over the three models. However, there is one important difference. The differentiated chances for each health category with respect to overeducation are smaller in model 2 and model 3 compared to model 1. For example, while individuals with a bad or very bad health report a higher overeducation chance of 1.179 units compared to individuals with a very good health in model 1, this chance has shrunken to 0.720 units in model 2 (0.344 units in model 3). Looking at the coefficients in model 3, the differentiated chances have become almost non-significant. Moreover, ρ is insignificant in model 3. This means that overeducation does not affect health and vice versa. Hence, the overeducation equation can be non-simultaneously estimated (see Table C1). In sum, these differences indicate that the choice for certain exclusion restrictions clearly affects the interpretation of the model.

The confirmed hypothesis that unhealthy workers exhibit a higher probability of being overeducated is limited due to the following shortcoming: applying the job analysis method to measure overeducation, means that individuals need to have a job before they can be labelled as overeducated. However, as already mentioned in the introduction, the relationship between the availability of employment opportunities and overeducation may not be straightforward. As Büchel & van Ham (2003) argue, unemployment can be chosen as a strategy to avoid overeducation. We are aware of the fact that analyzing overeducation while restricting the sample to the employed can lead to biased results. We tried to estimate a three-equation model using all five ESS data waves, introducing an employment equation in combination with the overeducation and health equation. Unfortunately, STATA reported computational problems. Nevertheless, to take a closer look at the sample selection bias, we introduce a Heckman probit model with sample selection, using the ESS data wave of 2010. First, the probability of being employed is calculated. Second, the probability of those in employment being overeducated is analyzed. Because the dependent variables in both the selection and the overeducation equations are binary, a bivariate probit model with sample selection is proposed.

Similar to a simultaneous equation ordered probit model, at least one instrument variable has to be selected that is believed it only affects the employment decision, i.e. does not affect the overeducation risk. We have selected the interaction effect between female and children as exclusion restriction. A first argument in favour of this choice is that this interaction effect has no significant effect on the overeducation risk in model 1, 2 and 3 (cf. method (b)). Second, although the labour force participation of women has increased strongly over the past few decades, it can be expected that women still find it more difficult to combine work and family responsibilities such as childcare. Hence, there is a possible trade-off between employment and having children, which has been empirically confirmed by OECD studies (OECD, 2004). The results of the bivariate probit model with sample selection are presented in Table B1 and briefly discussed below.

First of all, the significant coefficient of rho (ρ = –0.435) shows that there is a strong negative correlation between the error terms of the employment and overeducation equations. This indicates that workers with a low probability of employment run a high risk of overeducation when they find a job. This confirms the finding of Büchel & van Ham (2003), who argue that overeducation may function as a strategy to avoid unemployment, and vice versa.

The results of the selection equation, presented in the right column of Table B1, are all in line with the existing literature. Poor health decreases the probability of being employed. Foreigners also report lower probabilities of finding a job. The higher one is educated, the more chance of being employment. Age is non-linearly related to employment: with increasing age, the probability of being employed first increases and then falls. Having intimate contact increases the chance of finding a job. The place of residence, marital status and the country-specific variables do not seem to matter in the finding of a job. There are also no significant gender differences. Cohabitating and having children living at are associated with a higher probability of being employed. With respect to the exclusion restriction – the interaction effect between female and children – the negative, significant coefficient indicates that for women, having children living at home, decreases the probability of employment. This is in line with the study results of OECD (2004), i.e. women are characterized by a trade-off between employment and having children. Moreover, the significant interaction effect between female and cohabitation indicates that women are also affected by a trade-off between employment and cohabitation.

The results of the overeducation equation are presented in the left column of Table B1. The majority of the estimations are in line with the previous models. Most important, people with a poor health run a higher risk of overeducation, i.e. our hypothesis is again confirmed. The probability of being overeducated increases with higher educational credentials. The same is true for being a foreigner. Overeducation is again non-linearly related to age: with increasing age, the probability of being overeducated first decreases and then increases. There is only a small gender difference: women report a lower chance of being overeducated. Cohabitating seems to reduce this chance, while there are no remarkable patters with respect to marital status. The significant interaction effect between female and cohabitation is again positive, indicating that cohabitating has a more penalizing effect for men than for women. With respect to the country-specific variables, higher levels of educated workers are again associated with bigger proportions overeducated workers. The student-teacher ratios for primary and secondary education indicate that a low ratio in primary school and a high ratio in secondary school result in a lower chance of being overeducated when entering the labour market.

