Assessing the incidence and consequences of overeducation in the Netherlands

Olivier, Rémi

Assessing the incidence and consequences of overeducation in the Netherlands

Universiteit Leiden

Master thesis, June 2021

Rémi Olivier s1552430

Public Administration: Economics and Governance

Supervisor: Prof.dr. P.W.C. (Pierre) Koning

Second reader: Prof.dr. M.G. (Marike) Knoef

Abstract

This study uses data from the Labour Supply Panel ('Arbeidsaanbodpanel') to analyse the incidence of overeducation in the Netherlands and the returns to overeducation compared to non-overeducated workers in the same job level, as well as the costs to overeducation compared to non-educated workers in the same education level.

Our exploratory analysis combines descriptive statistics and linear probability models to illustrate the incidence of overeducation and to assess whether recent job polarization trends have led overeducation to be more prevalent among middle skilled workers, and whether economic downturns lead to an increase of overeducated workers. We find no evidence for such trends.

Our regression models on the effects of overeducation on earned income are consistent with previous studies in that there are returns to overeducation compared to non-overeducated workers in the same job level. However, we find large disparities in returns over different education levels and job sectors, with many showing no significant returns whatsoever. Costs to overeducation compared to non-overeducated workers of the same education level are more pervasive, and are especially high among those groups that show no significant returns compared to non-overeducated workers within the same job level. Furthermore, our fixed-effects models illustrate that any such returns might be the result of unobserved individual characteristics, rather than of an individual's formal education level.

Table of contents

1. Introduction 4

2. Theory: effectiveness of increasing education 6

2.1. Returns to education and overeducation 6

2.2. SBTC and Job Polarization: the role of education with changing labour demands 8

2.3. Crowding out effects 10

3. Research Design 12

3.1. Data source: Labour Supply Panel 12

3.2. Regression models 13

3.3. Robustness checks 16

3.4. Main variables 17

4. Exploratory analysis over job complexity, education levels and overeducation in the Netherlands 20

4.1. Descriptive statistics 20

4.2. Linear probability model: factors increasing chances of being overeducated 23

5. Results 28

5.1. Returns to education 28

5.2. Costs and returns to overeducation and undereducation 29

5.3. Returns to overeducation across job sectors 34

5.4. Returns to overeducation: changes over time 36

5.5.Robustness checks 38

6. Conclusion 41

References 44

1. Introduction

The extent to which investments in human capital through job training are an effective tool within Active Labour Market Policies is a highly researched topic (Psacharopoulos & Patrinos, 2018). The underlying theory behind investing in education and job training is that it would facilitate overcoming changes in demand in changing labour markets: as demands diminish in one sector or skill level and increase in others, education and job training can ease the transition of labour supply from the one to the other and, consequently, raise the overall productivity of the labour force. For many decades, these changes in demand have also been mirrored by Skill Biased Technological Change (SBTC) (Katz et al., 1999). Following the logic of the SBTC theory, it was expected that low-skilled job demands were to diminish, leading to an oversaturation of low skilled labour supply.

New insights into the mechanisms of SBTC however have put into the question the traditional hypothesis that low-skilled job demands are to reduce as a result of technological change, and have instead argued that mainly jobs in the middle segment, with many routine-type occupations, will diminish – a phenomenon known as 'job polarization' (Acemoglu & Autor, 2011). Job polarization may lead to an oversupply of middle educated workers, which would make them have to accept low skilled jobs. Accordingly, educating low skilled workers may prove an ineffective tool for increasing their chances in the labour market, as they would be increasingly competing with middle skilled workers for low skilled positions. Educating low skilled workers would only increase this competition for newly entering low skilled workers, worsening their chances on the low skilled labour market. This adds a new perspective on the question of crowding out effects and overeducation, and could lead to decreasing returns to education and overeducation. Studies on crowding out effects in low complexity occupations have shown middle skilled workers to compete with low skilled workers for low skilled positions, especially during economic downturns (Gautier et al, 1997; Polmann-Schult, 2005). These studies are however not directly linked to job polarization.

Despite increasing evidence in favour of job polarization effects in developed countries (Salvatori, & Manfredi, 2018), no recent in-depth analysis exists of the current overeducation situation in the Netherlands. The current study aims to fill in this gap. The goal of this study is to analyse returns to overeducation in the Netherlands. We first lay the groundwork for this in-depth analysis of overeducation in the Netherlands by developing an exploratory analysis of the factors influencing probability of overeducation, using data from the Dutch Labor

Supply Panel. In particular, we will investigate variation in the incidence of overeducation across sectors and over time, so as to uncover possible trends or business cycle effects. These analyses can also provide us important insights into the influence of either SBTC or job polarization effects in the Dutch labour market. For the main part of this study, we estimate the returns to overeducation in the Netherlands. At one extreme, one can think of the absence of any returns, with overeducated workers receiving similar wages as workers that are not overeducated. At the other extreme, overeducated workers may receive similar wage premiums as workers with similar education levels but higher job levels. By assessing where workers are positioned between these two extremes, we can evaluate the potential downside effects of overeducation.

For our analysis, we will use the bi-yearly 'Arbeidsaanbodpanel' or Labour Supply Panel of the years 2010-2014, during which the Netherlands experienced high rises in unemployment. This panel offers a representative sample of workers in the Netherlands and describes in detail their current, as well as their earlier, job situation and education. Datasets from this panel have been used before to analyse effects of education on wage in a study by Nelissen (2007). Whereas he used the sets to analyse to what degree the observed returns to education are the direct result of the education, and to what degree these are the result of personal traits, our study focuses specifically on those workers who are overeducated, to see what their returns to education are compared to workers in appropriate job levels.

The research question of this study is:

"What is the incidence of overeducation, how has it developed and what are the effects of overeducation on workers' wages in the Netherlands?"

This study is divided in three main parts. First, we review current literature on the returns to overeducation, as well as on SBTC and job polarization in developed countries. Based on this review, we will be able to formulate expectations on the effects of job polarization on crowding out trends and returns to (over-)education. Second, we develop an exploratory analysis model, containing descriptive statistics on the distribution of over- and undereducated workers in the Netherlands, as well as a linear probability regression model illustrating how certain characteristics of workers and the market affect chances of being overeducated. In the third and main part of this study, we analyse returns to overeducation by comparing the estimates of the costs of overeducation compared to non-overeducated

workers in the same education level, and the returns of overeducation compared to non-overeducated workers in the same job level.

2. Theory: effectiveness of increasing education

There are many welfare gains from education that we can think of, such as increased higher wages, increased tax revenues, self-actualisation, less crime and health benefits. For the purpose of this study however, we limit ourselves to types of welfare gains which have a very direct link to the labour market and overall productivity. First and foremost, studies in the past have indicated that increased education leads to increased wages, even at lower job complexities (Duncan & Hoffman, 1981; Psacharopoulos & Patrinos, 2018). Secondly, education and training can prepare workers for changing skill demands of the labour market, mainly as a result of technological change.

What follows is a review of existent literature on the returns to overeducation and of recent developments on the labour market that could have an effect on returns to education and overeducation. We start by looking at past studies on the returns to overeducation. Secondly, we will review studies on SBTC and job polarization, and formulate what effects we would expect from the latter on crowding out and returns to overeducation.

2.1. Returns to education and overeducation

While the link between education and earnings became the focus of empirical research in the early sixties of the 20th century (Psacharopoulos & Patrinos, 2018), Duncan and Hoffman (1981) were the first to study the effects of imbalances between a worker's education level and acquired job level. This started an extensive literature on the relative returns and costs to over-and undereducation (Rumberger, 1987; Sicherman 1991; Verdugo & Verdugo, 1989; Van der Meer, 2006) with a consensus arising over the decades that there are indeed returns to education, even if an individual works in a job below his education level. However, these returns are far lower than the returns to required education.

Hartog's (2000) macro-analysis on a number of these studies has found the following overlapping outcomes:

1. Returns to required education for a certain job are higher than returns to actual education.
2. Returns to overeducation are 50%-75% smaller than the returns to required education.
3. Returns to undereducation are negative, but in most cases smaller than the positive returns to overeducation.

Hartog (2000) evaluates the three main explanations for the observation that overeducation yields positive returns. The first two explanations, which are referred to as "Searching and Matching" and "Human Capital", focus on the supply side of labour and argue that higher wages for overeducated workers are the result of the underlying motivations of the worker for choosing to work in a lower-level occupation. The Searching and Matching explanation hypothesizes that this mismatch is caused by workers' (lack of) experience and imperfect information on the job market, whereas the Human Capital explanation hypothesizes that accepting lower-level occupations is a conscious investment in later upward occupational mobility. These two explanations are not mutually exclusive, and could both play a role in the distribution of wages and overeducated workers (Hartog, 2000).

The third explanation, "Assignment", focuses not on the choice of the worker, but on the demand side of the job market, which is why this explanation is more interesting for our study of overeducation in light of changing skill demands. In this respect, Teulings (1995) formulates a general equilibrium model based on this explanation, where both workers and jobs are ranked by quality (skill and complexity respectively), after which the highest quality worker is assigned to the highest quality job, and the assignment simply goes down the list, with the wages (assumed to be fixed per job) following along. From this model, we can understand effects of imbalances between skills supply and complexity demand on wages. In particular, some workers end up on positions too low for their skill level, but will on average still earn more than the lower educated workers who end up in lower ranked occupations. In other words, there would still be returns to attained education. However, if occupations of a certain complexity level were to disappear while education distributions remain the same, we can expect these returns to decrease, as larger parts of lower complexity levels will be occupied by overeducated workers, decreasing their average wage advantage compared to lower skilled workers.

Teulings' model is certainly most explicit in formulating this expectation. But even if we did not use this model as the foundational theory on returns to overeducation, Hartog (2000)

indicates that even the simplest forms of supply-and-demand models would still predict returns to (over-)education to fall if supply of skills rises faster than the demand. This expectation however is not confirmed by his own data-analysis, which shows an increase in the overall returns to overeducation in European countries.

More recent studies have also found positive returns to overeducation. In a recent comparative study, Johnes (2018) finds no evidence for smaller returns to overeducation among lower wage levels, and only signals significant lower returns for higher educated workers in Italy The interesting outlier of his study finds is the case of the United Kingdom, in which no returns to overeducation are found.

