Educational Attainment, Overeducation and Wages: Evidence from a Dynamic Model*

Lorenzo Navarini† Dieter Verhaest‡

July 6, 2023

Abstract

We estimate a dynamic discrete choice model to investigate the causal relationship between educational attainment, overeducation in the first job after graduation, and subsequent wages. Moreover, we adopt a novel decomposition approach in order to determine how overeducation risk affects the expected (unconditional) wage returns to educational attainment and their distribution. To this end, we rely on longitudinal Belgian data. We find that initial overeducation generates a wage penalty that persists at least until age 29. Even so, the effect of overeducation risk on the expected return to college is found to be moderate at best and, in some cases, even positive. This is partly due to a reduced overeducation risk that results from obtaining a bachelor's degree, most likely as a consequence of job polarisation. We also find that overeducation generates substantial heterogeneity in realised (ex-post) returns to education. Overall, these results suggest that overeducation is much more indicative of search and matching frictions on the labour market than of considerable overinvestments in higher education.

Keywords: Educational Mismatch; Overeducation; Dynamic Discrete Choice Model; Returns to Education; Educational Expansion; Heterogeneous Returns

*This paper benefited from helpful comments at various stages from Olivier De Groote, Koen Declercq, Jo Van Biesebroeck, Christophe Bruneel-Zupanc, François Poinas, Kristof De Witte, and several audiences at KU Leuven and at TSE. It also benefited from discussions during and after presentations at LESE 2022, LEER 2022, TIY 2022, FRDB Workshop for Fellows 2022, Applied Economics Conference 2022, and IWAEE 2023. Funding for this project was generously provided by the Research Foundation Flanders (FWO) - G079420N.

†KU Leuven, ECON Research Group, campus Brussels; Leuven Economics of Education (LEER); Email: lorenzo.navarini@kuleuven.be

‡KU Leuven, ECON Research Group, campus Brussels; Leuven Economics of Education (LEER);

1 Introduction

Supported by the overwhelming evidence regarding both the pecuniary and non-pecuniary returns for education (Oreopoulos and Salvanes, 2011; Heckman et al., 2018a, 2018b; Gunderson and Oreopolous, 2020), most developed countries have sought over recent decades to substantially increase the percentage of their population with a tertiary education degree. However, these benefits may be limited for a significant pool of graduates who start their careers in jobs that do not require a college degree (Groot and Maassen van den Brink, 2000; McGuinness, 2006; Verhaest and van der Velden, 2013; McGuinness et al., 2018). Indeed, these initially underemployed or so-called 'overeducated' graduates tend to earn a lower wage relative to adequately educated graduates who obtain similar degrees (Hartog, 2000; Barnichon and Zylberberg, 2019); plus, in terms of the non-pecuniary characteristics of their jobs, these graduates also seem to be worse off (Verhaest and Omey, 2009). To make matters worse, several studies also find that initial overeducation is persistent (Baert et al., 2013; Meroni and Vera-Toscano, 2017; Barnichon and Zylberberg, 2019) and that it leads to a greater probability of being unemployed later on (Sloane et al., 1999; Mavromaras et al., 2013).

Several explanations have been proposed as to why a portion of graduates are persistently overeducated and, as a result, may fail to fully capitalise on the potential benefits of higher education. One explanation is that overeducation is the result of search and matching frictions (Gautier et al., 2002; Dolado et al., 2009). Although this overeducation is often thought to be temporary, it may persist for several reasons, such as decreased on-the-job search (Holzer, 1987), locking-in effects due to job-specific human capital investments (Pissarides, 1994), negative signaling effects (McCormick, 1990), or a depreciation of underutilised skills (de Grip et al., 2008). According to this explanation, overeducation leads to heterogeneous realised (ex-post) returns to college and generates risk in the schooling decision (Leuven and Oosterbeek, 2011). Another suggested explanation is that overeducation is a consequence of heterogeneous skills across graduates. Indeed, many studies have found that overeducated workers score lower on ability tests or on their obtained GPA (Green et al., 2002; Agopsowicz et al., 2020), while others have argued that workers are overeducated without being overskilled (Allen and van der Velden, 2001; Chevalier, 2003; Green and McIntosh, 2007). Based on this explanation,

GLO; Email: dieter.verhaest@kuleuven.be

overeducation may thus also be a channel that generates the heterogeneity in expected (ex-ante) returns for college as found in several studies that do not focus on overeducation (see, e.g., Arcidiacono, 2004; Rodriguez et al., 2016).

A more controversial, but also relatively popular explanation, is that overeducation is the result of more general overinvestments in higher education (McGuinness, 2006; Leuven and Oosterbeek, 2011). Thus, employers may respond to an expansion of tertiary education by increasing their hiring requirements (Thurow, 1975; Charlot et al., 2005). In the longer run, however, labour markets are likely to generate more high-skilled vacancies in response (Gautier, 2002; Dolado et al., 2009; Ordine and Rose, 2017; Di Cintio et al., 2022). Moreover, as argued by Goldin and Katz (2008), technology has mostly been complementary to education over the last century[1]. And, according to the routinisation hypothesis (Autor et al., 2003; Goos et al., 2009), these technological advances have primarily served as substitutes for medium-skilled labour over recent decades, thus creating a polarised labour market. Given this observation, one may thus expect the attainment of a college degree to be just as effective a way to avoid overeducation. Indeed, a few descriptive studies conducted in the UK and Belgium have indicated that the probability of being overeducated is lower among the high-skilled than it is among the medium-skilled (Sloane et al., 1999; Verhaest and Omey, 2006)[2], while macro-level studies have typically failed to find a positive association between the share of highly skilled workers and the overeducation incidence (Verhaest and van der Velden, 2013; McGuinness et al., 2018; Delanay et al., 2020)[3]. However, it is unclear whether these findings are evidence of a causal link.

In this paper, we contribute to this discussion by investigating whether and how an increase in one's educational attainment affects one's likelihood of being overeducated and one's wage. To this end, we estimate a dynamic discrete choice model based on longitudinal data regarding young people's educational and early labour market careers

[1] Acemoglu (1998) claims that the increase in the number of high-skilled workers itself may have initiated technological advances that are complementary to their own employment.

[2] By looking at a large range of European countries, Lessear et al. (2015) found overeducation to be dominant among the medium-skilled workers in a few Southern European countries only. However, as the authors explain, this is likely due to the specific measure of overeducation (i.e. a so-called 'realised matches' measure) that was adopted. We revisit this point in the methods section.

[3] This conclusion often changes once the demand for high-skilled workers is controlled for in the analysis (e.g. Verhaest and van der Velden, 2013; Davia et al. 2017; Charalambidou and McIntosh, 2021). However, as argued, the high-skill job demand is likely to respond endogenously to the change in high-skill supply.

in Belgium. In this approach, career decisions are modelled as a sequence of choices that each depends on past decisions as well as on observed and unobserved characteristics (Cameron and Heckman, 2001; Carneiro et al., 2003; Heckman and Navarro, 2007; Heckman et al., 2016, 2018a, 2018b). Most of the individuals in our data entered the labour market between 1994 and 2003, a period for which the process of job polarisation is well documented (Goos et al., 2009). Furthermore, the Belgian case is particularly interesting because it combines a higher education system that is characterised by high levels of public subsidisation and low tuition fees with compulsory schooling until age 18. As a consequence, participation in higher education is relatively high and therefore very few young people enter the labour market without a higher secondary education degree.

Our modelling approach allows us to contribute in three main ways to the literature on educational participation, overeducation, and wages. First of all, it allows us to investigate the relationship between one's educational attainment, overeducation, and wages in a causal way. Not only is there a lack of evidence on the causal effect of educational attainment on overeducation, but the question of whether the wage penalty to overeducation presents a causal effect is still open to debate (Leuven and Oosterbeek, 2011). Several strategies have been adopted in order to address endogeneity problems. The first is to include ability-related test scores as controls in the wage equation (Chevalier and Lindley, 2009; Levels et al., 2014). Studies adopting this approach typically find that differences in skills explain only a small percentage of the penalty for overeducation. However, these test scores are unlikely to capture all unobserved differences that may matter in this context. Secondly, a few studies have used propensity score matching (McGuinness, 2008; McGuinness and Sloane, 2011) and in doing so concluded that overeducation affects wages negatively. However, whether the conditional independence assumption is fulfilled is dubious. A third strategy is to rely on fixed-effects panel data methods (Frenette, 2004; Dolton and Silles, 2008; Korpi and Tåhlin, 2009; Verhaest and Omey, 2012; Mavromaras et al., 2013). Generally speaking, this generates more mixed evidence on the importance of unobserved heterogeneity and produces estimates that may be biased due to endogenous job selection. One last strategy is to rely on instrumental variable regression, as is done by Korpi and Tåhlin (2009). However, as this strategy requires the use of valid instruments both for education and overeducation, adopting this method in this context is extremely challenging (Leuven and Oosterbeek, 2011). By exploiting the panel nature of our data and the sequentiality of choices, we are able to control for unobservable determinants in an alternative way.

Secondly, our modeling allows us to implement a new and also more comprehensive approach to gauge the importance of overeducation in explaining overall wage returns to education. The standard approach in the literature on overeducation and wages, introduced by Duncan and Hoffman (1981), is to replace years of education in the Mincer earnings equation with years of overeducation, years of required education, and years of undereducation. As the return to years of overeducation is usually found to be lower than that for years of required education, it is concluded that overeducation generates a wage penalty (Hartog, 2000; McGuinness, 2006; Leuven and Oosterbeek, 2011). However, these returns to years of overeducation and required education merely present returns conditional on one's match status and, therefore, do not take into account how one's match quality is affected by greater investment in education. We present a decomposition approach that attributes a part of the average unconditional wage return to education to a return in the case of perfect matching and another part to changes in match quality that may be induced by attaining more education. Furthermore, we show that this change in match quality may stem both from differences in the penalty to overeducation and differences in the likelihood of being overeducated across levels of education.

Thirdly, our modelling also enables us to investigate in greater detail whether overeducation is a channel that generates both heterogeneous expected and heterogeneous realised returns to college. Even if a percentage of the graduates are more likely to be overeducated due to lower skills levels, this should not imply that their return to college is negligible. Not only does the literature indicate that the wage return for college conditional on being overeducated is still positive (Hartog, 2000), there is also some evidence that employers prefer overeducated job seekers (Verhaest et al., 2018). Obtaining a college degree may therefore still improve one's ability to secure a medium-skilled job. By conditioning on both observable and unobservable characteristics in our model, we are able to investigate how differences in overeducation probabilities affect the full distribution of expected returns to college. Moreover, by simulating the matching process conditional on the estimated parameters of the model, we are also able to investigate how this matching affects the distribution of realised (ex-post) returns.

In line with the literature, we find that initial overeducation generates a persistent wage penalty. For instance, at age 23, this penalty is estimated to range from approximately 3% among those with a high-school or bachelor's degree to around 11% among those with a master's degree. However, the effect of overeducation risk on the expected return to college is found to be moderate at best with respect to obtaining a master's degree and even positive with respect to obtaining a bachelor's degree. This is partly due to an associated reduction in overeducation risk, probably as a consequence of job polarisation. Moreover, although we find that differences in overeducation probabilities reflect differences in expected (unconditional) wage returns across individuals, our results do not suggest that overeducation risk in and of itself reinforces this heterogeneity. However, we do find that overeducation risk generates substantial heterogeneity in realised (ex-post) returns to education. Overall, these results are far more consistent with overeducation being indicative of search and matching frictions rather than of considerable overinvestments in higher education.

The remainder of our paper is structured as follows. Section 2 introduces the decomposition approach that we use to analyse the role of overeducation risk in explaining the expected unconditional return to education. In Section 3, we describe the institutional setting of Flanders, the northern Dutch-speaking region of Belgium. Section 4 introduces the dataset and the measurement of our key variables. In Section 5, we outline our dynamic discrete choice model. In Section 6, we present the counterfactual simulation and the relative results of the treatment effects and the heterogeneous treatment effects. Finally, in Section 7, we discuss these results and conclude our paper.

2 Conceptual framework

In this section, we develop a conceptual framework to demonstrate how overeducation may both affect the average unconditional wage return to education and generate heterogeneous unconditional wage returns among individuals with the same level of educational attainment.

First, let us presume that the educational and labour market outcomes of an individual i can be summarised by means of the following three equations:

$$e_i = f(X_i, \varepsilon_i^e) \tag{1}$$

$$oe_i = g(e_i, X_i, \varepsilon_i^{oe}) \tag{2}$$

$$w_{a,i} = h(e_i, oe_i, X_i, \varepsilon_i^{w_a}) \tag{3}$$

with equation (1) reflecting the individual's educational attainment e_i and being a reduced-form equation of a more extended model of human capital accumulation, equation (2) determining the individual's overeducation status oe_i once they leave the educational system and enter the labour market (modelled as a binary outcome with $oe_i = 1$ when overeducated and $oe_i = 0$ when adequately qualified for the job) and, finally, equation (3) reflecting one's subsequent wage $w_{a,i}$ at age a. Additionally, we presume each of these three outcomes o to be a function of a similar set of exogenous characteristics and factors X_i (e.g. family background, gender, abilities, preferences, labour market conditions,...)[4]. Finally, each outcome is also presumed to depend on outcome-specific residual determinants ε_i^o that are modelled to be independent of one's characteristics X_i and one's prior endogenous outcomes. These residual determinants may, for instance, include outcome-specific preference shocks or, in the case of overeducation, random shocks due to search and matching frictions.

By substituting equation (2) in (3), we may now rewrite the wage as a function of educational attainment e_i, exogenous characteristics X_i and residual determinants $\varepsilon_i^{oe,a}$:

$$w_{ai} = h\Bigg(e_i, oe_i(e_i, X_i, \varepsilon_i^{oe}), X_i, \varepsilon_i^{a}\Bigg) = k(e_i, X_i, \varepsilon_i^{oe,a}) \tag{4}$$

With equation (4), we estimate the effect of educational attainment on wages unconditional of one's overeducation status. Henceforth, this allows us to identify the unconditional (total) wage return to education:

$$\frac{dw_{ai}}{de_i} = \frac{\partial w_{ai}}{\partial e_i} + \frac{\partial w_{ai}}{\partial oe_i}\frac{doe_i}{de_i} \tag{5}$$

where the first term in the right-hand side of the equation represents the direct effect of educational attainment on wages and the second term represents the indirect effect of educational attainment through its effect on overeducation. This indirect effect already

[4] To simplify the notation, we make the assumption that X_i is time-invariant. In a more extended version of the model (as is estimated in our paper), one can differentiate between common time-invariant exogenous factors and exogenous factors that are time-variant (e.g. labour market conditions at the moment of the outcome).

provides a first channel through which the presence of overeducation may affect one's wage return to education. However, as will be argued later in this section, the direct effect may also be affected by the presence of overeducation.