Some results are in contrast with previous models. Work experience, having children living at home, and place of residence are not significant. In other words, there is no evidence that spatial determinants of overeducation play a role, which sharply contrasts with the study results of Büchel & van Ham (2003). GDP per capita, physician density, air pollution and the interaction effects of these variables with age are also not significant. This can be viewed as a justification for their use as exclusion restrictions in model 1. In sum, the majority of results are in line with the previous models, i.e., the sample selection bias seems to be limited.

5. Conclusion

This dissertation used a cross-sectional simultaneous two-equation ordered probit model to examine the relationship between health and overeducation for a group of European countries. Using a combination of the European Social Survey waves of 2002, 2004, 2006, 2008 and 2010, three models are estimated using the full-information maximum likelihood method. In order to get unbiased and causal results, at least one element of the health equation was not included in the overeducation equation. In the first model (Table 4), exclusion restrictions are introduced based on substantial theoretical arguments. Model 2 (Table A1) includes all explanatory variables in both equations. In the third model (Table A2), exclusion restrictions were re-introduced, but they are identified by selecting those regressors who are only significant for the health equation, i.e. not for the overeducation equation in model 2. Only physician density and its interaction effect with age are selected by both strategies.

Although the choice for certain exclusion restrictions clearly affects the magnitude of the coefficients of the health categories, our hypothesis that poor health is associated with a higher probability of being overeducated is confirmed by all models. We present three theoretical explanations to interpret this result. First, overeducation may be a compensation for the fact that unhealthy workers lack certain human capital elements, such as ability or on-the-job training. Second, unhealthy workers may be characterized by a reduced spatial flexibility. Hence, these workers depend highly on the size of the labour market and are thus more prone to overeducation than their nondisabled counterparts. Third, unhealthy workers may be the victim of discrimination. Employer discrimination reduces the probability of employment, so that unhealthy individuals may be more likely to accept employment which does not fully utilize their skills or qualifications. As we do not present any empirical proof for these theoretical considerations, it might be useful for future research to examine in greater detail to what extent they are true.

However, there might be a fourth explanation. The positive correlation coefficient between the error terms in model 1 and model 2 indicates that workers with a high probability of being overeducated also report a better health, and vice versa. Hence, we argue that working in a job for which individuals are overeducated may be used as a strategy to improve their health status. To what extent overeducation can be used as a stepping stone to improve health deserves further research.

However, the operationalisation method – the job analysis method – has two major limitations. First, there is not always consensus about the level of education required for a given occupation and these levels are subject to constant change. This affects the validity of the job analysis method, which may eventually result in a systematic over- or underestimation of the overeducation risk for certain occupations. As Verhaest & Omey (2010) state, analyses based both on the job analysis method and subjective assessment measures should be conducted. However, this was not possible considering the nature of the ESS data.

Second, individuals need to have job before they can possibly be labelled as overeducated. However, it is possible that individuals use health conditions to justify unemployment or non-participation. In other words, the relationship between the availability of employment opportunities and overeducation may not be straightforward, as Büchel & van Ham (2003) argue. Those in employment comprise a selective group, and analyzing overeducation while restricting the sample to the employed could lead to biased results. Therefore, we also introduce a Heckman probit model with sample selection, using the ESS data wave of 2010. Although our central hypothesis is again confirmed, there is one important difference compared to the previous models: the place of residence dummies are not significant anymore. This means that spatial determinants of overeducation are less important than we initially expected. Overall, as the majority of results are in line with the previous models, we argue that the sample selection bias seems to be limited.

In summary, the approach used in this dissertation extends the existing literature by introducing a broader framework: we postulate that health, and not just disability, affects the overeducation risk. No distinctions were made between mental and physical health. The simultaneous equation model estimated in this dissertation provided more insights into the relationship between health and overeducation than a single equation model could, and the estimation results are also more efficient due to better controlling for unobserved heterogeneity. Nevertheless, although the interpretation of the effects of the explanatory variables does not change, the choice for certain exclusion restrictions clearly affects the magnitude of some coefficients.