The outcomes of aforementioned studies can be compared to those obtained from this study. As our data are more recent and there have been large developments in the labour market of western countries since Hartog's study, we can compare our results to those of 2000, and see if these recent trends have changed the returns we can expect from education if a worker is overeducated in the Netherlands.

2.2. SBTC and Job Polarization: the role of education with changing labour demands

Although the effects of technological change have been a subject of interest for labour economists since the early days of its existence, it was in the 1980s that Skill Biased Technological Change (SBTC) began to dominate as the main explanation for the increasing inequality in western countries (Card & DiNardo, 2002; Katz & Murphy, 1992). It was hypothesized that technological advances had increased demand for high skilled workers and decreased demand for low skilled workers, whose jobs could to a large extent be automated. As long as these skill demands were to increase faster than the upgrading of workers' skills, which is referred to as the 'Tinbergen (1975) race', this decline in demand on the one side of the skills spectrum and rise in demand on the other would explain the relative decreasing and increasing wages and the observed increased inequality in the United States (Katz & Murphy, 1992). This rise in inequality was not limited to the United States, as in the OECD countries (many of which have more generous minimum wages), unemployment among low skilled workers rose sharply between 1979 and 1992 (Berman et al, 1997).

From this understanding of the effects of SBTC, we can see that by educating low skilled workers governments can increase their average productivity and thereby increase their chances on the labour market. If successful, raising the educational level of low skilled workers would make skill supply keep up with increasing skill demands caused by technological change. This would, however, only hold if the labour market is characterized by decreased demands for low skilled, and increased demands for middle and high skilled workers (Manning, 2004).

New insights on the effects of technological change have put SBTC into question. Autor et al. (2015) demonstrated that not the skill level of the occupation decides whether it will be affected by technological change, but rather the degree to which it is a routine-type job. They demonstrate that manual-type low skilled occupations have remained fairly unaffected by technological change: we can indeed imagine that a restaurant or cleaning worker would not be severely affected by recent technological advancements. On the other hand, clerical office work such as accounting and basic analytics is largely substituted by automated systems. In other words, it appears that not the skill level decides to what degree workers are affected by technological change, but the degree to which their work is either complemented, substituted, or unaffected by new technologies (Autor et al, 2015). It has been argued that routine clerical work, which is mainly represented in the middle skill segment, is the main victim to current technological changes as it is most prone to being substituted by new technologies, whereas high complexity abstract jobs and low complexity manual jobs are either complemented by technological change or left unaffected (Manning, 2004).

These and other findings have led to a switch from the traditional SBTC theory to the "Routine-Biased Technological Change" (RBTC) theory, which predicts that mainly routine-type jobs are to be replaced by automation. As a result, this new model predicts that instead of technological advancements leading to lower demands for low skilled occupations, they will lead to a form of 'job polarization', increasing demand for both low skilled and high skilled labour but decreasing demand for labour of the middle segment (Autor & Dorn, 2013). Many empirical studies have found evidence for these job polarization trends in developed countries (e.g. Autor & Dorn, 2013; Salvatori & Manfredi, 2018) and it has become widely accepted that this trend is to shape the developments of the labour market of developed countries in the years to come (Salvatori & Manfredi, 2018; Borstlap, 2020).

As from job polarization we expect increasing demands for low skilled work and decreasing demands for middle skilled work, one could hypothesize that crowding out effects will arise within lower skilled positions, as more middle skilled workers will need to compete for these positions with lower skilled workers. This could in turn lead to increased unemployment among low skilled workers and higher overeducation levels among middle skilled workers, which as indicated by Hartog (2000) has far lower returns on wages than job required education. We would then expect decreasing overall returns to education, especially at intermediate education levels, but also at lower education levels as a result of crowding out effects. At higher education levels, we would expect returns to education to remain stable, as both SBTC and RBTC theories hypothesize increasing demands for high educated workers.

2.3. Crowding out effects

Thurow's (1975) job competition model hypothesises that in many cases workers do not compete with each other based on what wages they would accept, but on their relative costs of training for a certain position. The theory follows from the observation that cognitive positions are generally filled by workers who have been trained within the labour market, as opposed to those who enter it fresh. Based on that observation, Thurow states that "Thus, the labor market is not primarily a bidding market for selling existing skills but a training market where training slots must be allocated to different workers" (1975, p. 76). It is well known that in order to assess costs of training an employee for a position, employers look at the education level of the worker as an indication of the worker's ability to perform the tasks required by the position (Spence, 1973), and as we have seen in the literature analysis on returns to overeducation, there is evidence that higher educated workers on lower positions earn higher wages (Duncan & Hoffman, 1981; Hartog, 2000; Johnes, 2018).

Based on the job competition model, crowding out effects have been hypothesized to affect low educated workers, as higher skilled workers start to compete with them for positions below their skill level – and especially during economic downturns. If crowding out effects indeed play a role in explaining low skilled unemployment, educating these low skilled workers could still increase overall productivity (Duncan & Hoffman, 1981), but it would be far less effective in reducing low skilled vulnerability on the labour market, as "(…) the labor-market position of people receiving this schooling improves, but to the detriment of those who already have this additional education" (Teulings & Koopmanschap, 1989).

The initial argument behind the hypothesis that middle skilled workers were crowding out lower skilled workers was that as all developed countries steadily increased supply of skilled labour, the labour market has not been able to keep up with upgrading job skills (Pollmann-Schult, 2005). More recent studies have focused more specifically on cyclical crowding out effects, whereby middle skilled workers seek refuge in lower skilled positions during times of recession, crowding out lower skilled workers. Gautier et al. (1997) found evidence for crowding out effects in Netherlands between 1993 and 1995 at the worker inflow in lower job levels by estimating multinomial logit models for the distribution of workers starting in a new job over job levels, for each level of education, and from these estimates calculating job-level probabilities. They found that in a recession year, there is a higher probability for higher skilled workers to start in lower skilled positions. Pollmann-Schult (2005) found similar evidence for crowding out effects at the worker inflow in West Germany between 1984-2000 by analysing the transition of low skilled workers from unemployment: this study found that during cyclical downturns, re-employment chances deteriorated for low skilled workers more so than for skilled workers.

If the traditional hypothesis of SBTC were to hold and the demand for low skilled positions decreases over time, then one can indeed expect any crowding out effects to be mainly cyclical in nature and create no systematic imbalance in developed countries: educating more low skilled workers would move them away from those areas in the labour market that are expected to decrease. However, based on these new insights into the developments and polarization in the labour market, we could hypothesize crowding out lower skilled workers by educating people towards oversupplied labour areas instead of away from it to be likely to arise now more than ever. The prospect of job polarization destabilizing the labour market has of course not gone unnoticed by national and international governments (see for example Borstlap, 2020). It creates a complicated challenge for future governments as there is no one-size-fits-all approach to counter the imbalance as there was in the traditional SBTC theory. The success of policies aiming at increasing worker skills largely depends on the degree to which these efforts manage to mobilize workers from oversupplied to undersupplied areas in the labour market. In other words, its success depends on the degree to which it eases frictions in the labour market as a result of changing demands. Contrasting to this, crowding out effects combined with lack of increases in productivity would clearly indicate that, at least in some areas, investing in increasing workers' education would be highly ineffective (Teulings & Koopmanschap, 1989).

3. Research Design

As we are interested in seeing what effects recent labour market trends have had on the returns to overeducation in the Netherlands, we will need to gain detailed information on which characteristics of the worker and the labour market influence workers' chances of working in overeducated positions, as well as what the returns to education and overeducation are in specific job levels and education levels.

We will lay the groundwork for answering the research question of this study by characterizing overeducated workers in the Netherlands and seeing what trends we can observe over the period of our samples. Our first goal is to analyse descriptive statistics such as the distribution of overeducated workers over time, job levels and sectors in the Dutch labour market. To acquire more detailed information on what effects certain attributes of workers and the market have on the chances of the individual to work in a position below his education level, we specify a linear probability model. Based on this model, we will be able to get relevant insights into current trends of overeducation in the Netherlands.

The second part of the study focuses on the wage returns to education and overeducation, for which we will use various regressions. In these regressions, we consider the average costs to overeducation, as compared to workers with the same education level but in an appropriate job level, and the benefits to overeducation as compared to workers in the same job level but with an appropriate education level. We will compare both values as an interpretation of the relative gains to costs. This analysis will then be expanded to models containing interactions between overeducation and job levels, education levels, sectors and years, to see in which of these categories gains to costs are relatively favourable and in which categories they are not.

3.1. Data source: Labour Supply Panel

The datasets used for this research are sets from the 'Arbeidsaanbodpanel' (Labour Supply Panel), a biyearly time series panel originally organized by the OSA Institute for Labour Studies foundation, and after 2010 organized by the government research agency SCP (Social and Cultural Planning bureau) that collects data on households of the Dutch working aged population. The panel collects data on a sample of around 4500 participants every two years since 1986. These data were collected through face-to-face interviews from 1985-2002,

through surveys either online or on paper from 2004-2014, and solely through online surveys from 2016 onwards (SCP, 2019).

Data collection of the Labour Supply Panel takes place in two phases: during the first phase, the selected households are contacted by telephone for a screening interview, during which the bureau assesses which members of the household agree to participate. In the second phase, the surveys are sent to the participants, either online or by letter (SCP, 2019). In replacing participants who left the panel, the bureau makes sure that this sample gives as accurate a representation of the Dutch population as possible.

As measuring instruments undergo many changes throughout the history of this panel (e.g. as a result of updated education or occupation classifications), in our analysis three sets will be used that are highly similar in their way of measuring the variables that are of importance to this study: datasets 2010, 2012 and 2014.

The SCP reports a panel-to-panel attrition of about 20 to 30 per cent. Our analysis shows that within our set, 56,8% participated in all three years.