While equation (5) defines the return of an infinitesimally small change in the level of educational attainment, it is more natural to evaluate the return to more specific, discrete levels of educational attainment. With respect to educational attainment j, this unconditional return at age a for individual i with exogenous characteristics X_i may be defined as follows:

$$\Omega_{a,i,j} = \mathbb{E}\left[w_{a,i}|X_i, e_i = j\right] - \mathbb{E}\left[w_{a,i}|X_i, e_i = j-1\right] \tag{6}$$

with $\mathbb{E}\left[w_{a,i}\right]$ being the expected wage at age a. This unconditional total return is thus simply the difference between the expected wage given her obtained educational attainment j and the expected wage when the individual's educational attainment would have been $j-1$ only.

Rather than focusing on this unconditional return, the overeducation literature typically looks at the wage return conditional on one's match status. Depending on the overeducation status of individual i at educational level j, and one's status at the preceding level $j-1$, we can define the following four conditional wage returns accordingly:

$$\Omega^{M,M}_{a,i,j} = \mathbb{E}\left[w_{a,i}|X_i, e_i = j, oe_i = 0\right] - \mathbb{E}\left[w_{a,i}|X_i, e_i = j-1, oe_i = 0\right] \tag{7}$$

$$\Omega^{M,O}_{a,i,j} = \mathbb{E}\left[w_{a,i}|X_i, e_i = j, oe_i = 1\right] - \mathbb{E}\left[w_{a,i}|X_i, e_i = j-1, oe_i = 0\right] \tag{8}$$

$$\Omega^{O,M}_{a,i,j} = \mathbb{E}\left[w_{a,i}|X_i, e_i = j, oe_i = 0\right] - \mathbb{E}\left[w_{a,i}|X_i, e_i = j-1, oe_i = 1\right] \tag{9}$$

$$\Omega^{O,O}_{a,i,j} = \mathbb{E}\left[w_{a,i}|X_i, e_i = j, oe_i = 1\right] - \mathbb{E}\left[w_{a,i}|X_i, e_i = j-1, oe_i = 1\right] \tag{10}$$

where $\mathbb{E}\left[w_{a,i}|X_i, e_i = j, oe_i = 0,1\right]$ is the expected wage at educational level j when the individual i is either overeducated ($oe_i = 1$) or adequately matched ($oe_i = 0$).

While equation (7) describes the return to education presuming one would be adequately matched regardless of one's level of educational attainment ($\Omega^{M,M}_{a,i,j}$), equation (8) reflects the return to education when completing more education induces one's match status to switch from an adequate match to overeducation ($\Omega^{M,O}_{a,i,j}$). These two types of conditional returns are equivalent to the two types of returns that are typically reported in the literature on overeducation: the return to (years of) required education and the return to (years of) overeducation, respectively. Moreover, by subtracting the return to required education from the return to overeducation, we obtain the so-called wage penalty to overeducation that is frequently reported in the literature as well:

$$\psi_{a,i,j} = \Omega^{M,O}_{a,i,j} - \Omega^{M,M}_{a,i,j} = \mathbb{E}[w_{a,i}|X_i, e_i = j, oe_i = 1] - \mathbb{E}[w_{a,i}|X_i, e_i = j, oe_i = 0] \tag{11}$$

As shown in equation (6), this wage penalty of overeducation for educational attainment j is also equal to the difference in the expected wage while being overeducated and the expected wage while being adequately matched for educational attainment j.

The statistic of the wage penalty to overeducation is, along with the proportion of overeducated individuals, often used to gauge the importance of overeducation in reducing the wage return to education. However, without having completed more education, some may have been overeducated while others may even manage to improve their match status by completing more education. Henceforth, conditional returns $\Omega^{O,O}_{a,i,j}$ and $\Omega^{O,M}_{a,i,j}$ must be weighted in as well when assessing the importance of overeducation in explaining unconditional returns to education.

To assess more explicitly how important overeducation is in explaining the unconditional return, we implement a decomposition approach to this return. To this end, we first rewrite the expected wage at level of educational attainment j as a weighted average of the conditional wage when adequately matched and the conditional wage when overeducated:

$$\begin{aligned} &\mathbb{E}\left[w_{a,i}|X_i, e_i = j\right] = \\ (1 - P^{OE}_{i,j})\, \mathbb{E}\left[w_{a,i}|X_i, e_i = j, oe_i = 0\right] &+ P^{OE}_{i,j}\, \mathbb{E}\left[w_{a,i}|X_i, e_i = j, oe_i = 1\right] \end{aligned} \tag{12}$$

where $P^{OE}_{i,j}$ is individual i's probability of being overeducated when having obtained level of educational attainment j. Moreover, by using equation (11), we can rewrite equation (12) in the following way:

$$\mathbb{E}\left[w_{a,i}|X_i, e_i = j\right] = \mathbb{E}\left[w_{a,i}|X_i, e_i = j, oe_i = 0\right] + P^{OE}_{i,j}\psi_{a,i,j} \tag{13}$$

In addition, by adopting the same logic for the expected wage at level of educational attainment $j-1$, and by using equations (6) and (7), we obtain the following alternative formula for the unconditional wage return:

$$\Omega_{a,i,j} = \Omega^{M,M}_{a,i,j} + P^{OE}_{i,j}\psi_{a,i,j} - P^{OE}_{i,j-1}\psi_{a,i,j-1} \tag{14}$$

Finally, by once more adding and subtracting the term $P^{OE}_{i,j-1}\psi_{a,i,j}$ to the right-hand side of equation (14), we can decompose the unconditional wage return to education in the following three subcomponents:

$$\Omega_{a,i,j} = \underbrace{\Omega^{M,M}_{a,i,j}}_{\text{(A)}} + \underbrace{P^{OE}_{i,j-1}(\psi_{a,i,j} - \psi_{a,i,j-1})}_{\text{(B)}} + \underbrace{(P^{OE}_{i,j} - P^{OE}_{i,j-1})\psi_{a,i,j}}_{\text{(C)}} \tag{15}$$

where (A) represents the return that may be realised in case of perfect matching, (B) is a subcomponent that is attributed to the potential difference in overeducation penalty between educational attainment j and $j-1$, and (C) is a subcomponent that is attributed to the potential difference in overeducation risk between these two levels of attainment. The latter subcomponent is also equivalent to the indirect effect of education on wages as defined by equation (5).

Interestingly, the unconditional return collapses to component (A) when the expected match quality is identical across all levels of education. The sum of subcomponents (B) and (C) in equation (15), meanwhile, measures the contribution of any change in expected match quality that may be induced by investing in a higher level of education. Moreover, as overeducation does not affect the unconditional wage return in any other way than through (B+C), this sum therefore also serves as a reasonable measure of the importance of overeducation in explaining the unconditional return to education. Additionally, given that this component is merely driven by the difference in overeducation penalties and overeducation probabilities across levels of educational attainment, it is apparent that a focus on absolute overeducation penalties and probabilities may lead to misleading interferences regarding the importance of overeducation in this respect.

Our decomposition of the expected unconditional return may be implemented both for the average unconditional return and for the distribution of unconditional returns within a (sub-)population that is expected given the individuals' characteristics X_i. For instance, due to differences in innate abilities, individuals may differ in their overeducation risk (cf. equation (2)) and, therefore, also in their expected unconditional return. This distribution of expected unconditional returns is based on the assumption that one's individual overeducation status is not precisely known (i.e. it is the expected return prior to when the matching to a first job occurred). However, due to random shocks in overeducation (i.e. $\varepsilon_i^{oe} \neq 0$) (but also in wages), this distribution will deviate from the distribution of returns that are realised in practice. For instance, even if one's likelihood of being overeducated is small, search and matching frictions may still cause one to experience bad luck. Hence, to gauge the extent to which overeducation contributes to heterogeneous realised returns as well, one may also calculate the distribution of returns while presuming one's first match status is already known. This comes down to assigning to each individual one out of the four considered conditional returns (i.e. $\Omega_{a,i,j}^{M,M}$, $\Omega_{a,i,j}^{M,O}$, $\Omega_{a,i,j}^{O,M}$, and $\Omega_{a,i,j}^{O,O}$), based on a random draw of the individual's match probability distributions at each level of educational attainment. We will call this the distribution of returns conditional on random matching.

3 Institutional setting

To carry out estimations using our dynamic discrete choice model and to apply our decomposition approach, we rely on data related to the educational and early labour market careers of young individuals in Flanders, the Northern Dutch-speaking region of Belgium. In Flanders, compulsory education starts from September 1^{st} of the year in which the child turns 6 until their 18^{th} birthday or until June 30^{th} of the year in which the child turns 18. Primary education usually starts at the age of 6 and consists of 6 consecutive grades. Subsequently, at the age of 12 in the case of no delay, pupils enter secondary education. Secondary education consists of four tracks, namely the general track, the technical track, the art track, and the vocational track, with the technical or art tracks being introduced from grade 9 (i.e. the 3^{rd} grade in secondary education) onwards. Between the subsequent grades, students may downgrade from the general to the technical or art tracks, or from the technical or art tracks to the vocational track. From age 15 onwards, students may also opt for a part-time vocational track that may be combined with three to four days of apprenticeship training in a firm. After passing 6 grades in the general, technical, or art tracks, or 7 grades in the (full-time) vocational track, individuals may enter tertiary education without having to complete any entrance exam (except for medicine) or overcome any other entry barriers.

In the period before the Bologna reform, which is the period that is relevant for our sample, individuals entering tertiary education were able to choose between (i) a short-term programme at a vocationally-oriented college (called 'hogeschool' in Dutch), (ii) a long-term programme at such a college, or (iii) a more academically-oriented long-term programme at a university. While the short-term programmes lasted three years, the long-term programmes lasted four years or more. Moreover, the long-term programmes were subdivided into two stages with the first stage taking two or three years and leading to a so-called 'candidate' degree, although it was relatively rare for someone to leave tertiary education who managed to pass this stage. Since the Bologna reform, students have been able to undertake a so-called professional bachelor degree at a vocationally-oriented college and an academic bachelor degree at a university, with the latter providing direct access to an academic master's degree. Moreover, students may also start in an academic master's programme after having obtained a professional bachelor degree conditional on participating first in a bridging programme that usually takes one year. Both types of bachelor programmes last three years, while the length of a master's programme is at least one year. By law, the old short-term and long-term degrees have been declared to be equivalent to these new bachelor and master's degrees.

4 Data

4.1 Sample

Our model is estimated using the SONAR data. These data include representative samples of three cohorts (birth years 1976, 1978, and 1980) of approximately 3,000 individuals in Flanders who were surveyed for the first time when they were 23 years old. Moreover, these original surveys were supplemented with a number of follow-up surveys, completed at age 26 for the 1976 and 1978 cohorts and at age 29 for the 1976 and 1980 cohorts (the response rates are between 60% and 70%). The data include detailed information regarding schooling and labour market outcomes, which are gleaned by recording each educational choice from the age of 6 onwards and registering core information on one's labour market history on a monthly basis. In addition, the dataset includes a large set of indicators related to family background as well as information on one's overeducation status and wages both measured at the start of the first job as well as at the moment of the various surveys (ages 23, 26, and 29). To ensure the estimated model remains tractable, we remove from the initial sample those individuals (i) who experienced more than one year of delay at the start of their primary education (76 individuals) and (ii) those who have special needs that are catered for in schools providing special care (124 individuals). Moreover, we remove another 638 individuals with (iii) inconsistent, erroneous, or incomplete data regarding the exogenous variables (cf. infra) and their educational career. This leaves us with a final sample of 8,162 individuals, which is used to estimate the equations related to the educational outcomes.

4.2 Exogenous variables

At each stage of our model, we control for the following exogenous background characteristics of the individual: gender (one dummy), foreign origin (one dummy), years of education of the mother and the father (beyond the phase of primary education), number of siblings, year of birth, and day of birth within the calendar year. Most of these variables are standard background characteristics that are frequently included in dynamic discrete choice models on educational careers (e.g. Cameron and Heckman, 2001; Belzil and Poinas, 2010; Heckman et al., 2016, 2018a, 2018b; Baert et al., 2022). In addition, we control for the unemployment rate at the district level to account for differences in labour market conditions. This is a time-varying variable that is measured at the moment of each outcome. Table 2 includes descriptive statistics on each of these exogenous variables.

4.3 Educational attainment and track choices

Our dynamic model, which is a more extended version of the model introduced in Section 2 (equations (1)–(3)) and outlined in more detail in Section 5, includes in total 17 sequential outcomes related to the educational and early labour market career of the individuals. With respect to the educational career (cf. equation (1)), these outcomes include the delay at the start of primary and secondary education along with the enrolment, track choice, and attainment related to the following four critical stages of secondary and tertiary education: (i) lower secondary education, (ii) higher secondary education, (iii) lower tertiary education, and (iv) higher tertiary education. These four stages, along with their acronyms and relation to the ISCED classification, are summarised in Table 1.

Table 1: Codification of educational attainment

Code name	Description	ISCED Code	Freq.	Percent	Cum.
-	Less than Lower Secondary Education	ISCED 0 and 1	375	4.59	4.59
LSE	Lower Secondary Education	ISCED 2	551	6.75	11.35
HSE	Higher Secondary Education	ISCED 3 and 4	3,266	40.01	51.36
LTE	Lower Tertiary Education	ISCED 5 - Bachelor	2,390	29.28	80.64
HTE	Higher Tertiary Education	ISCED 5 - Master	1,580	19.36	100

We define individuals as having attained lower secondary education (LSE) if they have completed at least the fourth grade of secondary education, while higher secondary education (HSE) attainment is defined as having completed six grades of secondary education. Regarding tertiary education, lower tertiary education (LTE) attainment is defined by the completion of a short-term college degree or completion of at least the third grade of a long-term college or university degree. Although in the pre-Bologna system, many long-term programmes awarded a candidate qualification after just two grades, these qualifications are usually not considered to be equivalent to a bachelor's degree. By setting the bar at passing at least three grades at university, we follow the logic of the current system to obtain a bachelor's degree at university. Finally, those that have fully completed their long-term college or university degree (i.e. after four or more grades of tertiary education) are defined as having attained higher tertiary education (HTE). The latter level of educational attainment is equivalent to a master's degree in the current system.