Appendix A. Simultaneous equation ordered probit models 2 and 3

Table A1 Model 2: ordered probit model without exclusion restrictions

	Overeducation equation		Health equation	
	Coefficient	Robust std. error[a]	Coefficient	Robust std. error[a]
Individual characteristics				
Bad or very bad health	0.720***	0.253		
Fair health	0.495***	0.171		
Good health	0.221***	0.091		
Low secondary education	–0.575***	0.015	–0.068***	0.012
Upper secondary education	–0.706***	0.011	–0.053***	0.009
Post-sec. non-tert. education	0.306***	0.023	–0.047**	0.019
Second cycle tert. education	1.102***	0.028	–0.036*	0.018
Age	–0.017***	0.005	–0.007	0.004
Age squared	0.0002***	0.0001	–0.0001***	0.0000
Foreign origin	0.145***	0.016	–0.027**	0.013
Work experience	0.011***	0.004	–0.030***	0.003
Work experience squared	–0.0004***	0.0001	–0.0001	0.0000
Intimate contact	–0.042**	0.020	0.195***	0.013
Farm or countryside	0.093***	0.020	0.036**	0.016
Village	0.073***	0.013	–0.034***	0.010
Town or small city	0.078***	0.013	–0.057***	0.010
Suburbs	0.032	0.016	–0.005	0.013
Never married	0.019	0.017	–0.032**	0.014
Widowed	0.063*	0.034	–0.041	0.026
Divorced	0.022	0.021	–0.056***	0.016
Separated	–0.028	0.038	–0.021	0.031
Partnership	0.051**	0.024	–0.070***	0.019
Female	–0.065***	0.018	–0.144***	0.013
Cohabitation	–0.058***	0.019	–0.021	0.016
Children	–0.025*	0.015	0.017	0.011
Female*cohabitation	0.042**	0.021	0.060***	0.017
Female*children	0.012	0.019	–0.004	0.015
ESS2002	–0.242***	0.016	–0.111***	0.012
ESS2004	–0.175***	0.015	–0.107***	0.011
ESS2006	–0.153***	0.015	–0.129***	0.011
ESS2008	–0.109***	0.014	–0.031***	0.011
Country characteristics				
% upper secondary educ.	1.034***	0.038	0.159***	0.028
% post-sec. non-tert. educ.	1.671***	0.316	–4.231***	0.145
% first cycle tertiary educ.	1.789***	0.054	0.236***	0.042
% second cycle tertiary educ.	1.578***	0.273	–2.445***	0.192
Student-teacher ratio prim.	0.012***	0.003	–0.023***	0.002
Student-teacher ratio sec.	–0.022***	0.003	0.020***	0.002
GDP per capita	0.006***	0.002	–0.0001	0.001
GDP per capita*age	–0.0002***	0.0001	0.0006***	0.0000
Physician density (per 1000)	0.025	0.023	0.219***	0.017
Physician density*age	–0.0001	0.0005	–0.003***	0.0004
Air pollution	–0.043***	0.006	–0.011**	0.005
Air pollution*age	0.0009***	0.0002	–0.0004***	0.0001
Constant	–0.865***	0.136		
Cut value_2_1			–2.411***	0.097
Cut value_2_2			–1.085***	0.097
Cut value_2_3			0.348***	0.097
***Log pseudolikelihood* = -170223.17**				
Correlation coefficient (ρ^b) =	0.165***	0.067		

Source: own calculations based on the working respondents in ESS2002, 2004, 2006, 2008 and 2010. N = 108 211. Note: * p < 0.10; ** p < 0.05; *** p < 0.01. [a] standard errors adjusted for clustering on country level. [b] Wald test for both equations: chi-square = 15 196.64; d*f* = 43; p = 0.000.