Table 1: Participation within the sets 2010-2014

	Year			
Panel participation	2010	2012	2014	Total
Once	26.38	10.94	22.66	20.02
Twice	16.69	31.71	21.28	23.20
Three times	56.94	57.35	56.06	56.78

3.2. Regression models

In our exploratory analysis of overeducation in the Netherlands, we investigate how overeducation is distributed in the Dutch labour market and how it is evolving over time. Furthermore, we want to understand what attributes of workers and the market make it more likely for someone to be in a position that is lower than the worker's education level. For this analysis, we will use the following linear probability regression functions that we estimate with OLS:

Model 1:

$$\Pr(Overeducated_{it} = 1) = \alpha + Job\ level_{it} + \gamma\, X_{it} + e_{it}$$

Model 2:

$$\Pr(Overeducated_{it} = 1) = \alpha + Education\ level_{it} + \gamma\ X_{it} + e_{it}$$

As our overeducation variables are created from the variables education level and job level, controlling for both in one model creates multicollinearity issues. Therefore the functions containing overeducation as an explanatory variable will be represented in 2 models, one controlling for education level and one for job level. In these models, matrix X with control variables includes year effects, age effects (linear and exponential), gender effects, effects of weekly hours worked, economic growth and job sectors.

Both models estimate whether each of the explanatory variables increases or decreases the chances of a worker to be overeducated. The first model will give us a clear indication in which job level overeducation is most prevalent. The second model adds the variable education levels, which will illustrate what effects the different education levels have on the workers' chances of being overeducated.

In the main part of this study, our interest lies on the effect of overeducation on earned wages. For this, we use two estimation strategies. We first estimate a general model on the returns to education and job levels, where the ranking of education and job levels and their interactions are used as explanatories. For this model, we use two regressions:

$$Wage_{it} = \alpha + \beta\ Education\ level_{it} + \delta\ Job\ level_{it} + \eta\ Education\ level_{it} * Job\ level_{it} + \gamma\ X_{it} + e_{it}$$

$$Log(Wage_{it}) = \alpha + \beta\ Education\ level_{it} + \delta\ Job\ level_{it} + \eta\ Education\ level_{it} * Job\ level_{it} + \gamma\ X_{it} + e_{it}$$

Where the first function estimates absolute returns to job level, education level and the interaction term and the second model estimates relative returns. In these two regressions, job level and education level are expressed as continuous variables. We are particularly interested in the coefficient η, which is positive (negative) if returns to education are higher (lower) in more complex jobs. The other estimation strategy for the wage regression focuses on the

overeducation dummy. We then estimate average returns (compared to non-overeducated workers in the same job level) and costs (compared to non-overeducated workers in the same education level) to overeducation. For these estimates we use the following basic linear regression functions:

$$Wage_{it} = \alpha + Overeducated_{it} + Job\ level_{it} + \gamma\ X_{it} + e_{it}$$

$$Wage_{it} = \alpha + Overeducated_{it} + Education\ level_{it} + \gamma\ X_{it} + e_{it}$$

$$Wage_{it} = \alpha + Undereducated_{it} + Job\ level_{it} + \gamma\ X_{it} + e_{it}$$

$$Wage_{it} = \alpha + Undereducated_{it} + Education\ level_{it} + \gamma\ X_{it} + e_{it}$$

The first function estimates average returns to overeducation given a certain job level, whereas the second estimates average costs to overeducation given a certain education level. We will interpret both models as indicating the border values in the range between average costs and benefits to overeducation. The third and fourth functions are the same as the first and second, except in these functions we estimate the effects of undereducation on wages. While undereducation effects are not the main focus of this study, including these two models complements our understanding of the effects of formal education on earned income, as well as allowing for additional comparisons of our data with those of Hartog (2000), from which we would expect costs to undereducation to be smaller than the returns to overeducation.

We will allow for different interaction terms with overeducation in order to see what the returns and costs are to overeducation in different education levels, job levels, sectors and over our panel years.

3.3. Robustness checks

Age effects

As a means of robustness check for our estimates on the effects of overeducation on wages, we will run our main wage-related regression model without controlling for age effects. We do this as we suspect that at higher ages, workers are more likely to have benefitted from their job experience rather than their education level, while at lower ages workers are both more likely to earn lower wages and more likely to be overeducated.

Potential omitted variable bias

While our control variables are designed to control for the most obvious confounders in the relation between overeducation and wages, it is important for the correct interpretation of this study to also discuss potential unobservable distorting effects. That is because of the limitations to the assumption in this study that the wage earned by a worker is established by the labour market, based on an evaluation of either demands and supply (wage competition model) or costs to train the worker (job competition model). Personal and emotional characteristics of the worker, such as personal job aspirations and leisure preferences, might very well play a role in the earned wages as well. If these characteristics also affect chances of accepting positions below one's education level, this could distort the outcomes shown in our regressions. In fact, Nelissen's (2007) study of the effects of personal characteristics on wages illustrates that these characteristics may explain differences in wages better than formal education. To control for these possible distortions by unobserved individual characteristics, we will run a fixed effects regression model over the three panel years as a mean of robustness check of our results.

3.4. Main variables

Level of education

The SCP uses the Dutch soi-2006 classification to measure the datasets of 2010-2012. Statistics Netherlands (CBS) defines 7 main categories and several subcategories:

SOI-1 Education preceding the first level (usually 2 years)
SOI-2 Education at the first level (6 years)
SOI-3 Education at the second level, first stage (1-4 years)
3.1 Programmes not giving entry to level 4 3.2 Programmes giving entry to level 4.1 3.3 Programmes giving entry to levels 4.2 and 4.3
SOI-4 Education at the second level, second stage
4.1 Short programmes (up to 1 year after completion of level 3.3) 4.2 Intermediate programmes (between 1 and 3 years, full-time equivalents, cumulative at level 4) 4.3 Long programmes (3-5 years)
SOI-5 Education at the third level, first stage:
5.1 Shorter higher professional programmes (1-3 years) accessible to those having completed at least level 4.2 general or 4.3 vocational programmes (e.g. those leading to an associate's degree) 6 5.2 Higher professional programmes (4 years) accessible to those having completed at least level 4.2 general or 4.3 vocational programmes leading to a bachelor's degree 5.3 University programmes leading to a bachelor's degree
SOI-6 Education at the third level, second stage; mainly programmes leading to a master's degree or equivalent
SOI-7 Education at the third level, third stage; doctorate and other post-level 6 programmes (1-3 years).

(Statistics Netherlands, 2010)

The Labour Supply Panel has collected data on the highest obtained education degree by the participants. The first SOI-level is not represented in our dataset as none of the participants had this level as their highest obtained degree.

For all of our models but the linear probability model, it is preferable to simplify these classes, as otherwise some of the subcategories contain too few observations. We therefore align the education levels with the five occupation levels deemed appropriate for them, creating a new education variable of five classes. This simplified variable can also more easily be used as a continuous variable, which we include in our first regression on returns to education.

Level of occupation

The Labour Supply Panel uses the SBC-2010 classification to measure the distribution in the datasets 2010-2014. This classification distinguishes 5 job levels, which are assigned based on the required education level for the position. The first digit of this classification indicates the job level. 5 classes are distinguished:

1. Elementary job level: Jobs for which only basic education is required. The CBS assumes that Elementary job levels are most appropriate for the education levels 2-3.1. Although level 3.1 is part of the next main education level, this level of education is also seen as appropriate for an elementary job. This is because these types of education are usually low-level courses followed as a hobby (CBS, 2010).
2. Lower job level: Jobs for which a lower vocational education (of the Dutch mbo-1 or mbo-2 level) is required. The CBS assumes that SOI-levels 3.2, 3.3 and 4.1 are most appropriate for this job level.
3. Intermediate job level: Jobs for which a higher high school or vocational education (of the Dutch mbo-4, HAVO or VWO level) is required. The CBS assumes that SOI-levels 4.2 and 4.3 are most appropriate for this job level.
4. Higher job level: Jobs for which a higher professional education (equivalent to a Dutch HBO or university bachelor) is required. The CBS assumes that SOI-levels 5.1-5.3 are most appropriate for this job level.
5. Scientific job level: Jobs for which an academic master or doctorate degree is required. With this job level, the dutch measurement structure differs from the International Standard Classification of Occupations (ISCO), as higher job levels are assumed to be the highest level. The CBS however argues that in the Netherlands, there is a clear distinction that can be made between higher level jobs and scientific jobs (which would require an academic master). SOI-levels 6 and 7 considered appropriate for this job level.

Overeducation and undereducation

Hartog (2000) identifies three main ways of defining overeducation:

1. Job analysis: a systematic evaluation by professional job analists, an example of which being the Dutch CBS formulating appropriate education levels per job levels.
2. Worker self-assessment: an evaluation based on input from workers themselves. There are multiple concerns of biases, as respondents might overstate the

requirements of their occupations or simply reproduce the vacancy requirements (Hartog, 2000). Furthermore our panel does not offer precise enough data on workers' evaluation of their job complexity and we therefore will leave this method out of our analysis.

3. Realized matches: assessing appropriate education levels by estimating average education levels per job level. Although this method has the advantage of analysing the actual (realized) distribution of skilled workers over different job levels, it cannot be used for analyses of actual demand shifts in the labour market (Hartog, 2000). As the main goal of this research is to analyse returns to overeducation given job and education level, using a variable based on a regression of these two variables is not desirable.

In light of the above arguments, we will only use the first method of analysing overeducation. We consider a worker overeducated if his education level is higher than the appropriate education levels for his job level, and we consider a worker undereducated if his education level is lower than the appropriate education levels for his job level.

Although this method is also signalled by Hartog (2000) as the preferable method, there are some limitations. The main limitation is that the Statistics Netherlands uses only five classes of job levels, which leaves us with quite an unprecise definition of overeducation. Furthermore, this method forces us to operationalize overeducation and undereducation as a binary variable, meaning that effects from the degree of over- or undereducation cannot be represented in our model.

Economic growth

As illustrated by Gautier et al (1997) and Pollman-Schult (2005), there is evidence for an increase in overeducated workers during economic downturns. To see if differences over the panel years can be explained by differences the business cycle, we create the dummy variable *Economic Growth,* which in years of growth takes the value 1 and in years of decline takes the value 0. This effectively means that the years 2010 and 2014 take the value 1 and the year 2012 takes the value 0.