Enrolment in these four stages is defined as having enrolled in the third grade of secondary education (enrolment LSE), the fifth grade of secondary education (enrolment HSE), the first grade of tertiary education (enrolment LTE), and the fourth grade of tertiary education (enrolment HTE), respectively. Strictly speaking, individuals already enrol in lower secondary education from the first grade of secondary education onward. However, as this is the case for (almost) all individuals in our dataset, we adjust the definition towards enrolment in the third grade. The track choice refers to the (first) year of enrolment in each stage and distinguishes between the general track (in secondary education) or academic track (in tertiary education) and other tracks. The academic track in tertiary education is defined as including all programmes at university, while the non-academic track includes programmes at (more vocationally-oriented) colleges.

Table 2: Descriptive statistics

	(1)		(2)		(3)		(4)	
	Full Sample		Sample labour market outcomes		Adequately matched first job		Overeducated first job	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
A. *Exogenous variables*:								
Female	0.494	0.500			0.507	0.500	0.475	0.499
Foreign origin	0.056	0.231			0.051	0.220	0.057	0.232
Education Mother	5.738	3.437			5.729	3.408	5.362	3.353
Education Father	6.217	3.675			6.247	3.607	5.625	3.552
Number of siblings	1.669	1.422			1.663	1.377	1.662	1.506
Cohort 1978	0.338	0.473			0.329	0.470	0.338	0.473
Cohort 1980	0.345	0.475			0.325	0.468	0.373	0.484
Birthday date/100	1.718	1.002			1.711	0.999	1.731	1.019
B. *Endogenous variables*:								
B.1. Schooling outcomes								
Delay Primary School	0.015	0.123			0.014	0.118	0.015	0.122
Delay Secondary School	0.101	0.302			0.102	0.302	0.108	0.310
Lower secondary education: enrolment	0.991	0.095			0.987	0.114	0.997	0.056
Lower secondary education: general track	0.524	0.499			0.530	0.499	0.440	0.496
Lower secondary education: qualification obtained	0.954	0.209			0.931	0.253	0.987	0.113
Higher secondary education: enrolment	0.938	0.242			0.912	0.283	0.972	0.164
Higher secondary education: general track	0.442	0.497			0.445	0.497	0.353	0.478
Higher secondary education: qualification obtained	0.887	0.317			0.847	0.360	0.944	0.230
Lower tertiary education: enrolment	0.639	0.480			0.645	0.479	0.561	0.496
Lower tertiary education: academic track	0.214	0.410			0.205	0.404	0.151	0.358
Lower tertiary education: qualification obtained	0.486	0.500			0.527	0.499	0.346	0.476
Higher tertiary education: enrolment	0.215	0.411			0.200	0.400	0.165	0.371
Higher tertiary education: academic track	0.146	0.353			0.140	0.347	0.097	0.296
Higher tertiary education: qualification obtained	0.194	0.395			0.195	0.396	0.162	0.368
B.2. Labour market outcomes								
Overeducation first job			0.351	0.477	0.000	0.000	1.000	0.000
Wage selection at age 23			0.580	0.492	0.561	0.496	0.626	0.484
Hourly Wage at age 23			7.342	1.590	7.445	1.572	7.176	1.606
Wage selection at age 26			0.450	0.497	0.476	0.499	0.403	0.491
Hourly Wage at age 26			8.113	1.866	8.210	1.850	7.909	1.887
Wage selection at age 29			0.423	0.494	0.434	0.496	0.405	0.491
Hourly Wage at age 29			8.546	1.829	8.670	1.843	8.306	1.782
Observations	8162		7211		4648		2563	

Table 2 (Column (1)) also includes the descriptive statistics for each of these outcomes. Almost all individuals enrol in (99.1 percent) and attain (95.8 percent) a lower secondary education. In regard to higher secondary education, this slightly drops to 93.8 and 88.7 percent, respectively. When transiting to lower tertiary education, the drop is more substantial, with an enrolment rate of 63.9 percent and an attainment rate of 48.6 percent. Finally, 21.5 and 19.4 percent of the overall sample enrols in and attains a higher tertiary education. Only a small minority of the sample (11.3 percent) can thus be categorised as low-skilled (i.e. less than HS), while the medium- (HS degree) and high-skilled (at least a LT degree) represent 42.1 and 48.6 percent of the sample, respectively (see also Table 1). Regarding the track choice, the general or academic track is relatively more frequently chosen at the LSE and HTE stages, while the other (more vocational) tracks are more dominant at the HSE and LTE stages.

4.4 Overeducation

The main outcome of interest in our model is overeducation (cf. equation (2)), which is defined as having attained a level of education that is above the level of education that is required to perform one's job well. We focus on the overeducation status at the start of the first job with a standard labour contract, which excludes internships, apprenticeships, or student work. For the estimation of the equation related to the overeducation status in the first jobs, the sample is further reduced to 7,211 individuals. This is because there are 701 individuals for whom we have no data regarding a first job (either because they did not participate in the follow-up survey(s) or because they did not have a first job by age 29) and another 250 for whom the data on overeducation is missing (see Appendix A).

To measure overeducation, the literature has adopted a wide range of methods that can be subdivided into four broad categories (McGuinness, 2006; Verhaest and Omey, 2006; Leuven and Oosterbeek, 2011): (i) job analysis, (ii) direct self-assessment, (iii) indirect self-assessment, and (iv) realised matches methods. Job analysis methods are usually based on occupational classifications that define the required level of education based on the assessment of job experts. Self-assessment methods, meanwhile, rely on the assessment of the worker themselves, either by asking directly whether he or she is overeducated or indirectly by enquiring about the required level of education to carry out their job or to be hired for their job. Finally, realised matches methods rather measure the required level of education by the average or modal level of education within one's occupation.

Each of these methods has a number of disadvantages. Job analysis and realised matches methods, for instance, may insufficiently account for the heterogeneity of requirements within job categories with the same occupational title. Moreover, while the job analysis method requires frequent updates of the requirements to account for technological changes, job requirements measured by realised matches methods may be largely endogenous to the composition of the labour force in terms of their educational attainment. Finally, self-assessment measures are likely to be vulnerable to a wide range of cognitive biases. For instance, due to lack of expertise in this respect, individuals may find it difficult to gauge the true requirements of their job. Moreover, even if they manage to gauge these requirements correctly, they may answer in a socially desirable way and thus inflate their own status.

Given that our data enables us to measure overeducation based on these methods, we are able to circumvent these problems at least partially. In particular, we define individuals as being overeducated if they are classified as such based on at least two out of three deliberately chosen measures. The first measure adopts a job analysis approach. In our data, jobs have been coded based on the Standard Occupation Classification of Statistics Netherlands[5]. The classification groups jobs based on five functional levels, where each level represents one of the five considered levels of education in our model (cf. Table 1) that a worker ideally has to properly perform the tasks in the job. A comparison of these job requirements with one's level of education thus determines one's overeducation status. As shown in Table 3, using only this information, 52% of individuals are considered to be overeducated for their first job.

We complement this job analysis approach with information from one direct and another more indirect (but modified) self-assessment measure. The direct self-assessment measure is derived from the following survey question: 'According to your own opinion, do you have a level of education that is too high, too low or appropriate for your job?' When using this measure, 21.5% of the individuals are considered to be overeducated for their first job (Table 3). The indirect self-assessment measure is constructed based on

[5]Link to the dataset: Dutch Standard Classification of Occupations (SBC) 1992 (Last accessed: 15.02.2023)

the following survey question: 'What is (was), according to your own opinion, the most appropriate educational level to execute your first job?' As this question was not included in the survey for the 1976 cohort, we implement a modified procedure following Baert et al. (2013). First, we calculate the median self-assessed required level within each occupation based on the available information. To do this, we rely on the aforementioned five categories of education levels (cf. Table 1). Second, we extrapolate this median to all jobs in each occupation. Third, we class an individual as being overeducated if their attained level of education exceeds this median required level within their occupation. Using this procedure, 35.2% of the individuals are found to be overeducated for their first job.

Our choice to combine information on these three measures is based on three main arguments. First, we consider these measures to be the most closely connected with the concept of overeducation as defined in the literature. Given that hiring requirements may deviate from what is truly needed to do a job, this is less the case for self-assessment measures based on what is required to get the job or realised matches measures that mimic actual hiring behaviour. Second, given our focus on educational attainment and overeducation, we avoid including measures that are endogenous to the educational composition of the workforce. Third, these three measures result from three relatively independent assessments. While our first measure relies on the assessment of job experts and the second is based on the assessment of the worker themselves, our third measure can be interpreted as reflecting the opinions of one's co-workers within one's occupation. Hence, errors that are systematic across two or three of these measures are expected to be limited, while defining someone as overeducated if classified as such based on two out of three of these assessments should go a long way in accounting for other non-systematic errors.

Based on the combination of these three measures, we find 35.5% of the sample to be overeducated for their first job. In line with the idea that overeducated workers are a non-random sample of the population, we find that more often than not they are male and their parents are somewhat lower educated (Table 2 (Panel (A)). Moreover, in line with polarisation, we find them to be more likely to have obtained at least a higher secondary education degree and less likely to have also obtained a lower or higher tertiary education degree (Table 2 (Panel (B1)).

Table 3: Different measures of overeducation

Variable	Mean	SD
JA	0.520	0.499
DSA	0.215	0.411
ISA	0.352	0.478
BM	0.355	0.479

Notes: The overeducation measures are: Job Analysis (JA), Direct Self-Assessment (DSA), Indirect Self-Assessment (ISA) and Benchmark Measure (BM).

4.5 Wages

To maintain the sequentiality of our model, which is an important precondition to identify causal effects in dynamic discrete choice models, we analyse the wages at ages 23, 26, and 29 rather than at the start of the first job (cf. equation (3)). As a consequence, the estimated wage effects of overeducation in our model are to be interpreted as reduced-form effects that result from, among other things, its effect on one's later mismatch status. As shown by Baert et al. (2013), based on a subsample of the same SONAR data, overeducation is strongly persistent. Thus, if overeducation has a contemporaneous effect on wages, as is usually found to be the case in the literature, we can expect it to affect future wages as well. Moreover, this would also be consistent with a few studies on other countries that found that overeducated workers experience no more wage growth than other workers (Büchel and Mertens, 2004; Korpi and Tåhlin, 2009)[6].

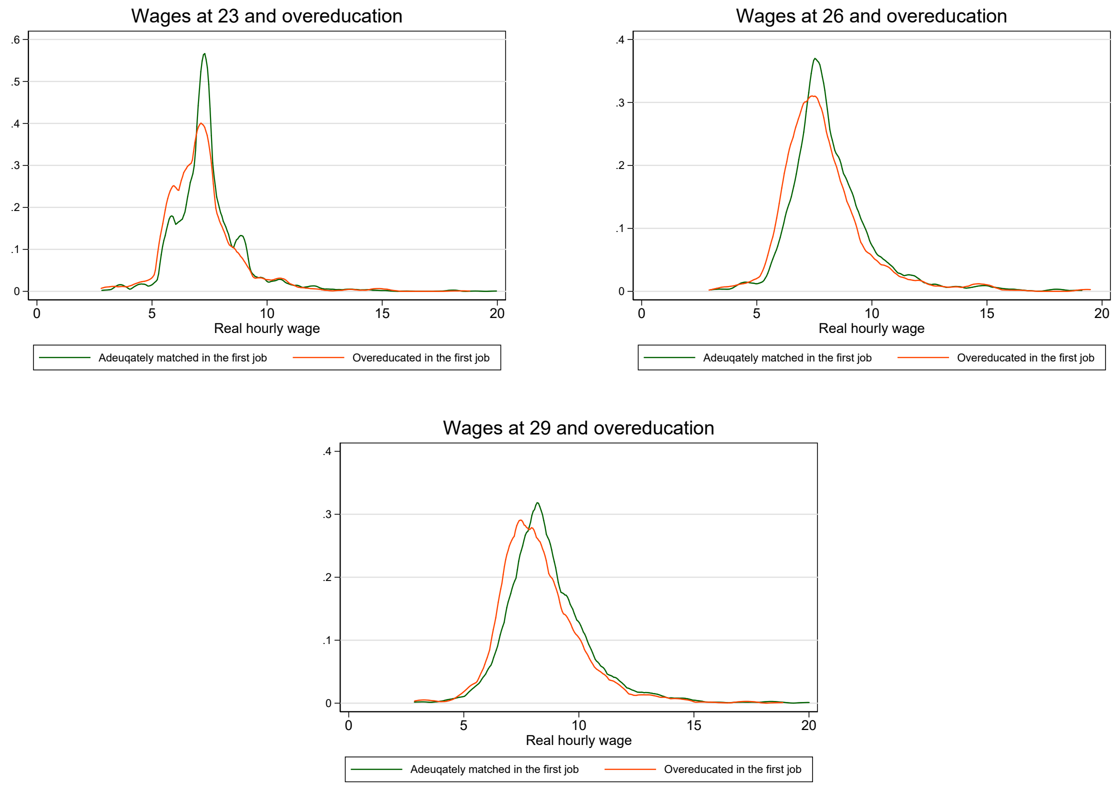
Respondents reported their official net monthly wage. While this was reported in intervals in the first survey of the first cohort, exact wages were reported in later surveys (if respondents refused to answer, they still had the option to report in intervals). For our analysis, we alter these reports in such a way as to log real hourly net wages (we rely on the midpoint for the interval reports). Due to missing data, the number of observations in the wage equations drops to 4,407, 3,379, and 3,142 for the ages 23, 26, and 29, respectively. With respect to age 23, data is missing for two main reasons. First, a significant proportion of the individuals were still in education or without jobs at this age. Second, even if employed, not all individuals were queried about their wage at age

[6] See Rubb (2006) or Roller et al. (2020) for contrasting findings.

23. In particular, for the 1978 and 1980 cohorts, those who were still in a first job that had started within the last year were precluded from answering these questions, while for the 1976 cohort, none of the individuals who were still in their first job (irrespective of when it started) were asked to indicate their wage. With respect to age 26 and age 29, meanwhile, missing data on wages are primarily caused by a lack of surveying (for the 1978 cohort at age 26, and for the 1980 cohort at age 29) or due to attrition. Missing data due to respondents' refusal to answer or because of wage outliers are less important for each of the three points of measurement. As these missing data are unlikely to be random, we account for this in our analysis by adding three selection equations to our model (cf. infra).

Figure 1 shows the wage distribution at each age depending on the match status in the first job. In line with initial overeducation having a persistent effect on wages, the wage distribution of the overeducated workers is each time positioned to the left of the wage distribution of adequately matched individuals. Based on our model, we will assess whether this difference truly reflects a causal effect of overeducation.