Table A2 Model 3: ordered probit model with exclusion restrictions based on model 2

	Overeducation equation		Health equation	
	Coefficient	Robust std. error[a]	Coefficient	Robust std. error[a]
Individual characteristics				
Bad or very bad health	0.344*	0.192		
Fair health	0.240*	0.129		
Good health	0.086	0.069		
Low secondary education	–0.575***	0.015	–0.067***	0.012
Upper secondary education	–0.708***	0.011	–0.052***	0.009
Post-sec. non-tert. education	0.313***	0.023	–0.047**	0.019
Second cycle tert. education	1.115***	0.025	–0.036**	0.018
Age	–0.016***	0.005	–0.007	0.004
Age squared	0.0002***	0.0001	–0.0002***	0.0000
Foreign origin	0.149***	0.016	–0.027**	0.013
Work experience	0.013***	0.003	–0.030***	0.003
Work experience squared	–0.0004***	0.0000	–0.0001	0.0000
Intimate contact	–0.059***	0.019	0.195***	0.013
Farm or countryside	0.086***	0.020	0.036**	0.016
Village	0.074***	0.013	–0.034***	0.010
Town or small city	0.081***	0.013	–0.057***	0.010
Suburbs	0.031**	0.016	–0.005	0.013
Never married			–0.033**	0.014
Widowed	0.050	0.032	–0.042	0.026
Divorced			–0.057***	0.016
Separated	–0.045	0.036	–0.022	0.031
Partnership	0.053**	0.023	–0.070***	0.019
Female	–0.053***	0.017	–0.144***	0.013
Cohabitation	–0.074***	0.016	–0.021	0.016
Children	–0.031**	0.015	0.017	0.011
Female*cohabitation	0.035*	0.021	0.060***	0.017
Female*children	0.014	0.019	–0.004	0.015
ESS2002	–0.234***	0.016	–0.111***	0.012
ESS2004	–0.166***	0.015	–0.107***	0.011
ESS2006	–0.143***	0.015	–0.129***	0.011
ESS2008	–0.105***	0.014	–0.032***	0.011
Country characteristics				
% upper secondary educ.	1.029***	0.038	0.160***	0.028
% post-sec. non-tert. educ.	2.094***	0.250	–4.225***	0.145
% first cycle tertiary educ.	1.786***	0.054	0.237***	0.042
% second cycle tertiary educ.	1.838***	0.245	–2.439***	0.192
Student-teacher ratio prim.	0.014***	0.002	–0.023***	0.002
Student-teacher ratio sec.	–0.026***	0.003	0.020***	0.002
GDP per capita	0.006***	0.002	–0.0009	0.001
GDP per capita*age	–0.0003***	0.0000	0.0006***	0.0000
Physician density (per 1000)			0.218***	0.017
Physician density*age			–0.003***	0.0004
Air pollution	–0.043***	0.006	–0.011**	0.005
Air pollution*age	0.0009***	0.0001	–0.0004***	0.0001
Constant	–0.713***	0.100		
Cut value_2_1			–2.414***	0.097
Cut value_2_2			–1.088***	0.097
Cut value_2_3			0.345***	0.097
***Log pseudolikelihood* = -170227.21**				
Correlation coefficient (ρ^b) =	0.065	0.051		

Source: own calculations based on the working respondents in ESS2002, 2004, 2006, 2008 and 2010. N = 108 211. Note: * p < 0.10; ** p < 0.05; *** p < 0.01. [a] standard errors adjusted for clustering on country level. [b] Wald test for both equations: chi-square = 14 723.46; d*f* = 39; p = 0.000.

Appendix B. Employment and overeducation

Table B1 Determinants of employment and overeducation: a bivariate probit model with selection