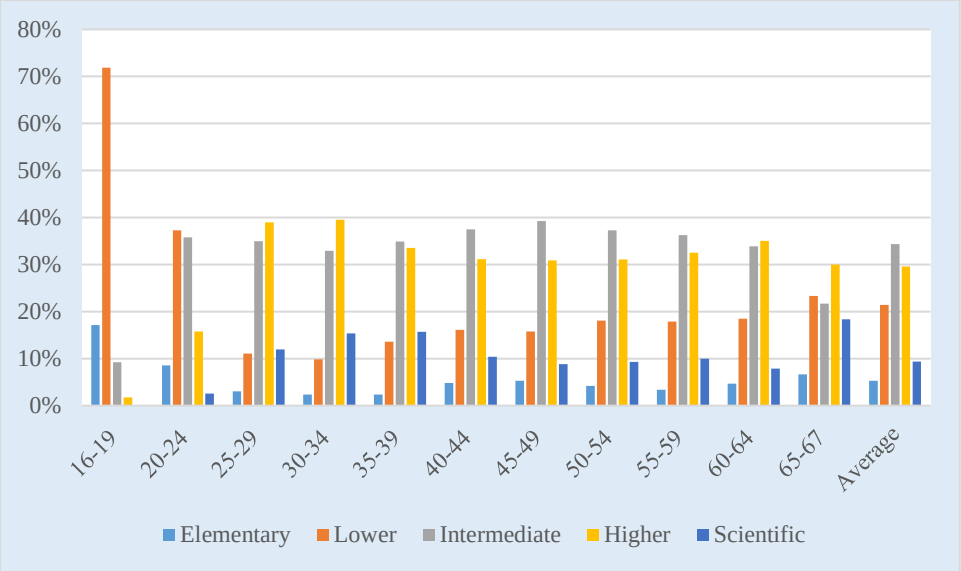
4. Exploratory analysis over job complexity, education levels and overeducation in the Netherlands

For the first part of this study, we seek to analyse in detail the distribution of overeducated workers and the factors influencing overeducation. We do so through descriptive statistics on over- and undereducation and by estimating a linear probability model to uncover which factors increase a worker's chance of being overeducated. All descriptives and estimations are based on the Labour Supply Panel of SCP.

4.1. Descriptive statistics

To start with, Figure 1 shows how job levels are distributed across age categories and education levels, as well as the average job level per education level.

Figure 1: Job level distribution over age categories



As Figure 1 demonstrates, the first age category works at far lower job levels than the other age categories, as they will not have completed their education yet if they follow an education of a higher level than SOI-4.3. This is also visible in the second category, but the categories beyond that give distribution of job levels fairly similar to each other. Interestingly, the distribution of participants aged 65-67 appears more polarized, with a higher percentage in job levels 4 and 5, as well as higher percentages in job levels 1 and 2.

Figure 2: Distribution of job levels over SOI-levels

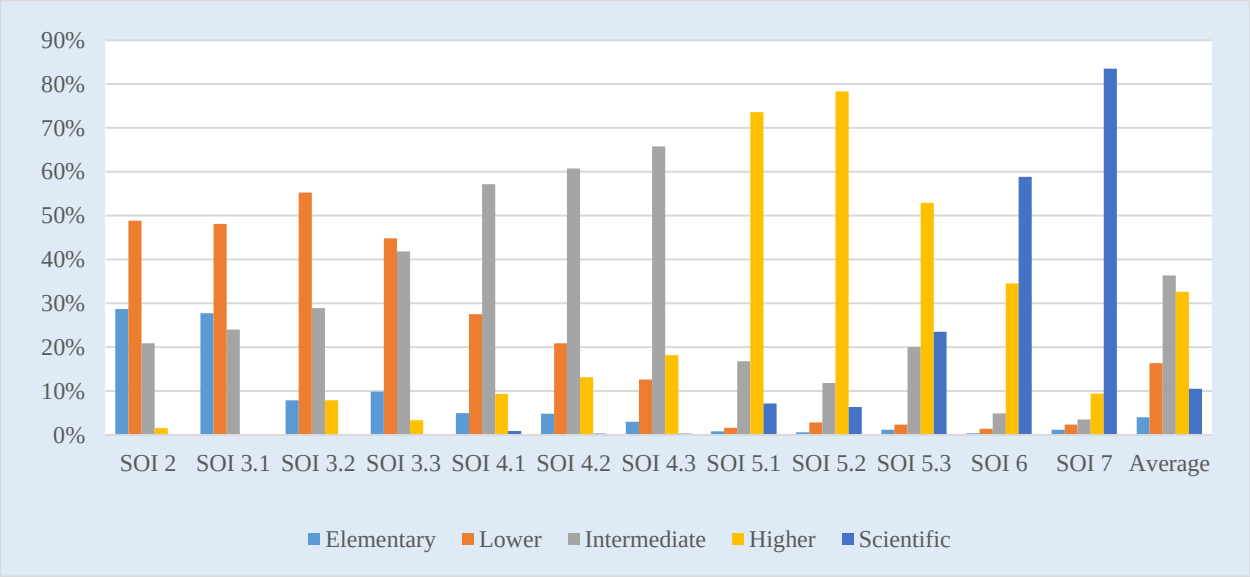


Figure 3: average job complexity per SOI education level

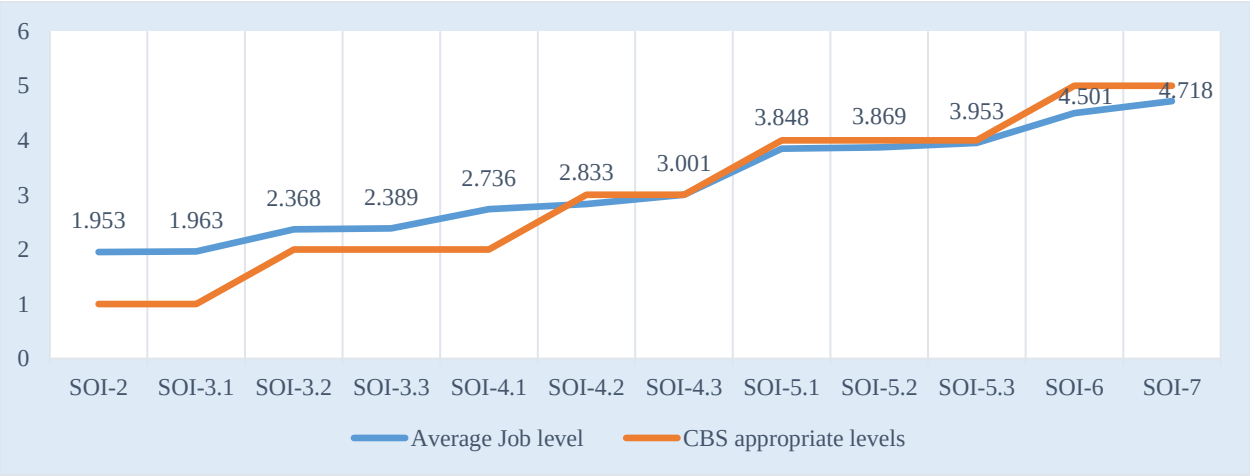
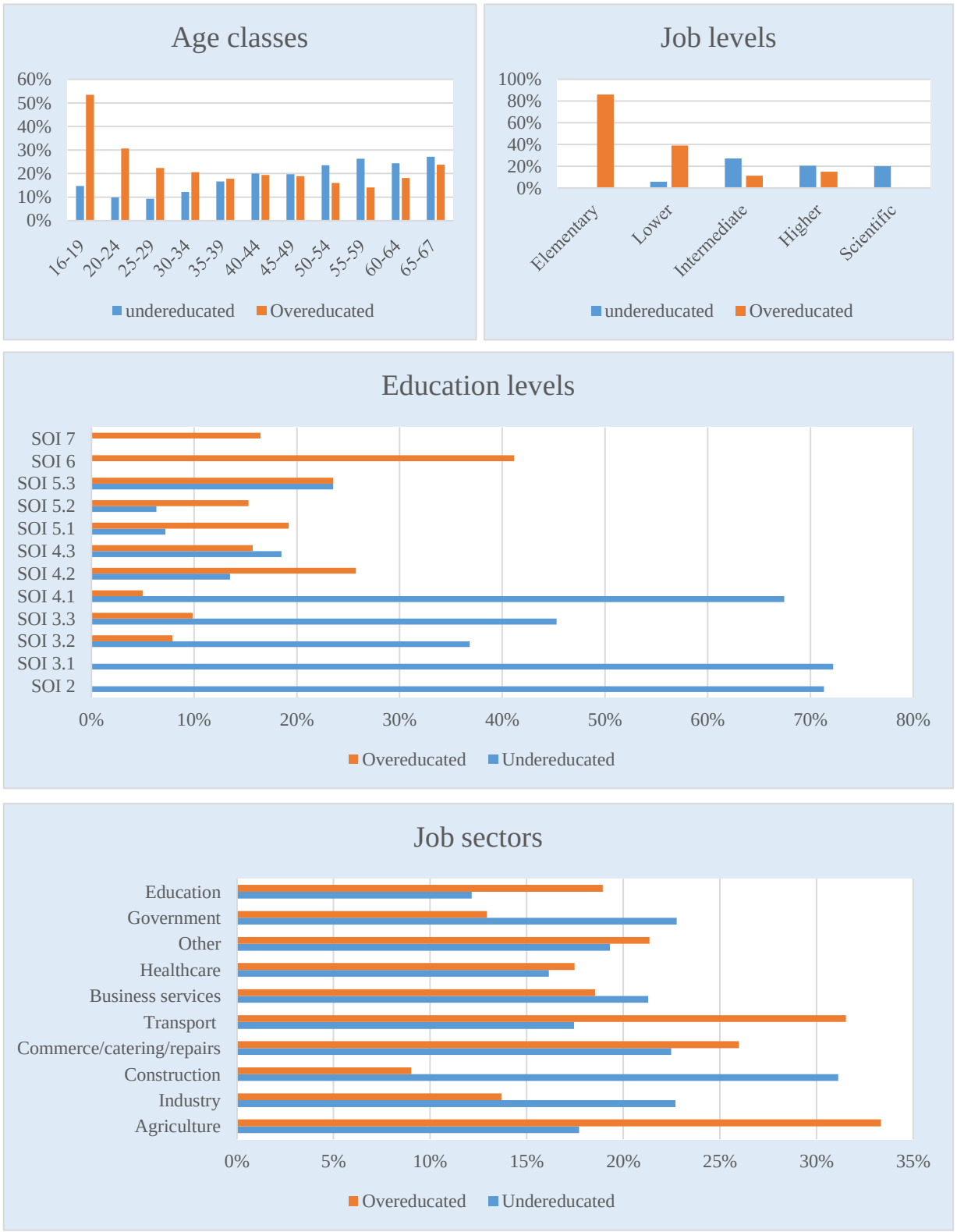


Figure 2 shows the distribution of job levels over given education levels. Figure 3 illustrates the average job complexity per SOI-level, interpreting the five job levels as a continuous variable indicating job complexity. It also shows the job complexity deemed appropriate for the education level by Statistics Netherlands. We see a gradual increase of average job complexity over the education levels, with a steep increase between the highest intermediate education SOI-level 4.3 and the first higher education SOI-level 5.1. By including the job level the CBS considers appropriate for each education level, we can see that this increase coincides with an increased appropriate job level.