Figure 1: Distribution of wages by overeducation status



5 Econometric strategy

In this section, we present our dynamic multistage model of educational choices and labour market outcomes. This is used to identify the impact of educational attainment on overeducation and its consequences on future wages by controlling for dynamic selection and unobserved heterogeneity.

5.1 Dynamic discrete choice model

We adopt a dynamic treatment effects approach[7] (Heckman and Navarro, 2007; Heckman et al., 2016), following the seminal papers of Cameron and Heckman (1998, 2001) and being applied and refined by, among others, Colding (2006), Belzil and Poinas (2010), Adda et al. (2010), Baert and Cockx (2013), Baert et al. (2017), De Groote (2018), Declercq and Verboven (2018), Heckman et al. (2018a, 2018b), Cockx et al. (2019), and Neyt et al. (2022). These dynamic models are characterised by a sequential structure of binary and ordered logit functions, with each choice opening up the possibility of performing particular future choices. This sequential structure is consistent with the organisation of the educational system, whereby obtaining access to a particular stage (e.g. tertiary education) is conditional on having succeeded in the previous stage (e.g. obtaining a higher secondary education qualification).

Our approach is a methodological middle-ground between the reduced-form treatment effect approach and the more structural dynamic discrete choice model approach: while agents are presumed to make choices and account for the consequences of these choices, as is the case in a fully structural approach, we do not need to explicitly identify and model the rules driving these choices, as in a reduced-form approach (Heckman and Navarro, 2007; Heckman et al., 2016, 2018a, 2018b). Hence, while it is not possible to estimate ex-ante individual valuations or expectations, our model leaves the door open to a broader set of explanations regarding what drives these choices than just perfectly forward-looking behaviour (Heckman and Navarro, 2007; Belzil and Poinas, 2010; Heckman et al., 2018a, 2018b). Another major advantage of this approach is that it does not require us to impose assumptions on the functional forms or distribution of the unobservables (Heckman et al.,

[7] In the previous literature, this approach is defined using a wide range of names: quasi-structural, semi-structural, quasi-reduced form, and black box approach, among others (Colding, 2006; Belzil and Poinas, 2010). We adopt the definition of dynamic treatment effects, as in Heckman and Navarro (2007) and Heckman et al. (2016, 2018a, 2018b).

2018a, 2018b). Moreover, it enables us to decompose the treatment effects into both direct and total effects associated with later educational choices (Heckman and Navarro, 2007; Heckman et al., 2016, 2018a, 2018b).

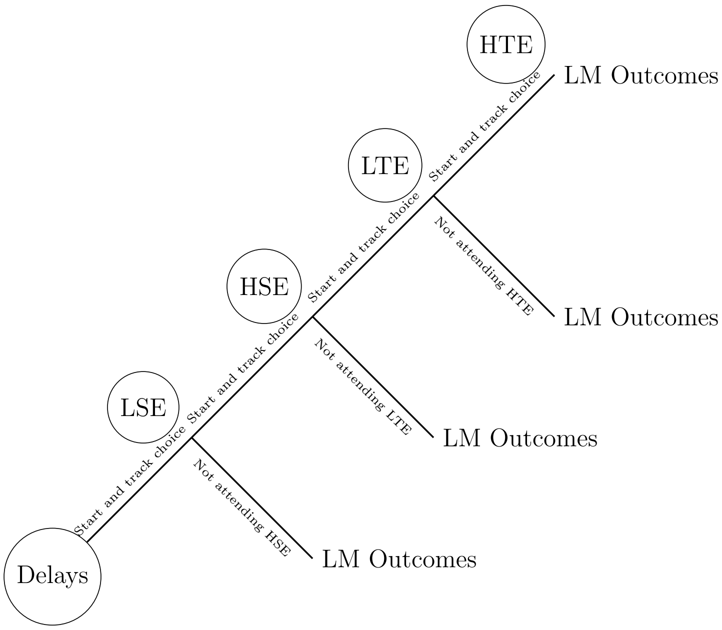
Our model, which is a more extended version of the model introduced in Section 2, is designed to capture the dynamic relationship between schooling choices, human capital formation, and labour market outcomes for each individual i. In line with our conceptual framework, we have three main choice and outcome sets to model: (i) educational choices and the process of attaining education (e_i), (ii) whether the individual is overeducated for their first job upon labour market entry (oe_i) and (iii) the realised subsequent wages at a specific age a ($w_{a,i}$).

In the first choice set, we consider a total of 10 sequential educational choices that individuals make from the age of 6 onwards. First, we include (i) the delay at the start of primary education and (ii) the delay at the start of secondary education. Next, we model the enrolment, track choice, and attainment with respect to each of the four considered stages in secondary and tertiary education: (iii) enrolment and track choice at the start of lower secondary education, (iv) lower secondary education attainment, (v) enrolment and track choice at the start of higher secondary education, (vi) higher secondary education attainment, (vii) enrolment and track choice at the start of lower tertiary education, (viii) lower tertiary education attainment, (ix) enrolment and track choice at the start of higher tertiary education, and (x) higher tertiary education attainment.

Enrolment and track choice at the start of each of the four stages of one's educational career are modelled as one and the same choice to preserve the sequentiality of the model. To this end, we rely on an ordered logit specification, with outcome value 2 indicating enrolment in the general or academic track, outcome value 1 indicating enrolment in another track, and outcome value 0 indicating no enrolment. Furthermore, as illustrated in Figure 2, access to each of the choices related to these four educational career stages (choices iii to x) is presumed to be conditional on the preceding educational choice: being able to obtain a degree at level j is conditional on having enroled and chosen a track at the same level, while being able to enrol and choose a track at level j is conditional on obtaining a degree at level $j-1$.

With respect to the second and third choice and outcome sets, we model a total of seven labour market outcomes: (xi) overeducation at the start of the first job, (xii and

Figure 2: A sequential dynamic model



Notes: Delays include both Delay in Primary Education and Delay in Secondary Education. Lower Secondary Education (LSE), Higher Secondary Education (HSE), Lower Tertiary Education (LTE) and Higher Tertiary Education (HTE) includes start, track choice and acquiring a degree. Lastly, Labour Market (LM) Outcomes include overeducation for the first job, wage selection, and log-wage equation for ages 23, 26, and 29.

xiii) a wage selection equation and the wages at age 23, (xiv and xv) a wage selection equation and the wages at age 26, and, finally, (xvi and xvii) a wage selection equation and the wages at age 29.

The full sequence of outcomes is represented by O and defined as $O = \{1, ..., S\}$, where each number corresponds to an outcome o, $o \in O$, with a total of S steps. For the sake of clarity, we denote $o = 11$ as $o = oe$, the overeducation outcome, and $o = \{13, 15, 17\}$ as $o = w_a$ for $a \in \{23, 26, 29\}$, the wage outcome measured at age a.

The optimal choice $\hat{c}_i^o$ of an individual i with respect to binary or ordered outcome $o \notin \{13, 15, 17\}$ is:

$$\hat{c}_i^o = c \in C^o \text{ if } \omega_c^o < U_{i,c}^o \leq \omega_{c+1}^o \tag{16}$$

where $U_{i,c}^o$ is the latent utility of choice c for outcome o, and ω_c^o and ω_{c+1}^o are threshold utilities (*cut-off values*) that determine the ordered choice.

In line with our framework, we presume these utilities to be determined by a vector of preceding endogenous choices V_i^o, a vector of exogenous characteristics X_i^o, as well as outcome-specific shocks $\varepsilon_{i,c}^o$ that are independent of the other exogenous and endogenous determinants. In the context of our dynamic model, we further subdivide the set of exogenous characteristics X_i^o into a set of invariant observed characteristics Z_i, variant observed characteristics R_i^o such as local labour market conditions, as well as unobserved exogenous determinants that are correlated with (some of) the preceding endogenous choices (abilities, motivations, preferences). We thus approximate $U_{i,c}^o$ by means of the following linear index:

$$U_{i,c}^o = \beta_0^o + Z_i\beta_Z^o + R_i^o\beta_R^o + V_i^o\beta_V^o + v_{i,c}^o \tag{17}$$

where $v_{i,c}^o$ is a residual term that captures both unobserved determinants that are common across two or more outcomes (see Section 5.2 for more details) and the aforementioned unobserved exogenous outcome-specific determinants $\varepsilon_{i,c}^o$.

With respect to the wage equations ($o \in \{13, 15, 17\}$), we consider a log-linear specification with a similar set of determinants for the binary outcomes. Moreover, to account explicitly for potential differences in overeducation penalties across levels of education, we also include interaction terms between the overeducation dummy, oe_i, and a subset of endogenous dummies, E_i, measuring one's educational attainment:

$$\log w_{a,i} = \beta_0^{w_a} + Z_i\beta_Z^{w_a} + R_i^{w_a}\beta_R^{w_a} + V_i^{w_a}\beta_V^{w_a} + oe_i\beta_{oe}^{w_a} + oe_iE_i\beta_{oe,E}^{w_a} + v_i^{w_a} \tag{18}$$

Finally, the vectors of endogenous variables, V_i^o, include all realised outcomes before outcome o, with the exception of three cases. First, by construction, these vectors do not include any of the previous outcomes that act as a selection variable for outcome o. This is the case for the four enrolment dummies which act as selection variables for the subsequent educational outcomes as well as for the selection dummies related to the wage equations. Second, in the wage equations, we do not include wages at earlier ages as determinant(s). We made this decision on the basis that wages are not consistently observed across all ages for all cohorts. The estimated effects in these wage equations are therefore to be interpreted as reduced-form effects that also take into account indirect effects through prior wages. Finally, as discussed in the next section, delay at the start of primary education is not added as a direct determinant of the labour market outcomes.

5.2 Selection bias and identification

If not adequately addressed, two different types of selection bias may emerge when estimating our model. First, there is classical selection bias resulting from the fact that the treated individuals may differ from the control group in a number of respects that are not covered by the observable exogenous variables. For instance, individuals who managed to attain a particular educational degree are likely to be different in terms of abilities and motivations relative to those who dropped out. In case these abilities and motivations also drive labour market outcomes, this would lead to a biased estimate of the labour market return to this degree. Second, the estimates may be biased due to dynamic selection bias. This is because of the increasing negative correlation between a treatment and the unobservable characteristics as students progress their educational careers (Cameron and Heckman, 1998). For instance, even if the selection into lower secondary education was aselective, this is unlikely to be the case for the selection into the subsequent stages of the educational system. Accordingly, among those who did not enrol into the subsequent stage, those who did enrol in lower secondary education would nonetheless be different in terms of unobservables from those who did not enrol in lower secondary education. The implication of this is that the estimated labour market effects of enrolling in lower secondary education conditional on enrolment in higher secondary education would nonetheless be biased.

To account for these two types of biases, we apply the following factor structure to the error term $v^o_{i,c}$:

$$v^o_{i,c} = \omega^o_k \eta_k + \varepsilon^o_{i,c} \tag{19}$$

in which η_k is a random effect, which is independent of the observed exogenous characteristics (Z_i and R^o_i) and independent of the outcome-specific residuals $\varepsilon^o_{i,c}$. This random effect is intended to represent any variation in the unobserved exogenous determinants that are not specific to one of the outcomes and that are not captured by the vectors of observed exogenous individual characteristics (Z_i and R^o_i). This approach is similar to that adopted by Cameron and Heckman (1998, 2001) to account for dynamic selection. Moreover, as all the treatments of interest (i.e. educational attainment and overeducation) are themselves modelled as outcomes of earlier choices (and, therefore, are dependent on the unobserved random effect), our approach also accounts for the former, more classical selection problem[8].

Following the literature on dynamic discrete choice models, we deploy a finite mixture distribution to model the unobserved random variable η_k (cf. Heckman and Singer, 1984; Arcidiacono, 2004)[9]. We assume that this distribution is characterised by an a priori unknown number of K different heterogeneity types with type-specific heterogeneity parameters ω^o_k for each outcome[10]. This prevents us from having to rely on strong distributional assumptions and, therefore, also minimises any bias resulting from misspecification in this respect (Heckman and Singer, 1984; Hotz et al., 2002).

To identify this unobserved component and, ipso facto, the treatment effects of interest, we rely on two different sources of information (cf. Heckman and Navarro, 2007; Heckman et al., 2016, 2018a, 2018b). First, we exploit the panel structure of the data by relying on the assumption that all treatments and outcomes are part of the same, more general human capital decision-making process. This implies that we have to solve an initial

[8] This is different for selection problems related to Z_i and R^o_i, as the random effect is assumed to be independent of these variables. However, this is not a problem as the effects of these variables are not the focus of our paper.

[9] It enters each likelihood contribution as a constant parameter, but, given the probability weight for each observation, it becomes a dummy capturing type-specific shocks.

[10] See footnote 5.

conditions problem (Keane and Wolpin, 1997; Cameron and Heckman, 1998, 2001; Keane et al., 2011). In the context of our model, this refers to the fact that this process may already have been initialised prior to enrolment in lower secondary education, which is the earliest choice of interest in the model. Hence, we decide to start the model already with a delay at the start of primary school (i.e. at the age of six) as the first outcome. This assumption regarding the initialisation of the process is substantially weaker than the assumptions made in many earlier studies using the same methodology (see, e.g., Hotz et al. 2002; Adda et al. 2010). Another implication is that the identification can be facilitated by adding to the model other decisions that are a crucial part of this decision process but are beyond the scope of the analysis (Cockx et al., 2017). Hence, we decide to also model the track choice, which is strongly selective in Flanders and generally considered to be an important determinant of subsequent educational and labour market outcomes.

As a second source of identification, we follow Arcidiacono (2005), Heckman and Navarro (2007), Heckman et al. (2016, 2018a, 2018b), and Ashworth et al. (2021) by also adding a set of exclusion restrictions. First, as the unemployment rate at the district level is a time-variant variable, the unemployment rate related to a specific outcome acts, de facto, as an exclusion restriction for the subsequent outcomes (cf. Heckman et al., 2018a, 2018b; Ashworth et al., 2021). Second, we add the delay at the start of primary education as an explanatory variable for the subsequent educational outcomes but not for the labour market outcomes (cf. Baert et al., 2022). We thus assume that the delay in primary education affects the labour market outcomes only indirectly through its effect on the delay at the start of secondary education. As the labour market effects of delay at the start of secondary education are unlikely to depend upon when it took place, this is a reasonable assumption.

5.3 Maximization and model selection

The estimation of this model is carried out by using an Expectation-Maximization (EM) algorithm (Dampster et al., 1977; Arcidiacono and Jones, 2003; Arcidiacono, 2005). This approach was originally formulated by Dampster et al. (1977), before being further developed by Arcidiacono and Jones (2003) and Arcidiacono (2005). It is composed of (i) an expectation and (ii) a maximization step, both of which are repeated until convergence is achieved.