	Overeducation equation		Employment equation	
	Coefficient	Robust std. error[a]	Coefficient	Robust std. error[a]
Individual characteristics				
Bad or very bad health	0.391***	0.084	–0.901***	0.075
Fair health	0.173***	0.042	–0.221***	0.048
Good health	0.014	0.030	0.040	0.030
Low secondary education	–0.545***	0.161	–0.308***	0.053
Upper secondary education	–0.758***	0.123	–0.057**	0.029
Post-sec. non-tert. education	0.203*	0.111	0.010	0.057
Second cycle tert. education	0.820***	0.093	0.300***	0.039
Age	–0.051*	0.031	0.256***	0.016
Age squared	0.0007*	0.0004	–0.003***	0.0002
Foreign origin	0.191***	0.057	–0.021***	0.063
Work experience	–0.006	0.017	–0.011	0.010
Work experience squared	–0.0001	0.0002	–0.0003**	0.0001
Intimate contact	–0.104***	0.035	0.132***	0.030
Farm or countryside	0.098	0.075	0.027	0.068
Village	0.038	0.041	–0.042	0.037
Town or small city	0.027	0.033	0.011	0.023
Suburbs	0.045	0.046	–0.043	0.031
Never married	0.050	0.042	0.050	0.078
Widowed	0.071	0.093	0.089	0.088
Divorced	–0.014	0.048	0.226***	0.084
Separated	0.044	0.089	–0.025	0.152
Partnership	0.067	0.040	0.129**	0.054
Female	–0.068*	0.040	–0.021	0.030
Cohabitation	–0.195***	0.048	0.357***	0.076
Children	–0.008	0.029	0.157***	0.027
Female*cohabitation	0.191***	0.046	–0.362***	0.045
Female*children			–0.338***	0.043
Country characteristics				
% upper secondary educ.	1.000***	0.247	0.219	0.295
% post-sec. non-tert. educ.	2.646***	0.980	0.057	1.210
% first cycle tertiary educ.	1.702***	0.355	0.303	0.370
% second cycle tertiary educ.	2.022***	0.780	–1.101	1.097
Student-teacher ratio prim.	0.034***	0.013	–0.018	0.019
Student-teacher ratio sec.	–0.039*	0.022	0.003	0.026
GDP per capita	0.002	0.008	–0.007	0.009
GDP per capita*age	–0.0002	0.0001	0.0002	0.0001
Physician density (per 1000)	0.048	0.120	–0.039	0.062
Physician density*age	–0.0002	0.003	–0.0003	0.001
Air pollution	–0.044	0.036	–0.002	0.032
Air pollution*age	0.0005	0.0005	0.0003	0.0006
Constant	0.343	0.688	–4.353***	0.463
***Log pseudolikelihood* = -33463.53**				
Correlation coefficient (ρ^b) =	–0.435***	0.145		

Source: own calculations based on the respondents of ESS2010. N = 48 275.
Note: * p < 0.10; ** p < 0.05; *** p < 0.01.
[a] standard errors adjusted for clustering on country level.
[b] Wald test for independent equations: chi-square = 6.77; d*f* = 1; p = 0.009.

Appendix C. Single probit regression of overeducation

Table C1 Determinants of overeducation: a single probit regression

	Overeducation equation	
	Coefficient	Robust std. error[a]
Individual characteristics		
Bad or very bad health	0.175***	0.047
Fair health	0.104***	0.033
Good health	0.020	0.025
Low secondary education	–0.261**	0.118
Upper secondary education	–0.653***	0.085
Post-sec. non-tert. education	0.371***	0.111
Second cycle tert. education	0.885***	0.114
Age	–0.0236	0.020
Age squared	0.0002	0.000
Foreign origin	0.156***	0.040
Work experience	–0.021	0.016
Work experience squared	–0.0003***	0.0001
Intimate contact	–0.083***	0.028
Farm or countryside	0.055	0.052
Village	0.051	0.038
Town or small city	0.045*	0.026
Suburbs	0.047	0.033
Never married	0.078*	0.041
Widowed	0.063	0.063
Divorced	0.069	0.050
Separated	0.005	0.072
Partnership	0.096***	0.028
Female	–0.084*	0.033
Cohabitation	–0.072	0.048
Children	–0.042	0.032
Female*cohabitation	0.090**	0.041
Female*children	0.014	0.038
Country characteristics		
% upper secondary educ.	1.297***	0.222
% post-sec. non-tert. educ.	3.096***	0.994
% first cycle tertiary educ.	1.929***	0.304
% second cycle tertiary educ.	1.065	0.864
Student-teacher ratio prim.	0.032**	0.013
Student-teacher ratio sec.	–0.039*	0.020
GDP per capita	–0.006	0.006
GDP per capita*age	–0.0001	0.0001
Physician density (per 1000)	0.056	0.102
Physician density*age	–0.0001	0.002
Air pollution	–0.048	0.030
Air pollution*age	0.0005	0.0006
Constant	–1.400	0.561
***Log pseudolikelihood* = -22 986.648**		
***Pseudo R^2* = 0.1895**		

Source: own calculations based on the respondents of ESS2010. N = 48 275.
Note: * p < 0.10; ** p < 0.05; *** p < 0.01.
[a] standard errors adjusted for clustering on country level.