We do see an interesting trend of the elementary and lower educated workers having a higher average job level than the norm of Statistics Netherlands, and workers from SOI 4.2 onward having a lower average job level. This could indicate that workers from SOI levels 4.2

onwards are more likely to be overeducated than undereducated. Figure 4 confirms this, showing how over- and undereducated workers are distributed over years, job levels, education levels and sectors.

Figure 4: Distribution of over- and undereducated workers in the Netherlands



As expected, Figure 4 confirms that the lower age categories show higher levels of overeducation compared to the other categories. The age category 65-67 also shows a relatively high percentage of overeducated workers. We also see a gradual increase of undereducated workers at older age categories, illustrating the fact that at later ages, more workers move up in job levels without getting higher formal education. It furthermore shows that the first two job levels are occupied to a very high degree by overeducated workers, with the first job level having 80.4 per cent overeducated workers. The highest level of undereducated workers are within the intermediate (SBC-3) job level, at 27.2 per cent.
As we see from figure 4, the amount of overeducated workers exceeds the amount of undereducated workers from education level SOI-4.2 onwards. This stems from a steep increase in overeducation and a steep decrease in undereducation at the same time.
As for the distribution over job sectors, we see some interesting disparities, such as a very high under- to overeducation ratio in the government, construction and industry sectors, and a high over-to undereducation ratio in the agriculture, transport and education sector. These differences in the distributions will still be visible in the regressions on returns to overeducation when we look at which sector yields the highest returns to overeducation.

4.2. Linear probability model: factors increasing chances of being overeducated

The linear probability model in table 2 provides a more detailed picture of the drivers of overeducation in the Netherlands. The first model in the table illustrates the effect of occupation levels on the chances of the worker being overeducated. We can see the similarities between these outcomes and the overall distribution of overeducated workers per job level in figure 4. SBC-1 and SBC-2 are most likely to employ workers that are overeducated, with relative increases of 83.0% and 38.5%, whereas SBC-3 has the lowest effect on chances of overeducation at 9.6%. Model 2 shows the effects of the different education levels on the chance of overeducation. Here too we see a clear link between this model and the descriptive statistics of figure 4. Contrasting to what we would expect from RBTC theory, there is no clear indication that middle-educated workers are more prone to being overeducated than higher educated workers. The biggest outlier is SOI-6, which has the highest effect on chances of overeducation with 42.9%. Besides that, the middle and high education levels all lie between around 15% and 25%, with no clear decreasing pattern over the levels. Both models indicate that age decreases chances of overeducation, but the (negative) squared age coefficient indicates that this effect gets smaller with increasing age.

The year 2014 has a significant effect in the first model, but when controlling for education levels this effect disappears.

Table 3 extends the model by controlling for economic growth, to see to what degree overeducation stems from economic downturns and crowding out effects. This model shows no evidence for cyclical overeducation, although admittedly a more complex model and definition of economic growth might show different results.

In table 4 we see what effects different sectors have on the chances of being overeducated. Interestingly, only the construction and industry sector show statistically significant coefficients, with negative effects.

Table 2: Estimates of overeducation probabilities (Linear probability model).

VARIABLES	(1)	(2)
Job level (reference value: SBC-5):		
SBC-1	0.830***	
	(0.0228)	
SBC-2	0.385***	
	(0.0151)	
SBC-3	0.0968***	
	(0.0133)	
SBC-4	0.131***	
	(0.0133)	
Education level (reference value: SOI-2)		
SOI 3.1		-0.0490
		(0.0614)
SOI 3.2		0.0444
		(0.0719)
SOI 3.3		0.0921**
		(0.0358)
SOI 4.1		0.0459
		(0.0402)
SOI 4.2		0.239***
		(0.0363)
SOI 4.3		0.148***
		(0.0355)
SOI 5.1		0.215***
		(0.0481)
SOI 5.2		0.151***
		(0.0355)
SOI 5.3		0.258***
		(0.0536)
SOI 6		0.429***
		(0.0366)
SOI 7		0.187***
		(0.0456)
Working Hours	-0.00160***	-0.00744***
	(0.000407)	(0.000427)
Age	-0.0405***	-0.0525***
	(0.00829)	(0.00895)
Age^2	0.00191***	0.00339***
	(0.000669)	(0.000720)
Gender	0.00965	-0.0319***
	(0.00899)	(0.00965)
Year (reference value: 2010):		
2012	0.0120	-0.00398
	(0.00915)	(0.00978)
2014	0.0424***	0.0149
	(0.00916)	(0.00979)
Constant	0.201***	0.472***
	(0.0373)	(0.0498)
Observations	8,453	8,399
R-squared	0.229	0.108

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Overeducation probability estimates: economic growth (linear probability model)

VARIABLES	(1) model 1
Economic growth	-0.0189
	(0.0250)
Job level (reference value: SBC-5):	
SBC-1	0.820***
	(0.0392)
SBC-2	0.382***
	(0.0256)
SBC-3	0.107***
	(0.0226)
SBC-4	0.107***
	(0.0228)
Interactions: job level*economic growth	
SBC-1 *economic growth	0.0158
	(0.0469)
SBC-2*economic growth	0.00408
	(0.0312)
SBC-3*economic growth	-0.0154
	(0.0277)
SBC-4*economic growth	0.0354
	(0.0280)
Age	-0.0402***
	(0.00829)
Age^2	0.00188***
	(0.000669)
Gender	0.00935
	(0.00899)
Year (reference value: 2010)	
2012, omitted	-
2014	0.0416***
	(0.00918)
Working hours	-0.00159***
	(0.000407)
Constant	0.217***
	(0.0401)
Observations	8,453
R-squared	0.230

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Overeducation probability estimates by job sector (linear probability model)

VARIABLES	(1) model 1
Job level (reference value: SBC-5):	
SBC-1	0.865***
	(0.0236)
SBC-2	0.416***
	(0.0161)
SBC-3	0.127***
	(0.0138)
SBC-4	0.139***
	(0.0133)
Age	-0.0350***
	(0.00830)
Age^2	0.00141**
	(0.000670)
Gender	-7.05e-05
	(0.00943)
Year (reference value: 2010):	
2012	0.0136
	(0.00911)
2014	0.0432***
	(0.00912)
Working hours	-0.00144***
	(0.000417)
Sector (reference: agriculture):	
Industry	-0.0837*
	(0.0465)
Construction	-0.122**
	(0.0486)
Commerce/catering/repairs	-0.0320
	(0.0465)
Transport	-0.0373
	(0.0476)
Business services	-0.00638
	(0.0461)
Healthcare	-0.0514
	(0.0462)
Other	-0.00523
	(0.0486)
Government	-0.0143
	(0.0468)
Education	0.0414
	(0.0470)
Constant	0.207***
	(0.0576)
Observations	8,411
R-squared	0.239

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

5. Results

5.1. Returns to education

We start by estimating returns to education in general, through an OLS model on monthly wages, controlling for continuous mean-centred job level and education levels and allowing for an interaction effect between the two. The intuition behind the interacted effect is that this should reveal higher returns to education at higher job levels, or reverse. In this regression, job level and education level are expressed in five classes and act as continuous variables. Model 2 regresses for percentual effects of the same variables on wage, by regressing on its natural logarithm.

Table 5: Returns to education

VARIABLES	(1) Wage	(2) log(Wage)
Job level mean-centred	221.9***	0.171***
	(11.43)	(0.00562)
Education level mean-centred	96.37***	0.0215***
	(11.12)	(0.00546)
Job level mean-centered.*Education level mean-centered	66.20***	0.00214
	(7.213)	(0.00354)
Age	199.2***	0.128***
	(16.61)	(0.00816)
Age^2	-10.92***	-0.00752***
	(1.336)	(0.000656)
Year (reference value: 2010):		
2012	53.55***	0.0381***
	(18.19)	(0.00894)
2014	108.5***	0.0646***
	(18.13)	(0.00891)
Gender	-294.0***	-0.122***
	(17.85)	(0.00877)
Working hours	35.93***	0.0295***
	(0.824)	(0.000405)
Constant	122.8*	6.024***
	(70.58)	(0.0347)
Observations	8,074	8,074
R-squared	0.517	0.679

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

We can see that, when controlling for job complexity and education level, there are statistically significant positive returns to further education, measured in absolute wage

levels. This illustrates that at higher job levels, the returns to one's education increase. Model 2 however shows that when we look at percentual change, these returns are no longer significant.
Given that the number of education and job levels are equal, we furthermore take from this model that job level has a stronger average effect on wage (221.9) than education level (96.4). This concurs with existent literature on education and overeducation (Hartog, 2000; Psacharopoulos & Patrinos, 2018).

5.2. Costs and returns to overeducation and undereducation

Table 6 gives more detailed insights into the variation of returns over different job and education levels. Model 1 regresses wages on the overeducation dummy and includes controls for education categories, meaning that it models for the average costs of overeducation compared to the average wage of given education levels. Our table shows that the average costs of overeducation are -302.9 euros per month, with a confidence interval of $p<0.01$. Model 2 regresses wages on overeducation and includes controls for job level, meaning that it models for the average returns to overeducation given a certain job level. Our table shows that these average returns are 118.5 euros and statistically significant. Model 3 regresses wages on overeducation including both job level and education level as controls. This implies that this model shows any additional returns to overeducation, given education level and job level. Similar to the linear and continuous model that was regressed on log wages, our model then shows no such significant additional returns. Model 4 includes an interaction term between overeducation and education level, while controlling for job level. This model illustrates at which education level overeducation yields positive returns given the job level. We can see that at the scientific education level, overeducation yields an average return of 228.7 euros per month, while at the intermediate education level there is a weakly significant average return of 61.2 euros per month at a confidence margin of 0.1. Finally, model 5 uses an interaction term between overeducation and job level, while controlling for the education level. This model shows us the costs of overeducation for each job level, given the education level. We see that overeducation at the intermediate and lower job level have the highest average costs at relative scores of -395.1 and -316.5, while the higher and elementary job level have average losses of -243.9 and -226.1.