In the expectation step, we compute the probability of each individual being in each heterogeneity type k, based on the likelihood value for each $k \in K$: $\mathcal{L}_i(Z_i, R_i, V_i, \omega_k; \theta)$. Indeed, for each type k, we know the type-specific likelihood and the total expected likelihood weighted by the probability of being in each type k, $\pi_{k,i}$:

$$\mathcal{L}_i(Z_i, R_i, V_i, \omega_k; \theta) = \sum_{i=1}^{I} \ln \left(\sum_{k=1}^{K} \pi_{k,i} \prod_{o=1}^{O} \mathcal{L}_i^o(Z_i, R_i^o, V_i^o, \omega_k; \theta) \right) \tag{20}$$

Bayes' rule implies that the probability of individual i being a type k, conditional on the observed variables, endogenous outcomes, and unobservables, is as follows:

$$\hat{p}_{k,i}(k|Z_i, R_i, V_i, \pi) = \frac{\pi_{k,i}\mathcal{L}_i(Z_i, R_i, V_i, \omega_k; \theta)}{\sum_{k=1}^{K} \pi_{k,i}\mathcal{L}_i(Z_i, R_i, V_i, \omega_k; \theta)} \tag{21}$$

In the maximization step, the conditional probabilities of being heterogeneity type k are treated as given, which allows us to optimise the full model by maximum likelihood.

$$\hat{\theta} = \arg\max_{\theta} \sum_{i=1}^{I} \sum_{k=1}^{K} \hat{p}_{k,i}(k|Z_i, R_i, V_i, \pi) \left(\sum_{o=1}^{O} \ln(\mathcal{L}_i^o(Z_i, R_i^o, V_i^o, \omega_k; \theta)) \right) \tag{22}$$

After the maximization step, we update the conditional probabilities and iterate to the next maximization. This process is repeated until convergence is achieved.

To identify the optimal number of heterogeneity types k, we re-estimate the model by gradually adding up to four types to the model. In Table 4, we report the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) values on each of these models. Based on these criteria, we select the model with three heterogeneity types ($K = \{1, 2, 3\}$) as our benchmark model. The proportions of the three types are 27.5%, 6.4%, and 66.3%, respectively. For $k = 2$ and $k = 3$, η_2 and η_3 enter the likelihood function as an additional intercept.

5.4 Counterfactual simulation

To gauge the treatment effects of interest and their confidence intervals, we rely on a counterfactual simulation strategy (Cockx et al., 2019). In each of the 999 draws of the simulation, the parameters used are randomly drawn from the asymptotic normal distri-

Table 4: Model selection using AIC and BIC

Model:	Number of parameters	Log-likelihood	AIC	BIC
K1	340	-35674.28	72028.55	72678.56
K2	357	-32152.00	65018.00	65700.51
K3	374	-29770.49	**60288.98**	**61003.99**
K4	391	-29778.63	60339.28	61086.79

Notes: Each model is named after Kn, with $n \in \{1, 2, 3, 4\}$, which is defined, as in this section, using the mathematical notation of k unobserved types. Therefore, Kn represents the model with n heterogeneity types.

bution of the model's parameters. Subsequently, for each of these draws, the probability types, estimated using the EM algorithm, are used to randomly assign a heterogeneity type to each individual in the sample. Thereafter, based on this novel set of parameters, we simulate the full sequence of schooling and labour market outcomes for each individual in the sample.

This counterfactual simulation strategy is also used to assess the quality of the model by generating the full set of outcomes and comparing it to the observed outcomes in the data. This is shown in Table 5. In most cases, the observed probabilities fall within the 95% confidence bounds of the simulated probabilities. Thus, the model fits the observed outcomes in the dataset relatively well.

A similar simulation strategy is adopted to gauge the composition of the three heterogeneity types. Table 6 displays the simulated outcomes when forcing all individuals to be in one of the three heterogeneity types, labelled as Type 1, 2, or 3. With respect to the two main types, a clear pattern emerges with Type 1 individuals having (relative to Type 3 individuals) a higher probability of experiencing a delay at the start of primary and secondary education, a lower probability of completing each level of educational attainment, a higher probability of being overeducated, and a lower average wage. This is consistent with Type 1 individuals being of lower ability relative to Type 3 individuals. Type 2 individuals, who are much less prevalent in the data, seem to form a more specific category, as they combine a high probability of overeducation with high wages.

Table 5: Goodness of fit

Variables	**Observed**	Simulation	95% CI	
(a) Delays:				
Delay in Primary Education	**0.015**	0.017	0.014	0.020
Delay in Secondary Education	**0.101**	0.104	0.097	0.111
(b) Educational choices:				
Start and Track Choice in LSE	**2.515**	2.508	2.498	2.518
LSE	**0.954**	0.952	0.948	0.957
Start and Track Choice in HSE	**2.379**	2.374	2.363	2.386
HSE	**0.887**	0.886	0.879	0.892
Start and Track Choice in LTE	**1.852**	1.852	1.837	1.867
LTE	**0.486**	0.489	0.478	0.499
Start and Track Choice in HTE	**1.361**	1.383	1.368	1.399
HTE	**0.194**	0.206	0.197	0.214
(c) Labour market outcomes:				
Overeducation	**0.353**	0.349	0.338	0.359
Wage Selection at 23	**0.539**	0.527	0.518	0.536
Log-hourly wage at 23	**1.974**	1.982	1.979	1.985
Wage Selection at 26	**0.414**	0.413	0.409	0.417
Log-hourly wage at 26	**2.072**	2.076	2.072	2.080
Wage Selection at 29	**0.385**	0.381	0.376	0.386
Log-hourly wage at 29	**2.126**	2.126	2.121	2.130

Notes: Educational attainments are defined as Lower Secondary Education (LSE), Higher Secondary Education (HSE), Lower Tertiary Education (LTE), Higher Tertiary Education (HTE). 95% CI indicated the 95 percent confidence intervals, as simulated using our approach.

Table 6: Probability types simulated models

	Overall	Type 1	Type 2	Type 3
		27.40%	*6.40%*	*66.20%*
(a) Delays:				
Delay in Primary Education	**0.017**	0.022	0.021	0.014
Delay in Secondary Education	**0.104**	0.127	0.110	0.095
(b) Educational choices:				
LSE	**0.952**	0.939	0.946	0.963
HSE	**0.886**	0.840	0.871	0.910
LTE	**0.489**	0.381	0.416	0.537
HTE	**0.206**	0.130	0.194	0.233
(c) Labour market outcomes:				
Overeducation	**0.349**	0.373	0.405	0.339
Log-hourly wage at 23	**1.982**	1.964	2.155	1.977
Log-hourly wage at 26	**2.076**	1.471	2.430	2.052
Log-hourly wage at 29	**2.126**	1.554	2.441	2.110

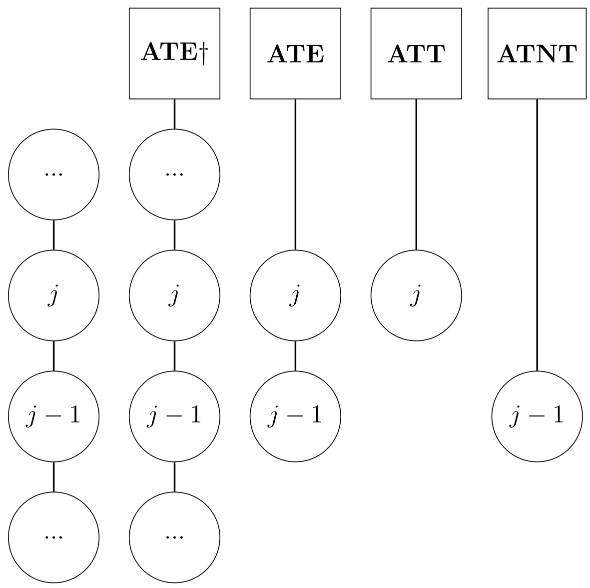
Notes: Educational attainments are defined as Lower Secondary Education (LSE), Higher Secondary Education (HSE), Lower Tertiary Education (LTE), Higher Tertiary Education (HTE).

5.5 Treatment Effects

As in Heckman et al. (2018a, 2018b), we define different treatment effects for analysing the impact of educational attainment on overeducation and wages. The first treatment effect to estimate is denoted as ATE†, which is the treatment effect computed over the entire population. This ATE is less relevant from a practical perspective because dynamic selection does not result in everyone having a reasonable likelihood of reaching each level of education. Therefore, we define a more credible treatment effect, ATE, which is computed over everyone at one of the two final nodes. For instance, for the likelihood of being overeducated and for the wage returns related to HTE, we compute the treatment effect over those who obtained either an LTE or an HTE as their maximum level of educational attainment.

Moreover, by calculating this separately over those with the treatment level of educational attainment and those with a level of educational attainment that is one level below the treatment level, we can also define the average treatment effect on the treated (ATT) and the average treatment effect on the non-treated (ATNT) (e.g. when the treatment obtains an HTE, ATT for those that obtained an HTE, and ATNT for those with an LTE

Figure 3: Definition of treatment effects



Notes: The first column represents the full sample, including individuals at j and $j-1$ and individuals included in other nodes (represented by circles containing "..."). Individuals are included in a given j educational attainment and in $j-1$ (i.e. the lower educational attainment, e.g. if j=HTE, then j-1=LTE). As described in the main text, ATE† is computed over the full sample, ATE over the individuals at the final nodes (j and $j-1$), ATT over individuals in j, and ATNT over individuals in $j-1$.

only). The difference between the ATT and the ATE is a measure of sorting on gains, while the difference between the ATNT and the ATE is a measure of sorting on losses (Heckman et al., 2018a, 2018b). These definitions are summarised in Figure 3, where j represents the treatment level of educational attainment and $j-1$ represents one level below this treatment level (e.g. if j is college, $j-1$ is high-school). The circles indicate which part of the sample is taken into account for the calculation of each of the treatment effects.

Finally, in addition to differentiating between ATE†s and ATEs, we also differentiate between direct ATEs and total ATEs, with total ATEs also taking into account that a certain level of educational attainment enables an individual to enrol in programmes at higher levels of educational attainment and, thereby, generate indirect effects.

6 Results

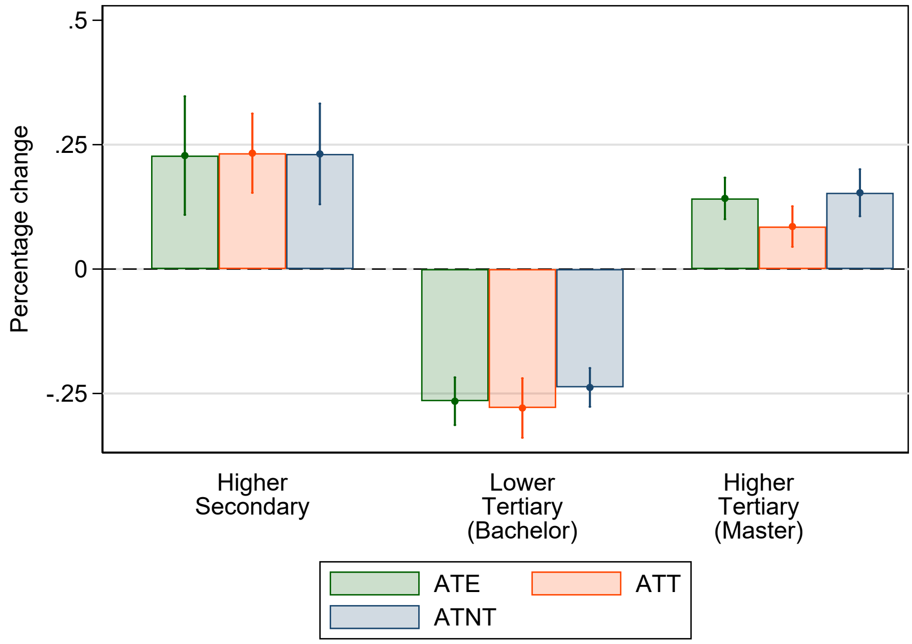
In this section, we present the simulated treatment effects of interest. First, we present results on the impact of educational attainment on overeducation. Second, we report the results on the wage returns conditional on one's match status as well as on the wage penalty for overeducation. Third, we simulate average unconditional wage returns to education and use the results in the preceding subsections to decompose these unconditional returns into various components. Fourth, we consider how overeducation generates heterogeneous wage returns. All of these simulations are based on our preferred model, of which the full set of estimated parameters is reported in the Appendix C. In our description, we primarily focus on the direct ATEs and highlight the main differences with the other definitions of treatment effect. In the final subsection, we also conduct some sensitivity analyses by using alternative sets of estimations.

6.1 Overeducation and educational attainment

Figure 4 shows the ATE of each of the considered levels of education on overeducation, conditional on having obtained the preceding level of attainment. For instance, the effect of a master's degree represents the effect relative to having obtained a bachelor's degree only.

The effect on the level of educational attainment is clearly non-linear. While entering

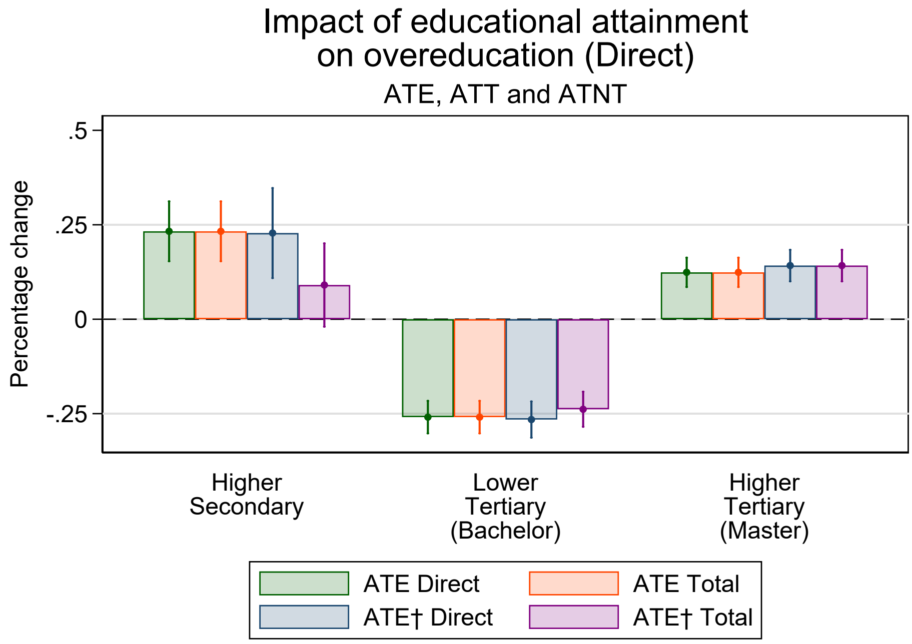
Figure 4: Impact of educational attainment on overeducation (ATE, ATT and ATNT)



the labour market with a high-school degree (HSE) is found to increase one's probability of being overeducated relative to entering the labour market with only a lower secondary degree (LSE), the opposite is true with respect to a bachelor's degree (LTE) relative to a high-school degree (HSE). Both effects are relatively substantial, with a high-school diploma increasing one's likelihood of being overeducated by approximately 22 percentage points and a bachelor's degree reducing this likelihood by around 25 percentage points. Finally, additionally investing in a master's degree further increases one's probability of being overeducated relative to having only obtained a bachelor's degree. However, as the latter effect is estimated to be roughly 11 percentage points only, a master's degree still reduces one's likelihood of being overqualified relative to a high-school degree by around 14 percentage points. These outcomes are clearly in line with a Polarised labour market and challenge the idea that overeducation is primarily a problem among tertiary education graduates.