Bibliography

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 – New Series, 53,* 434-452.

Battu, H. & Sloane, P.J. (2004). Overeducation and ethnic minorities in Britain. *The Manchester School, 72,* 535-539.

Battu, H., Belfield, C., & Sloane, P.J. (2000). How well can we measure graduate overeducation and its effects? *National Institute Economic Review, 171,* 82-93.

Bauer, T.K. (2002). Educational mismatch and wages: a panel analysis. *Economics of Education Review, 21,* 221-229.

Becker, G. (1964). *Human Capital: a theoretical and empirical analysis with special reference to education.* New York: Columbia University Press.

Benitez-Silva, H., Buchinsky, M., Chan, H.M., Cheidvasser, S. & Rust, J. (2004). How large is the bias in self-reported disability? *Journal of Applied Econometrics, 19,* 649-670.

Blazquez, M. & Malo, M.A. (2005). Educational mismatch and labour mobility of people with disabilities: the Spanish case. *Revista de Economia Laboral, 2,* 31-55.

Blechinger, D. & Pfeiffer, F. (2000). "Technological change and skill obsolescence: the case of German apprenticeship training", in: Heijke, H. & Muysken, J. (2000). *Education and training in a knowledge-based economy.* London: MacMillan.

Borghans, L. & de Grip, A. (2000). *The overeducated worker? The economics of skill utilisation*. Cheltenham: Edward Elgar.

Brunner, E. & Maruyama, K. (2011). Health and sustainability: international ecological study of carbon dioxide emissions and life expectancy. *Journal of Epidemiology and Community Health, 65 (1),* 442-443.

Büchel, F. (2002). The effects of overeducation on productivity in Germany – the firms' viewpoint. *Economics of Education Review*, *21,* 263-275.

Büchel, F. & Battu, H. (2002). *The theory of differential overqualification: does it work?* Bonn: IZA Discussion Paper Series.

Büchel, F., de Grip, A., & Mertens, A. (2003). *Overeducation in Europe. Current issues in theory and policy.* Cheltenham: Edward Elgar.

Büchel, F. & van Ham, M. (2003). Overeducation, regional labour markets, and spatial flexibility. *Journal of Urban Economics, 53,* 482-493.

Budria, S. & Moro-Edigo, A.I. (2009). The overeducation phenomenon in Europe. *Revista Internacional de Sociologia, 67,* 329-345.

Budria, S. (2011). Are educational mismatches responsible for the 'inequality increasing effect' of education? *Social Indicators Research, 102 (3),* 409-437.

Cai, L. (2010). The relationship between health and labour force participation: evidence from a panel data simultaneous equation model. *Labour Economics, 17 (1),* 77-90.

Chevalier, A. (2003). Measuring over-education. *Economica, 70,* 509-531.

Cobb, S. (1976). Social support as a moderator of life stress. *Psychosomatic Medicine, 38,* 301-314.

Cooper, R. (2009). States with more physicians have better-quality health care. *Health Care, 28 (1),* p.91-102.

DeLeire, T. (2001). Changes in wage discrimination against people with disabilities: 1984-93. *The Journal of Human Resources, 36,* 144-158.

Dhondt, S. et al. (2012). Health impact assessment of air pollution using a dynamic exposure profile: implications for exposure and health impact estimates. *Environmental Impact Assessment Review, 36,* 42-51.

Dolton, P.J. & Silles, M.A. (2002). *The determinants of graduate over-education*. Oxford: Department of Economics.

Dolton, P.J. & Silles, M.A. (2008). The effects of overeducation on earnings in the graduate labour market. *Economics of Education Review, 27,* 125-139.

Dolton, P.J. & Vignoles, A. (2000). The incidence and effects of overeducation in the UK graduate labour market. *Economics of Education Review, 19,* 179-198.