These five models each provide us a specific piece of information about the returns to overeducation. By combining these pieces, we get a more generalizable picture of returns to

overeducation in the Netherlands. Let us start by combining the information of model 1 and 2. In particular, model 1 indicates a -302.9 average loss to overeducated workers compared to workers in the same education level who work at an appropriate job level, and model 2 shows a 118.5 average gain to overeducated workers compared to workers in the same job level who have a more appropriate education level. We can interpret these disparities as the wage spread between overeducated and non-overeducated worker. Specifically, in a range of 302.9+118.5=421.4, 118.5 is made up of gains to overeducation and 302.9 of losses (or, in percentages: 28.1% gains and 71.2% losses). Stated differently, overeducated workers would earn 303 euros more if they would be rewarded equally as equally educated workers in suitable job position, but they would earn 118 euros less when rewarded equally to lower educated workers in similar job positions.

Model 4 and 5 indicate more precisely in which job and education levels overeducation has the strongest gains and losses, respectively. In the scientific education level we see a strongly significant average gain of 228.7, while the other education levels show no significant gains, except for the intermediate education level at a weakly significant average gain of 61.15. Model 5 shows the average losses to overeducation per job level. These are significant at every job level, with the highest loss at SBC-level 3 with -395.1 euros per month.

Table 7 shows the outcomes of similar model specifications as in table 6, but now with a focus on undereducated workers. In model 1 we see average benefits of 264.7 euros to undereducation compared to workers in the same education level who are not undereducated. The costs to undereducation compared to workers in the same job level who are not undereducated can be inferred from the second model, at -112.0. Concurrent with existing literature, we see that there are costs to undereducation, but these are smaller than returns to the job level (which can be seen as an equivalent to Hartog's (2000) 'required education'). There is one exception at job level SBC-2, which shows no significant returns. Furthermore, if we compare the results of table 6 to table 7, we see that the average benefits to overeducation (118.5) are about as high as the costs to undereducation (-112.0). The overall bandwidth of costs to benefits is 264.7+112.0=376.7, where 29.7% is made up of costs and 70.3% is made up of benefits, mirroring the results of the analysis of costs and benefits to overeducation.

The returns to undereducation are quite high at the higher job levels, especially at SBC-4 with 354.0 euros per month. The costs to undereducation increase at higher education levels as well, with a 263.2 euros per month average loss at the higher education level.

Table 6: Returns to overeducation on monthly wages:

VARIABLES	(1) model 1	(2) model 2	(3) model 3	(4) model 4	(5) model 5
Overeducated	-302.9***	118.5***	-17.39		
	(20.31)	(21.85)	(31.07)		
Education level (reference value: elementary):					
Lower education	120.6**		22.75		113.2*
	(58.54)		(59.52)		(58.61)
Intermediate education	330.0***		102.6*		329.2***
	(57.79)		(62.32)		(57.87)
Higher education	603.2***		125.2*		613.4***
	(58.18)		(69.09)		(58.26)
Scientific education	1,096***		369.4***		1,080***
	(60.75)		(82.56)		(61.75)
Job level (reference value: SBC-1)					
SBC-2		94.03**	26.47	89.99	
		(42.60)	(44.51)	(55.29)	
SBC-3		351.6***	214.8***	331.7***	
		(42.75)	(50.06)	(57.44)	
SBC-4		725.8***	531.0***	687.5***	
		(43.37)	(60.18)	(58.24)	
SBC-5		1,197***	814.0***	1,173***	
		(48.73)	(79.45)	(61.59)	
Interactions: overeducated*education level					
Overeducated*lower				105.9	
				(76.79)	
Overeducated*intermediate				61.15*	
				(36.33)	
Overeducated*higher				43.47	
				(37.24)	
Overeducated*scientific				228.7***	
				(35.16)	
Interactions: Overeducated*job level					
Overeducated*SBC-1					-226.1***
					(42.50)
Overeducated*SBC-2					-316.5***
					(31.51)
Overeducated*SBC-3					-395.1***
					(39.88)
Overeducated*SBC-4					-243.9***
					(44.39)
Year (reference value: 2010)					
2012	55.65***	51.84***	50.75***	53.43***	56.58***
	(18.41)	(18.24)	(18.19)	(18.23)	(18.41)
2014	113.1***	105.5***	102.7***	105.2***	111.9***
	(18.35)	(18.21)	(18.17)	(18.20)	(18.37)
Gender	-285.2***	-290.2***	-291.9***	-290.0***	-286.9***
	(18.06)	(17.91)	(17.87)	(17.89)	(18.08)
Working hours	37.68***	36.18***	36.22***	36.07***	37.79***
	(0.820)	(0.827)	(0.826)	(0.827)	(0.828)
Age	207.4***	194.9***	195.9***	194.1***	206.6***
	(16.80)	(16.66)	(16.62)	(16.65)	(16.82)
Age^2	-11.36***	-10.72***	-10.68***	-10.69***	-11.30***
	(1.352)	(1.340)	(1.338)	(1.339)	(1.354)
Constant	-299.2***	-309.1***	-251.9***	-277.3***	-297.4***
	(88.46)	(78.54)	(91.65)	(87.00)	(88.58)
Observations	8,074	8,074	8,074	8,074	8,074
R-squared	0.506	0.516	0.519	0.517	0.507

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Returns to undereducation on wages:

VARIABLES	(1) model 1	(2) model 2	(3) model 3	(4) model 4	(5) model 5
Undereducated	264.7***	-112.0***	-47.32*		
	(21.82)	(19.49)	(28.74)		
Education level (reference value: Elementary)					
Lower education	143.4**		-0.0141		-4.644
	(58.94)		(59.41)		(80.07)
Intermediate education	405.2***		50.75		204.5**
	(58.99)		(63.78)		(80.34)
Higher education	714.2***		46.04		524.2***
	(59.91)		(72.80)		(80.69)
Scientific education	1,156***		267.5***		969.9***
	(62.37)		(81.01)		(82.20)
Job level (reference value: SBC-1)					
SBC-2		44.89	39.75	42.68	
		(41.48)	(41.81)	(41.60)	
SBC-3		290.0***	255.0***	290.1***	
		(39.83)	(43.39)	(40.03)	
SBC-4		659.4***	596.6***	651.7***	
		(40.50)	(52.57)	(40.74)	
SBC-5		1,116***	903.0***	1,144***	
		(45.21)	(64.76)	(46.36)	
Interactions: Undereducated*education level					
Undereducated*elementary				-70.82	
				(69.12)	
Undereducated*lower				-114.0***	
				(26.92)	
Undereducated*intermediate				-66.96**	
				(33.71)	
Undereducated*higher				-263.2***	
				(58.43)	
Interactions: Undereducated*job level					
Undereducated*SBC-2					-83.84
					(113.3)
Undereducated*SBC-3					166.0***
					(32.84)
Undereducated*SBC-4					354.0***
					(31.83)
Undereducated*SBC-5					318.7***
					(53.46)
Age	211.5***	195.0***	196.1***	195.1***	212.0***
	(16.87)	(16.65)	(16.62)	(16.65)	(16.85)
Age^2	-11.71***	-10.67***	-10.68***	-10.66***	-11.81***
Year (reference value: 2010)	(1.358)	(1.340)	(1.337)	(1.340)	(1.356)
2012	51.98***	53.11***	51.09***	51.34***	50.35***
	(18.50)	(18.23)	(18.19)	(18.24)	(18.49)
2014	100.3***	109.8***	103.3***	108.3***	99.30***
	(18.45)	(18.18)	(18.17)	(18.19)	(18.44)
Gender	-281.8***	-292.0***	-292.3***	-290.9***	-280.6***
	(18.14)	(17.91)	(17.87)	(17.91)	(18.12)
Working hours	38.43***	36.24***	36.21***	36.26***	38.40***
	(0.819)	(0.827)	(0.826)	(0.827)	(0.818)
Constant	-508.6***	-210.9***	-235.1**	-212.5***	-321.2***
	(88.54)	(75.84)	(91.86)	(75.82)	(103.8)
Observations	8,074	8,074	8,074	8,074	8,074
R-squared	0.502	0.516	0.519	0.517	0.503

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

5.3. Returns to overeducation across job sectors

We now run the same OLS models on monthly wages, but we add an interaction term between overeducation and job sector. We next conduct similar bandwidth analyses of gains and losses to overeducation. In these models, we focus on one single interaction effect for all education and job complexity levels, and models 1 and 2 we differentiate only in terms of controls. Therefore, we can compare costs and returns to job sectors by their individual bandwidth. The relative returns in these bandwidths are shown in the third column of the regression table.

Table 8 shows the results of the interaction term with job sectors, with model 1 including controls for job level and model 2 for education level. The third column illustrates what percentage (if any) the gains are to overeducation within the range of gains to losses to overeducation from model 1 and 2. It should be noted that these estimates will be far less precise in cases of insignificant gains and/or losses in model 1 and 2. As we can see in model 1, the agricultural, construction, transport and education sector and 'other' sectors interacted with overeducation show no significant returns in the monthly wage. For agriculture, it is possible that this is the result of a small number of observations. The other sectors with no significant returns to overeducation all count about 500 observations, with the education sector as an outlier with more than 1,000 observations; Sectors with significant gains have more observations, from 907 (government) to 2,273 (healthcare). The government and business services sectors yield relatively steep returns to overeducation. Column 3 shows these returns to account for around 50% in the range of gains to losses to overeducation, which is considerably higher than the overall average returns of 28.12% as found in table 6. The second model, controlling for education level, shows significant losses to overeducation given the education level in all sectors but the agriculture sector (which, as mentioned before, is likely due to too little observations). We see especially high losses in the transport sector (-456.5) and the education sector (-477.8). This observation has even more impact when we consider that both sectors showed no significant returns in model 1, meaning that in these sectors, the variance between overeducated and non-overeducated workers is explained fully by losses to overeducation. This indicates that in these job sectors, there are no discernible extra returns to overeducation compared to appropriately schooled workers in the same job level, while there are very high costs to overeducation compared to appropriately schooled workers in the same education level.