In Figure 4, we further differentiate between the ATT and the ATNT. In line with Heckman et al. (2018a, 2018b), we find evidence of sorting on gains at the higher stages of one's educational career. Individuals sort on their expected benefits from obtaining a tertiary education degree in terms of experiencing lower levels of overeducation. This may be attributed to high-ability individuals expecting better-matched and, henceforth, better-paid jobs when participating in higher education.

Figure 5: Impact of educational attainment on overeducation (ATE and ATE†, Direct and Total effects)



Finally, in Figure 5, we also look at how the estimated effects change when the sample is extended beyond the final nodes (ATE†) and when total (instead of direct) effects are considered. For the bachelor's and master's levels, the results are largely similar. The treatment effect of starting and obtaining an HSE, meanwhile, is clearly lower when the total ATE† effect is considered. This is driven by the dynamic nature of our model in the sense that obtaining an HSE degree does not only directly increase one's risk of overeducation but also grants access to higher levels of educational attainments that are associated with a lower risk of overeducation. The difference in outcomes between the total ATE and total ATE† is in line with the aforementioned sorting on gains: those for which the effect of obtaining a higher education degree on overeducation is lower are more likely to select themselves into higher education and are, therefore, less likely to be included in the calculation of the ATE with respect to obtaining an HSE.

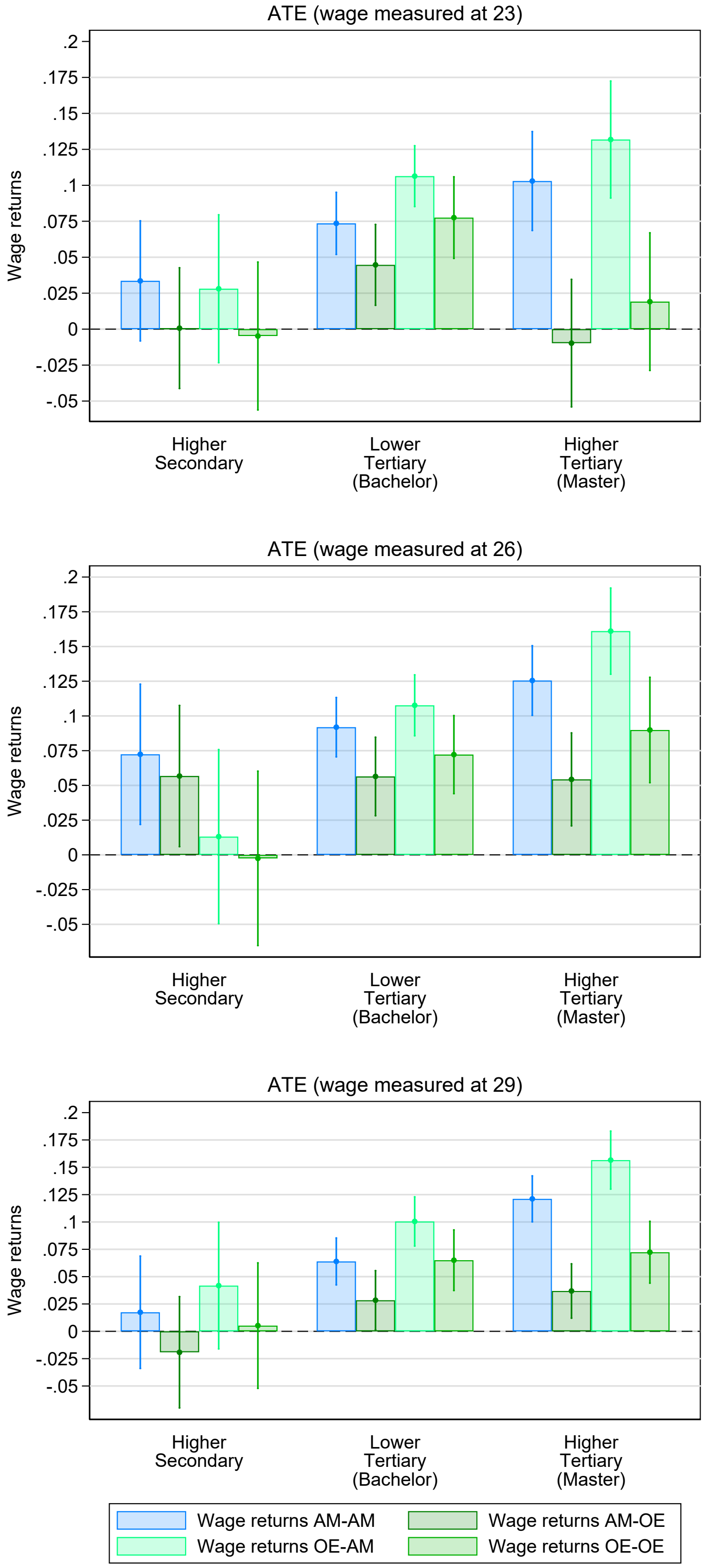
6.2 Conditional wage returns to education

In Figure 6, we report the direct ATE on wages depending on one's match status at the attained level of education (j) and the match status one may have obtained at the preceding level ($j-1$). This delivers the following four conditional returns: (a) the wage return when being adequately matched in both j and $j-1$, (b) the wage returns to overeducation at level j while being adequately matched at level $j-1$, (c) the wage return to being adequately matched when starting from an overeducation status and, lastly, (d) the wage return when being overeducated in both j and $j-1$. These conditional returns are reported for each of the three wage observations.

The first type of conditional return (a), which is the return to educational attainment assuming one is always perfectly matched, is found to gradually increase depending on the level of educational attainment. For instance, at age 23, the wage returns to obtain a higher secondary, a lower tertiary, and a higher tertiary education degree presuming one is always adequately matched are 3.4%, 7.4%, and 10.3%, respectively. At age 29, these conditional returns are 1.7%, 6.4%, and 12.1%, respectively. Moreover, the second type of wage return (b), which is conditional on being overeducated at level j and being adequately matched at the preceding level, is, in most cases, positive. Nonetheless, we find this return to be consistently lower than the return conditional on being adequately educated at both levels of attainment. For instance, at age 29, the wage return to obtaining a higher tertiary degree (relative to a lower tertiary degree) is estimated to be equal to 3.7% only in cases where this additional investment leads to an individual being overeducated.

These first two types of conditional returns are equivalent to the wage return to adequate education (a) and the return to overeducation (b), as typically reported in the literature on overeducation. Moreover, by subtracting these returns, we obtain the results on the overeducation wage penalty, which are reported in supplementary Figure 7. These penalties are statistically significant at each age and at each level of educational attainment. At age 23, they range from 2.9% for those with a bachelor's degree to 11.3% for those with a master's degree, although the latter penalty drops somewhat to 8.4% if measured at the age of 29. Note that these effects represent the effects of the match status at the start of the first job. Therefore, besides indicating that the overeducation penalty is real, these findings also suggest that initial overeducation generates a long-lasting scarring

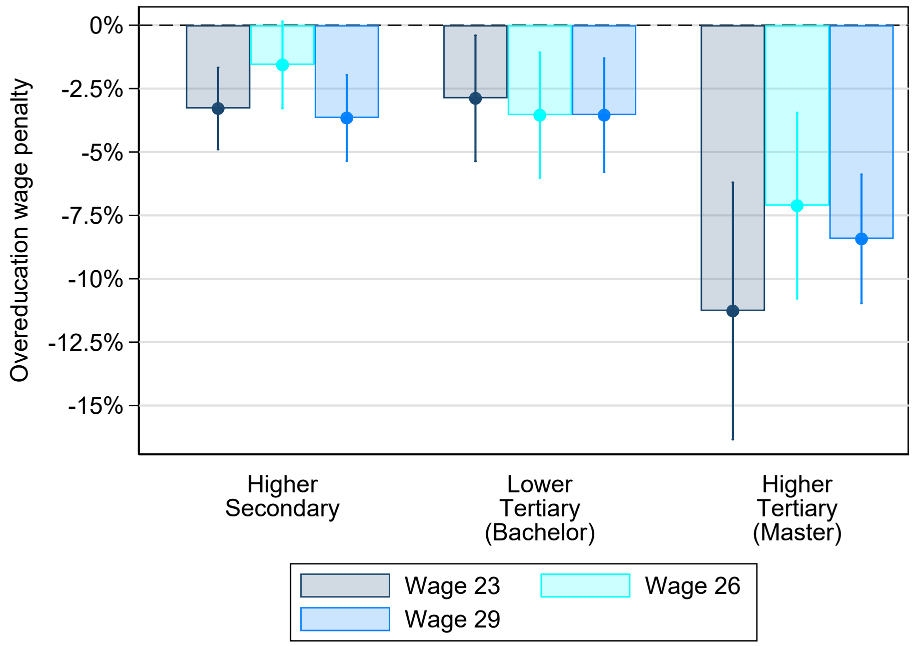
Figure 6: Conditional wage returns (at 23, 26 and 29 years)



Notes: We simplify the notation from Equations 7, 8, 9, and 10, and we refer to the following: (i) Wage returns AM-AM as $\Omega^{M,M}_{a,i,j}$, (ii) Wage returns AM-OE as $\Omega^{M,O}_{a,i,j}$, (iii) Wage returns OE-AM as $\Omega^{O,M}_{a,i,j}$ and (iv) Wage returns OE-OE as $\Omega^{O,O}_{a,i,j}$.

effect.

Figure 7: Overeducation Wage Penalty by educational attainment



Investing in more education may not only induce people to stay adequately educated or to become overeducated; it may also improve their match status (case (c)) or increase their likelihood of staying overeducated (case (d)). As shown in Figure 6, the wage returns conditional on these match statuses are usually positive, at least when they concern investment in tertiary education. For instance, at age 29, we find that the wage return to a higher tertiary education degree (relative to a lower tertiary degree) is equal to 15.7% in cases where the individual also manages to improve their match status through this investment. Similarly, the return to a higher tertiary education degree is still 7.2% for overeducated workers in cases where this additional investment would have induced one to remain overeducated. Moreover, while the former conditional return comfortably exceeds the return conditional on being adequately educated at both the considered and the preceding level of attainment, the latter return usually exceeds the return for those who become overeducated. Henceforth, the standard measure of the wage return to education for those who are overeducated may provide an underestimation of their true wage return to education.

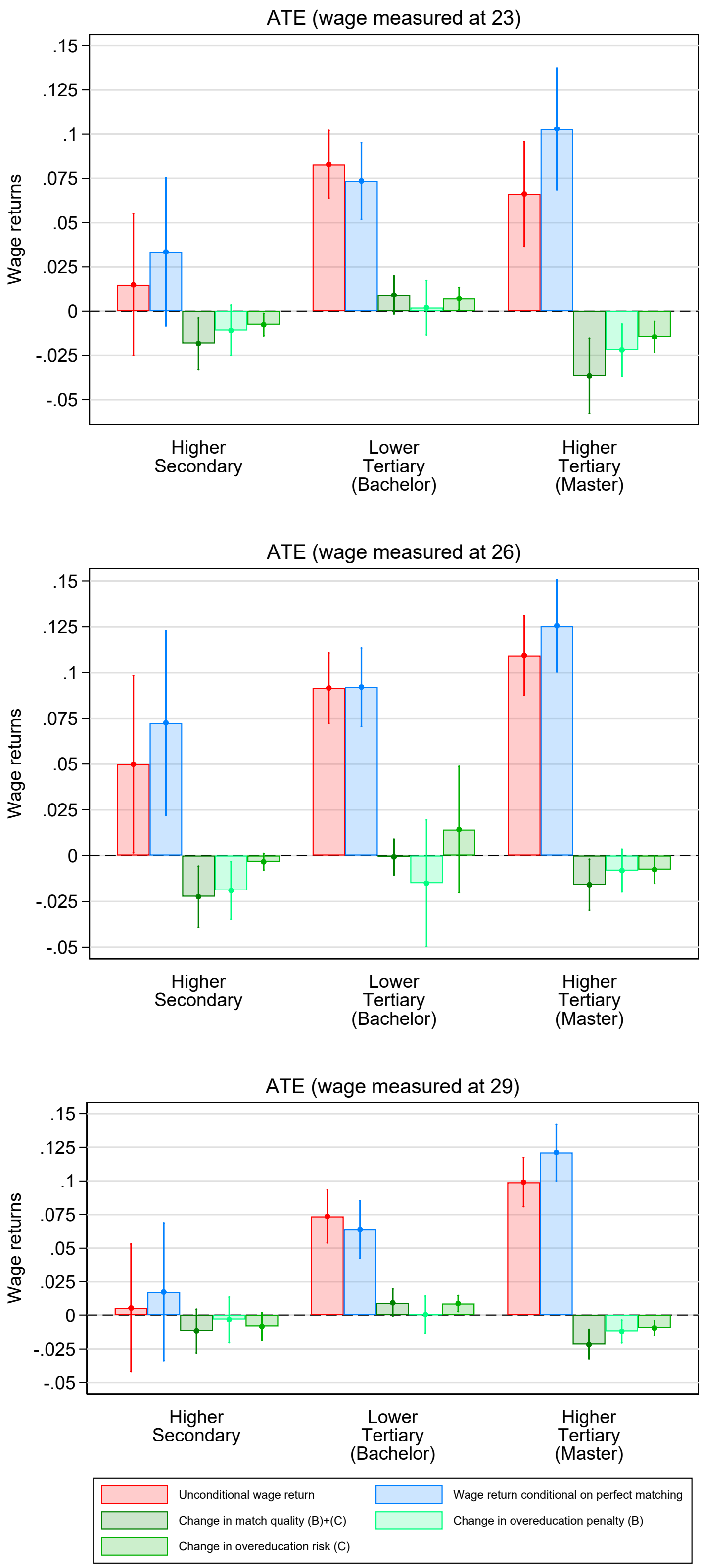
In Appendix B, we also report the results on these conditional returns while relying on our alternative treatment indicator ATE† and while also taking into account the indirect effects of additional educational investments on subsequent levels of educational attainment. Overall, our conclusions are largely similar when relying on these alternative treatment effect definitions. The main differences once more pertain to the higher secondary education level. Several of the estimated conditional returns to this level of attainment are small and statistically insignificant when relying on the direct ATE definition. This is likely due to labour market institutions, such as collective bargaining and minimum wages, which may generate strong wage compression at the lower end of the wage distribution. When relying on the total ATE† definition, however, these returns become much more substantial and statistically significant. Also, this result is consistent with the wage returns to a higher secondary education being mainly indirect, as obtaining a higher secondary education degree opens the door towards tertiary education.