Fleming, C.M. & Kler, P. (2008). I'm too clever for this job: a bivariate probit analysis on overeducation and job satisfaction in Australia. *Applied Economics, 40,* 1123-1138.

Fogel, R.W. (2004). *The escape from hunger and premature death.* Cambridge: Cambridge University Press.

Frank, R.H. (1978). Why women earn less: the theory and estimation of differential overqualification. *American Economic Review, 68 (3)*, 360-373.

Galuska, A.D., Will, J.C., Serdula, M.K., & Ford, E.S. (1999). Are health care professionals advising obese patients to lose weight? *JAMA, 282,* 1576-1578.

Green, F., McIntosh, S., & Vignoles, A. (1999). *Overeducation and skills: clarifying the concepts*. London: London School of Economics and Political Science.

Groot, W. & Maassen van den Brink, H. (2000). Overeducation in the labor market: a meta-analysis. *Economics of Education Review*, *19,* 149-158.

Halaby, C.N. (1994). Overeducation and skill mismatch. *Sociology of Education, 67,* 47-59.

Hartog, J. (2000). Over-education and earnings: where are we, where should we go? *Economics of Education Review, 19,* 131-147.

Hoffmann, E. (1999). *International statistical comparisons of occupational and social structures: problems, possibilities and the role of ISCO-88.* Geneva: International Labour Office.

ILO (International Labour Office). (2011). *International Standard Classification of Occupations (ISCO-88)*. Geneva: International Labour Office.

ILO. (n.d.). International Standard Classification of Occupations (ISCO-08) – Conceptual Framework. URL: <http://www.ilo.org/public/english/bureau/stat/isco/docs/annex1.doc>. (25/03/2012).

Jones, M.K. (2006). Is there employment discrimination against the disabled? *Economic Letters, 92,* 32-37.

Jones, M.K., Latreille, P.L., & Sloane, P.J. (2006). Disability, gender and the British labour market. *Oxford Economic Papers, 58,* 407-449.

Jones, M., Mavromaras, K., Sloane, P.J., & Wei, Z. (2011). *Disability and job mismatches in the Australian labour market*. Bonn: IZA Discussion Paper Series.

Jones, M. & Sloane, P.J. (2010). Disability and skill mismatch. *The Economic Record, 86,* 101-114.

Keshk, O. (2003). Simultaneous equations model: what are they and how are they estimated. URL: <http://polisci.osu.edu/prism/resources/statsmodels/done_method_lunch.pdf>. (05/07/2012).

Kucel, A. & Byrne, D. (2008). Are over-educated people insiders or outsiders? A case of job search methods and over-education in UK. URL: <http://www.esri.ie/UserFiles/publications/20080929121714/WP258.pdf>. (12/07/2012).

Layes, A., Asada, Y. Kephart, G. (2012). Whiners and deniers – What does self-rated health measures? *Social Science & Medicine, 75,* 1-9.

Lindeboom, M. & Kerkhofs, M. (2009). Health and work of the elderly: subjective health measures, reporting errors and endogeneity in the relationship between health and work. *Journal of Applied Econometrics, 24,* 1024-1046.

Madden, D. (2004). Labour market discrimination on the basis of health: an application to UK data. *Applied Economics, 36,* 421-442.

Mavromaras, K., Oguzoglu, U., Black, D., Wilkins, R. (2006). *Disability and employment in the Australian labour market*. URL: <http://melbourneinstitute.com/downloads/labour/3-06%20Final%20Report%20accepted%203%20sep%2007.pdf>. (08/06/2012).

McGoldrick, M. & Robst, J. (1996). Gender differences in overeducation. A test of the theory of differential overqualification*. American Economic Review, 86,* 280-284.

Mehta, A., Felipe, J., Quising, P., & Camingue, S. (2011). Overeducation in developing countries: how can we test for it, and what does it mean? *Economics of Education Review, 30,* 1334-1347.

OECD (Organisation for Economic Co-operation and Development). (2004). *Female labour force participation: past trends and main determinants in OECD countries.* Paris: OECD.