Table 8: Returns to overeducation on wage over job sectors:

VARIABLES	(1) control: job level	(2) control: education level	(3) Returns to overeducation[1]
Job level (reference value: SBC-1)			
SBC-2	92.43**		
	(43.93)		
SBC-3	345.4***		
	(44.30)		
SBC-4	720.8***		
	(45.19)		
SBC-5	1,190***		
	(50.08)		
Education level (reference value: elementary)			
Lower		115.3**	
		(58.74)	
Intermediate		328.9***	
		(58.00)	
Higher		600.3***	
		(58.38)	
Scientific		1,097***	
		(61.13)	
Interactions: overeducated*Job sector			
Overeducated*Agriculture	233.8	-174.4	57.3%
	(153.0)	(153.0)	
Overeducated*Industry	134.6**	-282.1***	32.3%
	(61.66)	(61.06)	
Overeducated*Construction	192.7	-220.8*	46.6%
	(123.1)	(123.7)	
Overeducated*Commerce/catering/repairs	105.7**	-284.2***	27.1%
	(45.84)	(44.32)	
Overeducated*Transport	-23.81	-456.5***	5.0%
	(57.69)	(55.11)	
Overeducated*Business services	239.4***	-201.0***	54.4%
	(43.44)	(43.83)	
Overeducated*Healthcare	85.31**	-303.0***	22.0%
	(41.18)	(39.84)	
Overeducated*Other	116.0	-305.1***	27.5
	(78.69)	(79.10)	
Overeducated*Government	209.7***	-228.2***	47.9%
	(65.34)	(66.14)	
Overeducated*Education	-0.595	-477.8***	0%
	(55.28)	(57.32)	
Age	196.8***	209.0***	
	(16.71)	(16.86)	
Age^2	-10.74***	-11.33***	
	(1.343)	(1.355)	
Year (reference value: 2010)			
2012	47.34***	50.56***	
	(18.25)	(18.41)	
2014	105.6***	112.9***	
	(18.19)	(18.33)	
Gender	-284.6***	-280.6***	
	(17.98)	(18.14)	
Working hours	36.20***	37.77***	
	(0.831)	(0.826)	
Constant	-322.9***	-317.2***	
	(79.54)	(88.71)	
Observations	8,041	8,041	
R-squared	0.518	0.509	

[1] The third column of this regression table shows the relative gains to overeducation in the given job sector.

5.4. Returns to overeducation: changes over time

We now extend our model by allowing an interaction effect between overeducation and panel years. Table 9 shows our results.

Model 1 and 2 are set up in similar fashion to model 1 and 2 of table 6, except that we now allow year dummies to interact with overeducation. As we can see, the year 2010 yields percentual gains of 28.4% in the gains to losses range. This deteriorates to 25.5% in 2012 and then recovers in 2014 with 30.1%. The differences between these years however are not significant. In model 3 and 4, we test whether these percentual changes of gains to overeducation follow the trends of economic growth: we do see higher gains and lower costs during economic growth and lower gains and higher costs during economic downturns, but these higher gains too are not statistically significant.

Table 9: Returns to overeducation on monthly wage over panel years:

VARIABLES	(1) control: job level	(2) control: education level	(3) Economic growth: job level	(4) Economic growth: education level
Job level (reference value: SBC-1)				
SBC-2	94.17**		94.21**	
	(42.60)		(42.60)	
SBC-3	351.5***		351.8***	
	(42.76)		(42.75)	
SBC-4	725.6***		725.9***	
	(43.39)		(43.37)	
SBC-5	1,197***		1,197***	
	(48.74)		(48.73)	
Education level (reference value: Elementary)				
Lower		120.5**		120.5**
		(58.54)		(58.54)
Intermediate		329.8***		330.0***
		(57.80)		(57.79)
Higher		603.3***		603.3***
		(58.19)		(58.18)
Scientific		1,096***		1,095***
		(60.76)		(60.76)
Interactions: overeducated*year				
Overeducated*2010	114.5***	-289.0***		
	(36.44)	(35.49)		
Overeducated*2012	107.1***	-312.7***		
	(35.10)	(34.38)		
Overeducated*2014	131.5***	-305.8***		
	(33.09)	(32.60)		
Interactions: overeducation*economic growth				
Overeducated*Economic growth=0			107.3***	-312.7***
			(35.10)	(34.37)
Overeducated*Economic growth=1			123.9***	-298.1***
			(25.62)	(24.40)
Age	195.0***	207.5***	195.0***	207.5***
	(16.66)	(16.81)	(16.66)	(16.80)
Age^2	-10.72***	-11.37***	-10.72***	-11.37***
	(1.340)	(1.353)	(1.340)	(1.353)
Year (reference value: 2010)				
2012	53.13***	59.74***	54.71***	58.16***
	(20.03)	(20.23)	(19.55)	(19.73)
2014	102.3***	116.0***	105.4***	113.0***
	(20.09)	(20.28)	(18.21)	(18.35)
Gender	-290.0***	-285.5***	-290.2***	-285.3***
	(17.92)	(18.07)	(17.91)	(18.06)
Working hours	36.17***	37.68***	36.18***	37.68***
	(0.827)	(0.820)	(0.827)	(0.820)
Constant	-308.8***	-301.1***	-310.2***	-300.0***
	(78.69)	(88.55)	(78.60)	(88.49)
Observations	8,074	8,074	8,074	8,074
R-squared	0.516	0.506	0.516	0.506

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

5.5. Robustness checks

We have developed models as means of robustness checks for our main results. We start by running our main regression model estimating the returns to overeducation on monthly wage, but omitting controls for age effects. Table 10 shows the results of this regression.

Table 10: Returns to overeducation on monthly wage, without age effects:

VARIABLES	(1) model 1	(2) model 2
Overeducated	-340.7***	73.98***
	(20.86)	(22.35)
Education level (reference value: Elementary)		
Lower	117.2*	
	(60.30)	
Intermediate	292.3***	
	(59.43)	
Higher	566.8***	
	(59.85)	
Scientific	1,059***	
	(62.45)	
Job level (reference value: SBC-1)		
SBC-2		76.35*
		(43.76)
SBC-3		332.0***
		(43.91)
SBC-4		708.9***
		(44.56)
SBC-5		1,173***
		(50.04)
Year (reference value: 2010)		
2012	36.70*	32.69*
	(18.95)	(18.72)
2014	99.81***	90.87***
	(18.89)	(18.68)
Gender	-332.9***	-337.9***
	(18.41)	(18.17)
Working hours	35.37***	33.88***
	(0.837)	(0.841)
Constant	708.0***	636.7***
	(74.66)	(60.70)
Observations	8,074	8,074
R-squared	0.475	0.489

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

When we compare the results of this model to those in our main regression table, we see a small but insignificant rise in costs, and a small but insignificant fall in returns to overeducation.

We also run a fixed-effects model, to check for potential omitted variable biases from invisible characteristics. The results of this model offer quite a different image of overeducation from our main model. Table 11 shows the results of this model.

This table shows no significant returns to overeducation both when including controls for education and for job levels. Moreover, these education and job levels show no significant effects on monthly wages as well. We consider two possible explanations for these results being so different from those in our main regression table. The first explanation would be that there are indeed omitted variables in our main regression, which actually explain the differences in wages between overeducated and appropriately educated workers. In other words: although overeducated workers do indeed earn less compared to appropriately educated workers in the same education level, this is not the result of their overeducation but of unobserved individual characteristics. This would be more in line with findings by Nelissen (2007), who found evidence that individual characteristics have a far greater effect on earned income than formal education. A second explanation would be that when workers move to higher job levels, there is a delay in the wage levelling with the average wage earned by workers at the same job level of the same education level. This would also explain why both education levels and job levels show no significant effects on earned wages.

Table 11: Returns to overeducation on monthly wages, fixed effects model:

VARIABLES	(1) model 1	(2) model 2
Overeducated	8.525	12.30
	(19.64)	(23.20)
Education level (reference value: Elementary)		
Lower	-70.41	
	(60.45)	
Intermediate	-96.59	
	(64.48)	
Higher	-59.72	
	(72.33)	
Scientific	-74.11	
	(86.41)	
Job level (reference value: SBC-1)		
SBC-2		21.04
		(46.99)
SBC-3		34.07
		(52.94)
SBC-4		67.64
		(59.85)
SBC-5		39.19
		(71.73)
Age	163.4***	163.3***
	(36.89)	(36.84)
Age^2	-12.24***	-12.35***
Year (reference value: 2010)	(2.475)	(2.474)
2012	64.26***	63.12***
	(12.26)	(12.26)
2014	167.5***	166.5***
	(16.55)	(16.59)
Working hours	15.46***	15.38***
	(1.134)	(1.136)
Constant	797.5***	688.1***
	(153.1)	(150.6)
Observations	8,074	8,074
R-squared	0.109	0.109
Number of koppelnr	4,326	4,326

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

6. Conclusion

As a result of the dynamic nature of the labour market, policymakers should constantly be on the lookout on the implications of demand changes in the market for the effectiveness of current socio-economic policies. While older evidence points at the effectiveness of education in term of increasing average wages and productivity, one might question the relevance of this in a labour market that has become more flexible, particularly when it comes to the role of overeducation. To answer these questions, this study has provided an up-to-date framework of overeducation in the Netherlands, using survey data from the Dutch Labour Supply panel. In doing so, we provide further evidence for existent claims, as well as offering new evidence on where in the labour market overeducation is most prevalent and what returns we can expect over different education levels, job levels and job sectors.

Based on linear probability models, we show no clear evidence that overeducation is more prevalent in the middle-educated segments of the working population, as we would have expected to see if there were strong trends of job polarization. This does not mean that such trends do not exist at all: an analysis of overeducated workers with an education focused on routine-type occupations might very well show that job polarization is prevalent in the Netherlands. However, so far any such trends have not had as clear an effect on middle-skilled workers as to massively push them into lower job positions far more often than, for example, higher skilled workers. We also found no strong evidence for cyclical overeducation, but here too, more specific models might show small effects that our models could not uncover. We have seen that the job sector a worker belongs to is generally no significant explanatory factor for overeducation, with the exception of the industry and the construction sector, which both have a significant negative effect on the chances of being overeducated. A possible explanation could be that education towards these sectors is generally quite sector-specific, and not easily transferrable to other types of occupations. Presumably, higher educated workers are less likely to switch to manual work at lower job complexity levels.