6.3 Unconditional wage returns: decomposition

To investigate how overeducation affects the average unconditional wage return, we implement a decomposition approach. In Figure 8, we report this decomposition while relying on the direct ATE definition. While the first bar reports the unconditional return, the next two bars represent its decomposition into a part that reflects the return in the case of fixed match quality across levels of attainment and another due to changes in match quality. The last two bars represent a further decomposition of the change in match quality component in one subcomponent due to changes in wage penalties for overeducation and another due to changes in overeducation risk.

While the unconditional wage returns to obtaining a master's degree are consistently positive, they are lower than those in the case of perfect matching. For instance, at age 23, the unconditional return is equal to 6.6% relative to a return of 10.3% in the case of perfect matching. Almost two-thirds of this difference (2.2 %-points) is caused by the larger overeducation penalty for master's relative to bachelor's degrees while the remaining part (1.4 %-points) is attributable to the larger overeducation risk among master's graduates. Over time, however, this difference clearly drops even if master's graduates also experience a further increase in the wage return to education in the case of perfect matching. For instance, at age 29, the unconditional return is equal to 9.9% relative to a return of 12.1% in the case of perfect matching. This is largely due to a drop in the relative importance of the difference in overeducation penalties between bachelor's and master's graduates.

Figure 8: Decomposition of change in match quality



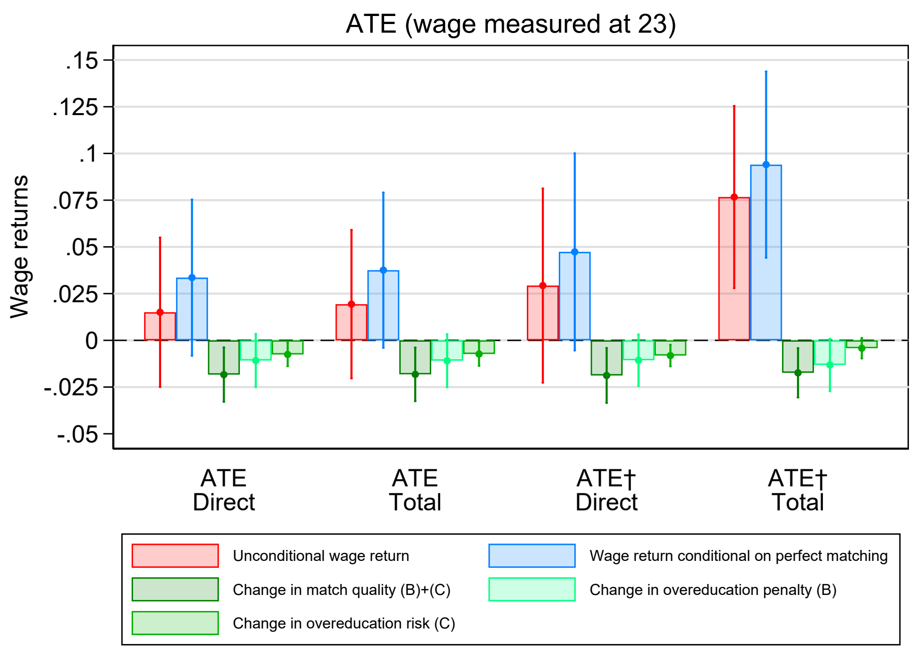
Notes: The decomposition approach is explained in Section 2 and in Equation 15: each bar included in this graph can be referred directly to this equation.

Indeed, in the previous section, we reported a clear drop in the overeducation penalty over time for master's graduates.

For bachelor's graduates, the results are clearly different, with their unconditional wage return being at least equal to (at age 26) or even larger than the wage return in the case of constant match quality across levels of educational attainment. For instance, at age 29, their unconditional wage return is 7.4% relative to a wage return of 6.4% conditional on perfect matching. Our decomposition suggests this is largely due to differences in overeducation probabilities between those with a high-school and those with a bachelor's degree. Indeed, in Section 6.1, we reported that investing in a bachelor's degree causes one's overeducation risk to drop substantially.

Finally, in regard to obtaining a high-school degree, the estimated unconditional returns are lower but also less precise. Hence, its estimate is only statistically significant at age 26. In addition, as a result of a low return in the case of perfect matching, this is also due to a significant drop in match quality relative to when one would have entered the labour market without a high-school degree. For instance, at age 23, this drop in match quality is estimated to reduce the average unconditional return by approximately 1.8 %-points.

Figure 9: Decomposition of change in match quality (ATE and ATE†, Direct and Total effects)



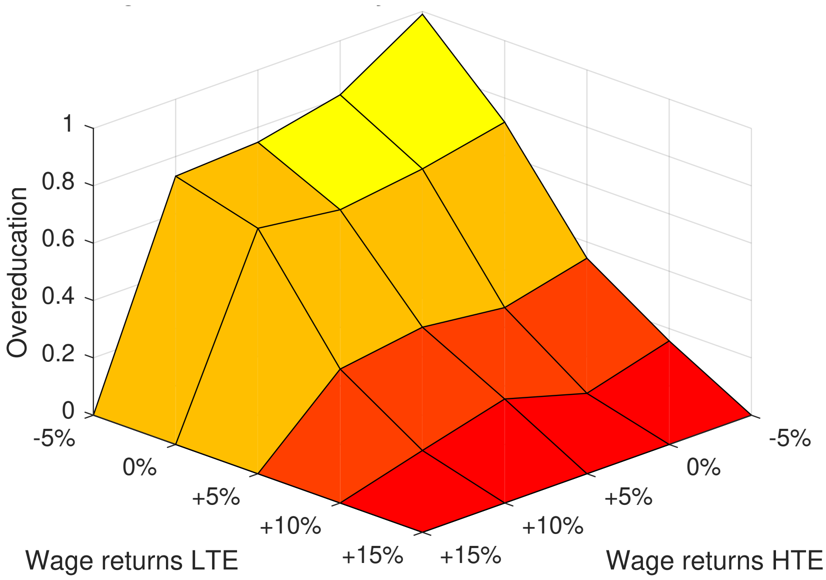
6.4 Heterogenous Wage Returns to Education

In Figure 9, we further report results on this decomposition for obtaining a high-school degree while relying on alternative indicators of the treatment effects. The results pertain to wages at age 23. The results for these alternative indicators for ages 26 and 29 and for the other levels of educational attainment are reported in Appendix B. As a high-school degree provides access to higher education, the results are once again more favourable when relying on the total ATE† indicator. While the direct unconditional return is estimated to be equal to 1.5% only when relying on the direct ATE indicator, the total ATE† return is estimated to be 7.7%. This is primarily due to the higher estimated return conditional on perfect matching, which is estimated to be equal to 9.4% relative to 3.4% only when relying on the direct ATE indicator. The estimated effect caused by changes in match quality, meanwhile, is fairly similar across the alternative indicators. Even though overeducation penalties are usually larger at higher levels of obtained education, this is levelled out by the lower overeducation risk that is associated with obtaining a higher education degree relative to a high-school degree only.

As a first indication of how overeducation may be associated with heterogeneous returns to education, Figure 10 compares the simulated overeducation probabilities with the simulated wage returns at age 23 for obtaining an LTE and HTE relative to obtaining a high-school degree only. To this end, we rely on the full set of simulated outcomes for all individuals. A clear negative relationship emerges between the predicted risk of being overeducated and the unconditional wage return to higher education. Overall, this is consistent with overeducated individuals having, on average, less favourable traits that reduce the potential benefits they receive from obtaining a higher education degree.

To determine whether the differences in overeducation risk across individuals are merely a reflection or also a cause of these heterogeneous returns, we also report, in Figure 11, the simulated between-individual distribution of the unconditional ex-ante wage returns to education along with the simulated between-individual distribution of their two components: the wage returns to education conditional on perfect matching, and the wage components due to changes in match quality. Each observation in the presented distributions represents the expected value of these three variables for one individual in the sample, which are calculated by averaging, within each individual, over their 1,000 simulated values. We report separate results for each level of educational attainment and

Figure 10: Heterogeneous wage returns and overeducation (conditional on HSE)



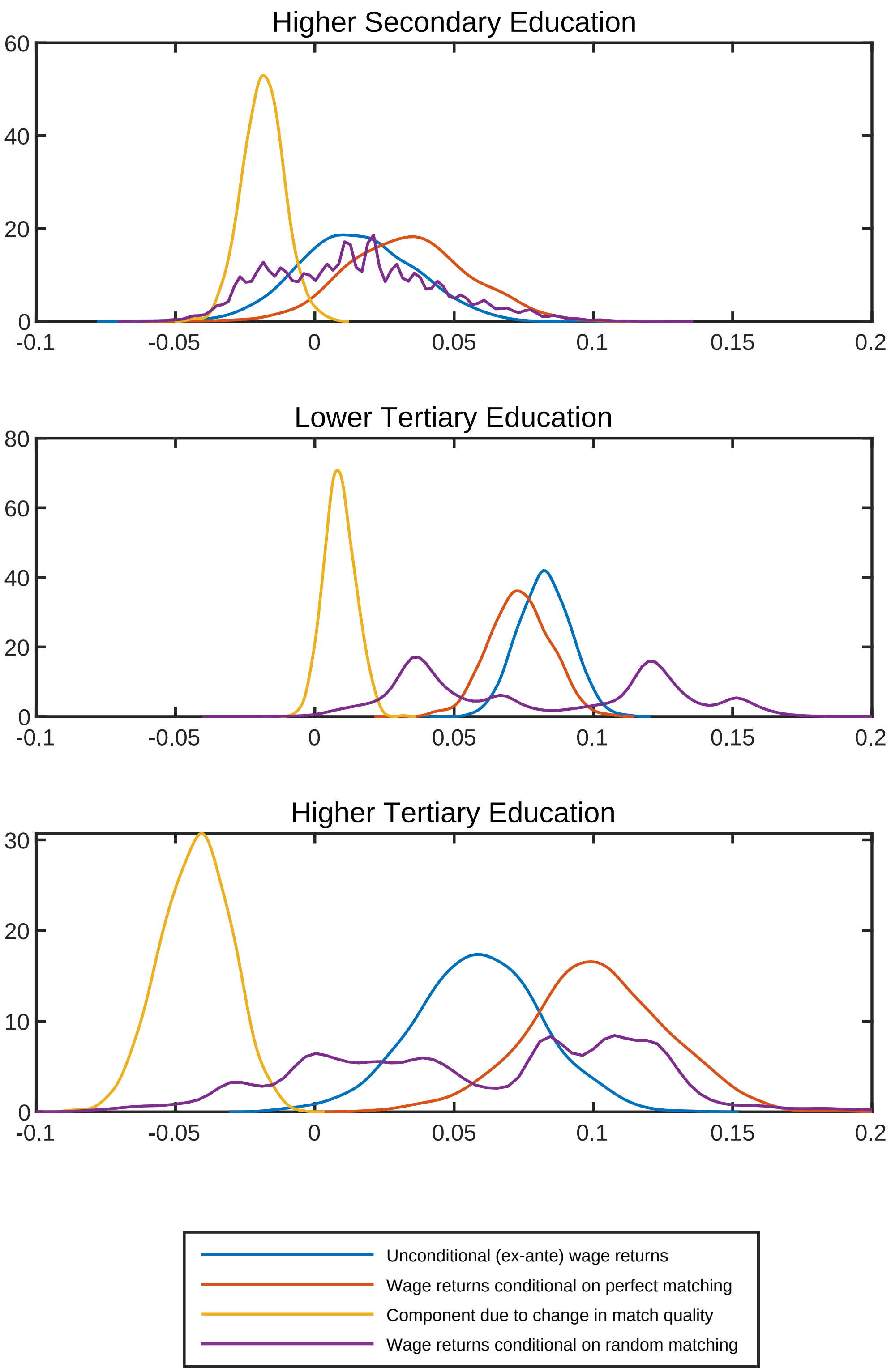
Notes: This figure is obtained using the full sequence of simulated outcomes. In this framework, we discretise wage returns and compute the fraction of overeducated individuals for each composition of the resulting discretised wage returns bins. We use wage returns at the age of 23 so as to avoid issues with cohort data.

concentrate on ATE wage returns at age 23. The results for the other age observations are reported in Appendix B.1.

We first focus on the distributions of the two components. As can be seen, the between-individual heterogeneity in wage returns conditional on perfect matching is estimated to be substantial. This heterogeneity in conditional returns is the most substantial with respect to obtaining a master's degree and somewhat less pronounced with respect to obtaining a bachelor's degree. Moreover, our model also suggests the expected wage component due to changes in expected match quality to be heterogeneous across individuals, albeit to a lesser extent relative to the heterogeneity in components in the case of perfect matching. Furthermore, in line with the results reported in the previous sections, we find this expected component due to changes in match quality to be negative for most individuals with respect to obtaining a high-school or a master's degree. With respect to obtaining a bachelor's degree, meanwhile, this component is positive for most individuals.

As expected, we find the between-individual heterogeneity in these two components to also translate to a substantial between-individual heterogeneity in the unconditional wage returns. Nonetheless, this heterogeneity is, for each of the three considered levels of educational attainment, fairly similar to the heterogeneity in returns conditional on perfect

Figure 11: Simulated distributions of unconditional wage returns, their decomposition and realized wage returns (age 23)



matching. Overall, this indicates that the differences in overeducation risk mainly serve to reflect the heterogeneity in unconditional returns rather than reinforce this heterogeneity. Nonetheless, in line with our findings in the earlier sections, we find that overeducation affects the location of the distributions of unconditional returns. As a result of the negative component due to changes in match quality when obtaining a high-school or master's degree, we find that the distribution of the unconditional wage return for obtaining such a degree is situated to the left of the distribution of its returns conditional on perfect matching. This translates to negative unconditional returns for a non-negligible part of the sample. What is more, the unconditional returns for obtaining a master's degree are clearly lower relative to its returns conditional on perfect matching. However, this unconditional return remains substantial for most of the individuals. Finally, in line with the associated improvement in average match quality, we find that the unconditional return for obtaining a bachelor's degree exceeds its return conditional on perfect matching for most individuals.