Robertson, D.S. (2006). Health effects of increase in concentration of carbon dioxide in the atmosphere. *Current Science, 80 (12),* 1607-1609.

Roodman, D. (2009). *Estimating fully observed recursive mixed-process models with cmp.* CDG Working Paper 168. Washington D.C..: Center for Global Development.

Ross, C.E. & Wu, C.L. (1995). The links between education and health. *American Sociological Review, 60,* 719-745.

Rundell, K.W. (2012). Effect of air pollution on athlete health and performance. *British Journal of Sports Medicine, 46 (6),* 407-412.

Sajaia, Z. (n.d.). Maximum likelihood estimation of a bivariate ordered probit model: implementation and Monte Carlo simulations. URL: <http://www.adeptanalytics.org/download/ado/bioprobit/bioprobit.pdf>. (28/05/2012).

Sarma, S. & Peddigrew, C. The relationship between family physician density and health related outcomes: the Canadian evidence. *Cahiers de Sociologie et de Demographie Medicales, 48 (1)*, p.61-105.

Schneider, S. (2007). *Measuring educational attainment in cross-national surveys: the case of the European Social Survey.* Oxford: Nuffield College.

Sicherman, N. (1991). Overeducation in the labor market. *Journal of Labor Economics, 9,* 101-122.

Sicherman, N. & Galor, O. (1990). A theory of career mobility. *Journal of Political Economy, 98*, 169-192.

Sloane, P. J. (2003). "Much ado about nothing? What does overeducation literature really tells us?", in: Büchel, F., de Grip, A., & Mertens, A. (2003). *Overeducation in Europe. Current issues in theory and policy.* Cheltenham: Edward Elgar.

Smith, J.P. (1999). Healthy bodies and thick wallets: the dual relation between health and economic status. *Journal of Economic Perspectives, 13,* 145-166.

Stern, S. (1996). Semiparametric estimates of the supply and demand effects of disability on labour force participation. *Journal of Econometrics, 71,* 49-70.

Subramanian, M. & Canning, D. (2009). Short and long-term relationship between physician density on infant mortality: a longitudinal econometric analysis. Boston: Harvard School of Public Health.

Swift, R. (2011). The relationship between health and GDP in OECD countries in the very long run. *Health Economics, 20,* 306-322.

The World Bank. (2011). *CO2 emissions (metric tons per capita)*. URL: <http://data.worldbank.org/indicator/EN.ATM.CO2E.PC>. (25/05/2012).

UNESCO (United Nations Educational Scientific and Cultural Organization). (2006). *ISCED97. International Standard Classification of Education 1997*. Paris: UNESCO.

UNESCO. (2010). *Beyond 20/20*. URL: <http://stats.uis.unesco.org/unesco/TableViewer/TableView.aspx?ReportId=3336&IF_Language=eng>. (28/06/2012).

UN (United Nations). (2011). *International Human Development Indicators*. URL: <http://hdrstats.undp.org/en/countries/profiles/BEL.html>. (29/04/2012).

Verbeek, M. (2008). *A guide to modern econometrics*. Chichester: John Wiley & Sons.

Verhaest, D. & Omey, E. (2010). The determinants of overeducation: different measures, different outcomes? *International Journal of Manpower, 31,* 608-625.

Verhaest, D. & Omey, E. (2012). Overeducation, undereducation and earnings: further evidence on the importance of ability and measurement error bias. *Journal of Labor Research, 33 (1),* 76-90.

Verhaest, D. & van der Velden, R. (2012). Cross-country differences in graduate overeducation and its persistence. Forthcoming in 'The European Sociological Review'.

White, A.M., Philogene, G.S., Fine, L., & Sinha, S. (2009). Social support and self-reported health status of older adults in the United States. *American Journal of Public Health, 99,* 1782-1878.

WHO (World Health Organisation). (2011a). *About WHO.* URL: <https://apps.who.int/aboutwho/en /definition.html>. (14/07/2012).

WHO. (2011b). *Global Health Observatory Data Repository – Physician density per 1000.* URL: <http://apps.who.int/ghodata/?vid=92100>. (08/07/2012).

Zwinkels, W. (2001). *The employment situation of people with disabilities in the European Union.* EIM: Business and Policy Research.