Our first OLS model, in which we have defined returns to education and job levels as the effects of both on monthly wages, illustrates that the job level has higher returns to wages than the education level, which is in line with findings by Hartog (2000). In our models estimating returns to overeducation, we have defined returns to overeducation as the effect of overeducation on wages compared to non-overeducated workers in the same education level

or in the same job level. Our findings on the returns and costs to overeducation are similar to the studies discussed by Hartog (2000). We see positive returns to overeducation given the job level, but compared to workers of the same education level working at an appropriate job level, overeducated workers earn less. These costs are also generally higher than the returns to overeducation given the job level. Furthermore, our results follow Hartog's (2000) findings that undereducation has negative returns on wages, and while he finds that most studies show smaller costs to undereducation than returns to overeducation, our model showed the costs to undereducation mirroring the returns to overeducation.

Interestingly, we have found that returns to overeducation are not visible in all education levels, with only the scientific and intermediate education levels showing significant positive returns. Whether a worker can expect positive returns to overeducation compared to non-overeducated workers in the same job level thus depends largely on the education level the worker has obtained. These disparities in returns to overeducation are also visible in models estimated on samples that are stratified by job sectors: The government sector and business services sectors show relatively high returns to overeducation given the job level, higher even than the costs to overeducation given the education level. Yet other job sectors like the education and transport sectors show no significant returns whatsoever, while the costs to overeducation given the education level are relatively high. This illustrates that, while in the labour market as a whole there are returns to overeducation, there are many areas in the labour market where no such returns are visible.

This study has shown that positive returns to overeducation strongly depend on the education level and job sector of the worker. This means that, depending on the direction and the level of the education, some overeducated workers end up with no returns to their superfluous education. But what does that mean for the effectiveness of education in increasing wages and productivity?

Answering this question requires a more nuanced approach of analysing returns to education and skills, as is further exemplified by recent policy proposals from the Commission on Regulating Work (Commissie Regulering van Werk), led by Hans Borstlap. In January 2020 the Commission presented its findings, underscoring the expectation that mainly middle class jobs will be replaced by technological advancements, and that these developments require the Dutch labour force to be able to adapt to new skill requirements. As for training of workers, the Commission recommends creating a government funded 'personal development budget

(Persoonlijk Ontwikkelbudget)' that workers can use for training and further education. The Commission stresses the importance of a clear link between the training and education offer and the current and future demands of the labour market, and therefore recommends setting up an agency that oversees the modules that can be paid for from the personal development budget (Borstlap, 2020).

These policy proposals show that the debate surrounding the effectiveness of schooling has to be held on more dimensions than just the friction between job level and education level. Upgrading workers' skills in order to keep up with changing demands in the job market will often mean expanding one's education or transitioning to a different sector, but not necessarily increasing one's education level; in other words, formal education is not the only factor in explaining returns to skills and potential. In fact, based on the results of this study, the effects of formal education levels are quite limited in all but the highest education levels, especially for overeducated workers. This also nuances the question whether overeducation is actively harmful or not: Although we do see that not all workers receive the full returns to their education as a result of overeducation, we cannot say whether these are actual cases of unutilized potentials, because the formal education level might be too simplistic a parameter for defining one's potential. The degree to which a worker has upgraded or expanded skills without obtaining a higher education level, as well as individual non-cognitive characteristics, might need to be included into the evaluation of unfulfilled potentials as well. Our fixed-effect models have illustrated the degree to which controlling for these can change the outcomes of returns to education and overeducation, although the fact that the job level has no effect on wages in this model as well does obfuscate whether these invisible characteristics are really the reason for the lack of returns to (over)education, or if this is the result of lagged effects in wage corrections.

Further research into the effects of updating skills and of non-cognitive characteristics rather than formal education and overeducation can offer additional insights into which of these factors increase or decrease returns on monthly wages. Secondly, further research would be needed on the effects of job polarization and crowding out effects on overeducation in the Netherlands. This could, together with further analysis into the role of non-cognitive characteristics in defining wages, certainly enhance our understanding of the degree to which overeducation is the result of factors of the market and to which degree it is the result of individual characteristics. This will in turn provide a more nuanced answer to the question whether overeducation is a harmful phenomenon in the Dutch labour market.

References

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. *In Handbook of labor economics* (Vol. 4, pp. 1043-1171). Elsevier.

Ashenfelter, O., Layard, R. (1986). *Handbook of labor economics*. Vol. I: (Handbooks in economics; 5. 821684574). Amsterdam: Elsevier Science Publishers.

Autor, D. H., & Dorn, D. (2013). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *The American Economic Review*, 103(5), 1553-1597.

Autor, D. H., Katz, L. F., Krueger, A. B. (1998). Computing Inequality: Have Computers Changed the Labor Market? *The Quarterly Journal of Economics*, 113(4), 1169-1213.

Autor, D. H., Levy, F., Murnane, R. J. (2003). *The skill content of recent technological change. The Quarterly Journal of Economics,* 118(4), 1279-1333.

Autor, David H, Dorn, David, Hanson, Gordon H. (2015). Untangling Trade and Technology: Evidence from Local Labour Markets. *The Economic Journal* (London), 125(584), 621-646.

Berman, E., Bound, J., Machin, S. (1998). Implications of Skill-Biased Technological Change: International Evidence. *The Quarterly Journal of Economics*, 113(4), 1245-1279.

Borstlap, H. (2020). *In wat voor land willen wij werken? Naar een nieuw ontwerp voor de regulering van werk.* The Hague: Commissie Regulering van Werk.

Buyst, E., Goos, M., Salomons, A. (2018). Job polarization: An historical perspective. *Oxford Review of Economic Policy*, 34(3), 461-474.

Card, D., & DiNardo, J. E. (2002). Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles. *Journal of Labor Economics*, 20(4), 733-783.

Centraal Bureau voor de Statistiek. (2010). *Standaard Beroepenclassificatie 2010* (Editie 2010). Consulted on https://www.cbs.nl/nl-nl/onze-diensten/methoden/classificaties/onderwijs%20en%20beroepen/beroepenclassificatie--isco-en-sbc--/standaard-beroepenclassificatie-2010--sbc-2010--/-/media/af9524ec88894c10adbcba7b28203037.ashx

Duncan, G. J., & Hoffman, S. D. (1981). The incidence and wage effects of overeducation. *Economics of Education Review*, 1(1), 75-86.

Factors." *Quarterly Journal of Economics* 107: 35-78.

Gautier, P.A, Pomp, J.M, Zijl, M. (1997). *Does crowding out explain low-skilled unemployment?* The Hague: Centraal Planbureau.

Goos, M., Manning, A., Salomons, A. (2009). Job Polarization in Europe. *The American Economic Review*, 99(2), 58-63.

Hartog, J. (2000). Over-education and earnings: Where are we, where should we go? *Economics of Education Review*, 19(2), 131-147.

Johnes, G. (2019). The Incidence of and Returns to 'Overeducation': PIAAC Evidence on the G7. *Minerva* (London), 57(1), 85-107.

Josten, E., Portegijs, W., Merens, A., de Voogd-Hamelink, M. (2019). *Arbeidsmarkt in kaart: wel- en niet-werkenden – editie 1*. The Hague: Sociaal en Cultureel Planbureau.

Katz, L., Murphy., K.M. (1992) Changes in Relative Wages, 1963-1987: Supply and Demand

Khalifa, S. (2012). Job Competition, Crowding out, and Unemployment Fluctuations. *Macroeconomic Dynamics*, 16(1), 1-34.

Manning, A. (2004). We Can Work It Out: The Impact of Technological Change on the Demand for Low-Skill Workers. *Scottish Journal of Political Economy*, 51(5), 581-608.

Nelissen, J. (2007). Scholing alleen werkt niet. *Economische statistische berichten*, *92*, 328-331.

Pierrard, O., & Sneessens, H. R. (2008). Biased technological shocks, relative wage rigidities and low-skilled unemployment. *Scottish Journal of Political Economy*, 55(3), 330-352.

Pollmann-Schult, M. (2005). Crowding-out of unskilled workers in the business cycle. *European Sociological Review*, 21(5), 467-480.

Psacharopoulos, G., & Patrinos, H. A. (2018). Returns to investment in education. *Education Economics*, 26(5/6), 445-458.

Rumberger, R. W. (1987). The impact of surplus schooling on productivity and earnings. *Journal of Human Resources*, 22 (1), 24–50.

Salvatori, A., & Manfredi, T. (2019). *Job polarisation and the middle class: New evidence on the changing relationship between skill levels and household income levels from 18 OECD countries*. Paris: OECD Publishing.

Schaart, R., Bernelot, M., Westerman, S. (2008). *The Dutch standard Classification of Education, SOI 2006.* Den Haag: Centraal Bureau voor de Statistiek. Consulted on https://www.cbs.nl/-/media/imported/documents/2008/24/2008-the-dutch-standard-classification-of-education-soi-art.pdf?la=en-gb

Sicherman, N. (1991). Overeducation in the labor market. *Journal of Labor Economics*, 9 (2), 101–122.

Sociaal en Cultureel Planbureau (2020). *Arbeidsaanbodpanel* [panel database: 2010, 2012, 2014]. Consulted on https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:65614

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, 87(3), 355-374.

Teulings, C. N. (1995). The Wage Distribution in a Model of the Assignment of Skills to Jobs. The Journal of Political Economy, 103(2), 280-315.

Teulings, C., & Koopmanschap, M. (1989). An econometric model of crowding out of lower education levels. *European Economic Review*, 33(8), 1653-1664.

Thurow, L. (1975). *Generating inequality: Mechanisms of distribution in the U.S. economy*. New York: Basic Books.

Tinbergen, J. (1975). *Income distribution : Analysis and policies*. Amsterdam: Elsevier.

Van den Berg, G.J, Gautier, P.A, Van Ours, J.C, Ridder, G. (2002). Worker turnover at the firm level and crowding out of lower educated workers. *European Economic Review*, 46, 523-538.

Van der Meer, P. H. (2006). The validity of two education requirement measures. *Economics of Education Review* 25: 211–219.

Van Ours, J.C., & Ridder, G. (1995). Job matching and job competition: Are lower educated workers at the back of job queues? *European Economic Review*, 39(9), 1717-1731.

Verdugo, R., & Verdugo, N. (1989). The impact of surplus schooling on earnings. *Journal of Human Resources*, 24 (4), 629–643.