These unconditional expected wage returns partly depend on one's overeducation risk in the treated level of educational attainment relative to the (lower) control level of attainment. However, depending on one's effective match status at each level of attainment, realised (i.e. ex-post) returns may be lower or higher. For instance, even if one has a high risk of overeducation at the treated level of attainment given one's individual characteristics, one may still manage to be adequately matched due to idiosyncratic matching shocks (i.e. $\varepsilon_i^{oe} \neq 0$). To test for the impact of these idiosyncratic shocks, we also simulate the distribution of returns that may emerge from a random matching process w.r.t. the first job while relying on each individual's estimated overeducation probabilities[11]. The resulting distributions are added to the graphs in Figure 11. As can be seen, the returns resulting from this simulated random matching are far more heterogeneous than the unconditional expected returns. This is particularly the case for obtaining a bachelor's and a master's degree. For instance, in the latter case, a substantial proportion of the individuals have a return conditional on random matching that is well above 10% while others have conditional returns that are just as negative. Overall, this is in line with the notion that overeducation is a consequence of idiosyncratic matching shocks (e.g. due to search

[11]Note that, as our aim is to disentangle the effect of idiosyncratic shocks in overeducation from other effects, we do not account for idiosyncratic shocks in wages. Therefore, this distribution does not represent the final distribution of realised returns.

and matching frictions) and, henceforth, is a source of heterogeneous realised returns to college.

6.5 Sensitivity analyses

We end with two sensitivity analyses related to our model. Our first analysis focuses on the procedure used to measure overeducation. This procedure was based on a combination of three independent measures of overeducation. To gauge the impact of this decision, we re-estimate the model based on each of the three separate overeducation wage measures, as described in Section 4.5.

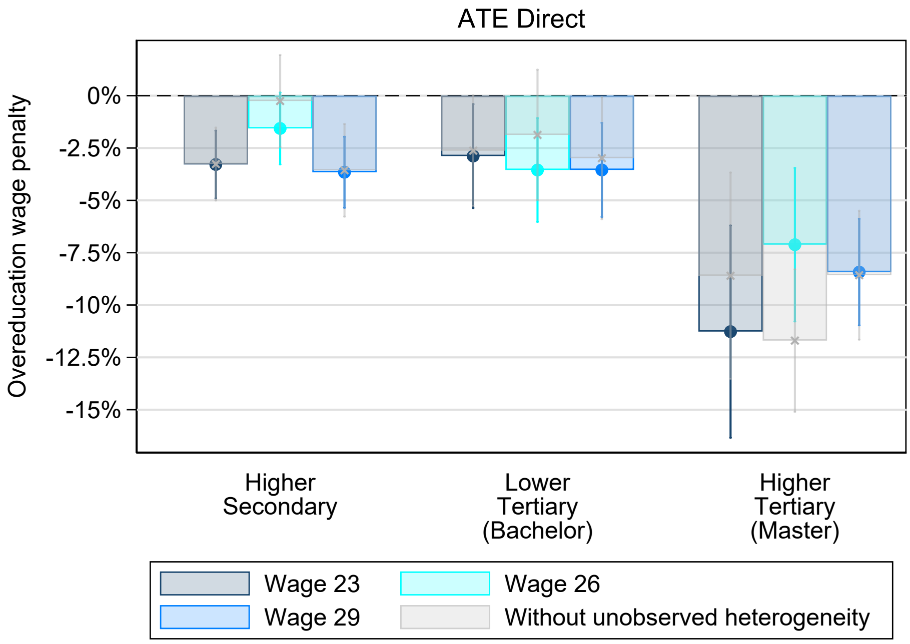
Table 7 presents a selection of ATEs based on these alternative estimates. First, we report results pertaining to the impact of educational attainment on the risk of becoming overeducated. Reassuringly, the direction of the effects is the same across the adopted measures and in line with the benchmark results – that is, while both obtaining a high-school degree and a master's degree is found to increase one's chances of being overeducated (relative to obtaining the previous level) based on each measure, obtaining a bachelor's degree is always found to be associated with a lower overeducation risk. Nonetheless, the estimated effects are different in size across the various measures and, in two cases with respect to obtaining a high-school degree, also statistically insignificant. Second, the finding that there is a wage penalty for overeducation is fairly consistent across all measures and levels of attainment. Only with respect to obtaining a master's degree while relying on the job analysis measure is this penalty estimated to be statistically insignificant. Nonetheless, these penalties are often smaller when relying on the independent measures relative to when relying on the benchmark measure. Third, the results regarding the decomposition are fairly consistent, with the change in match quality negatively affecting the unconditional return for a high-school and master's degree and positively affecting the return for a bachelor's degree. However, for this analysis, the size of the estimated ATEs depends on the adopted measures, with the benchmark measure usually delivering more sizeable and statistically significant estimates on both components of the unconditional wage return. For instance, unlike when relying on the benchmark measure, the other measures do not generate a statistically significant component of the unconditional wage return for obtaining a bachelor's degree that is attributed to a change in match quality.

Table 7: Sensitivity analysis on the overeducation measure

	Educational attainment:	Overeducation measure: **BM**	JA	ISA	DSA
(a) *Effects of Educational Attainment on Overeducation*					
ATE Direct	HSE	**0.228*****	0.089	0.226***	0.055
		(0.060)	(0.063)	(0.063)	(0.058)
	LTE	**-0.265*****	-0.129***	-0.336***	-0.116***
		(0.024)	(0.023)	(0.025)	(0.020)
	HTE	**0.142*****	0.249***	0.106***	0.085***
		(0.021)	(0.019)	(0.020)	(0.017)
(b) *Overeducation Wage Penalty*					
Wage 23	HSE	**-0.033*****	-0.031***	-0.019**	-0.031***
		(0.008)	(0.009)	(0.008)	(0.009)
	LTE	**-0.029****	-0.032***	-0.031**	-0.036**
		(0.013)	(0.011)	(0.013)	(0.014)
	HTE	**-0.113*****	-0.021	-0.112***	-0.078***
		(0.026)	(0.026)	(0.028)	(0.025)
(c) *Unconditional Wage Returns Decomposition*					
Unconditional WR	HSE	**0.015**	0.004	0.014	0.002
		(0.020)	(0.021)	(0.020)	(0.019)
	LTE	**0.083*****	0.080***	0.087***	0.073***
		(0.010)	(0.010)	(0.010)	(0.009)
	HTE	**0.066*****	0.060***	0.065***	0.078***
		(0.015)	(0.015)	(0.015)	(0.014)
WR conditional on PM	HSE	**0.034**	0.024	0.027	0.008
		(0.021)	(0.023)	(0.022)	(0.019)
	LTE	**0.074*****	0.077***	0.083***	0.071***
		(0.011)	(0.012)	(0.011)	(0.010)
	HTE	**0.103*****	0.061**	0.093***	0.092***
		(0.018)	(0.024)	(0.016)	(0.015)
Change in MQ	HSE	**-0.019*****	-0.020**	-0.013	-0.007*
		(0.007)	(0.010)	(0.008)	(0.003)
	LTE	**0.010***	0.003	0.004	0.002
		(0.006)	(0.007)	(0.006)	(0.003)
	HTE	**-0.037*****	-0.000	-0.028***	-0.014**
		(0.011)	(0.019)	(0.009)	(0.007)

Notes: The overeducation measures represent respectively: the Benchmark Measure (BM), Job Analysis (JA), Indirect Self-Assessment (ISA) and Direct Self-Assessment (DSA). Educational attainments are defined as: Higher Secondary Education (HSE), Lower Tertiary Education (LTE, or Bachelor) and Higher Tertiary Education (HTE, or Master). At last, the measures used in the decomposition are the following: unconditional (ex-ante) wage returns (Unconditional WR), wage return conditional on perfect matching (WR conditional on PM) and change in match quality (Change in MQ).

Figure 12: Overeducation Wage Penalty: sensitivity analysis on unobserved heterogeneity



Notes: Overeducation wage penalty by educational attainment computed for both the model without unobserved heterogeneity (in grey) and the benchmark model with three heterogeneity types.

Overall, these results show that, while not leading to radically different conclusions, the choice of our measure influences the magnitude of the estimated effects. As argued in the method section, we believe our benchmark measure to be more accurate and, henceforth, to deliver less biased estimates. In any case, our finding that the benchmark measure often produces stronger overeducation wage penalties as well as stronger effects on both components in the wage decomposition, is consistent with the other three measures being more prone to random measurement errors.

As a second sensitivity analysis, we test whether accounting for unobserved heterogeneity matters. To do this, we re-estimate our model while considering one heterogeneity type only. The results regarding the wage penalty for overeducation are summarised in Figure 12. With the exception of the estimates with respect to wages at age 26 when having obtained a master's degree, we find that not accounting for unobservables either scarcely affects the estimated wage penalty for overeducation or even leads to an underestimation of this penalty.

Upon first glance, this seems surprising given that our results suggested that overeducated individuals have less favourable traits and lower expected unconditional returns for education. However, in addition to accounting for classical unobserved heterogeneity, our models also account for dynamic selection. As shown in Table 6, individuals of Type 3, who enjoy the highest levels of educational achievement combined with the lowest risk of being overeducated (see Section 5.4), are also more likely to be selected in the wage equations. Meanwhile, the opposite is true for individuals of Type 1, who experience lower levels of educational attainment combined with a more elevated risk of overeducation. A likely explanation for this selection is that individuals with less favourable traits are less inclined to take up jobs as the offered wages in these jobs are less likely to exceed their reservation wage. This effect may be reinforced if job seekers are also more reluctant to take up jobs for which they are overeducated. The estimates of the overeducation wage penalty thus suggest that the upward bias due to this reservation wage effect is, in absolute terms, at least as large as or even greater than the negative bias caused by classical unobserved heterogeneity.

7 Conclusions

Based on detailed longitudinal Belgian data, we constructed a dynamic discrete choice model to investigate the relationship between educational attainment, overeducation, and wages. We relied on the literature on dynamic treatment effects and estimated a sequential dynamic model of educational choices and labour market outcomes (Heckman and Navarro, 2017; Heckman et al., 2016, 2018a, 2018b). This allowed us to contribute in three main ways to the literature. First, we contribute to the discussion regarding whether the relationships between educational attainment, overeducation, and wages are causal. Second, we implemented a new decomposition approach, which enabled us to investigate the relationship between overeducation and the unconditional wage return to education in a more comprehensive manner. Third, we explored whether overeducation is a channel that may generate both heterogeneous expected and heterogeneous realised returns for education.

With respect to our first contribution, our results suggested that being overeducated at the start of one's first job generates a significantly negative wage penalty. At age 23, this penalty was estimated to range from around 3% among those with a higher secondary or a lower tertiary education degree to approximately 11% among those with a higher tertiary degree. Overall, this confirms the findings of the already fairly extensive literature on this topic (Hartog, 2000; McGuinness, 2006; Leuven and Oosterbeek, 2011; Barnichon and Zilberberg, 2019). However, most of these earlier studies either relied on standard regression analysis or accounted for endogeneity based on identification strategies that have been strongly criticised (Leuven and Oosterbeek, 2011). By modelling overeducation as a function of all relevant past educational choices and by using the panel data structure of the data to estimate the unobserved heterogeneity component, we circumvented these problems. Interestingly, we also found this wage penalty for overeducation in the first job to persist up until age 29. This is consistent with several other studies that found overeducation to be strongly persistent (Baert et al., 2013; Meroni and Vera-Toscano, 2017; Barnichon and Zylberberg, 2019) or that overeducated workers tend to experience no more wage growth than adequately educated workers (Büchel and Mertens, 2006; Korpi and Tåhlin, 2019). Moreover, it is also consistent with the findings of a more general literature on the scarring effects of graduating during a recession or experiencing a bad labour market entry (Gregg, 2001; Oreopoulos et al., 2012, Cockx and Ghirelli, 2016).

Nevertheless, our results on our newly developed decomposition approach also revealed that this overeducation penalty generates a misleading picture of the importance of overeducation in explaining the unconditional average wage return for education. This is due to this unconditional return being affected by the change in overeducation penalty and overeducation risk when investing in more education rather than by the level of the overeducation penalty and risk per se. In fact, with respect to obtaining a bachelor's degree, we even found some evidence that the unconditional average wage return for education exceeds the return that would have been realised in the absence of a mismatch. This is primarily due to our finding that obtaining a bachelor's degree (relative to obtaining a high-school degree only) may be a way to reduce one's risk of overeducation. Moreover, for master's degrees, the impact of overeducation on its unconditional return seemed to be moderate at best. Although we found master's degrees to be associated with a reduced match quality in the first job relative to bachelor's degrees (but not relative to high-school degrees), their unconditional average wage return was still estimated to be substantial. For instance, at age 29, we found their unconditional average return to be around 10% relative to a return of roughly 12% if there was no change in match quality. The unconditional average return for obtaining a high-school degree, meanwhile, was found to be much more limited, due to, among other things, the increased overeducation risk and penalty that is associated with obtaining such a degree relative to having obtained a lower secondary education degree only. Overall, these findings do not suggest that overeducation is indicative of considerable overinvestments in higher education. Instead, they are consistent with a Polarised labour market (cf. Autor et al., 2003; Goos et al., 2009) in which obtaining a higher education degree may be a viable way to avoid overeducation.

Even if the impact of overeducation on the average unconditional wage return of obtaining a higher education degree is moderate or even positive, this does not mean that obtaining such a degree is an efficient strategy for all individuals. Indeed, in line with several other studies (e.g. Arcidiacono, 2004; Rodriguez et al., 2016), we found the between-individual heterogeneity in this unconditional return to be substantial. Moreover, we found these unconditional returns to be smaller among those individuals who face a higher overeducation risk. Nonetheless, this heterogeneity in unconditional wage returns ultimately proved to be fairly similar to the heterogeneity in wage returns conditional on perfect matching. Overall, this suggests that, while differences in overeducation probabilities may reflect differences in unconditional wage returns across individuals (e.g. due to differences in abilities), overeducation in itself is not a channel that further reinforces this heterogeneity. Indeed, even if individuals have a higher likelihood of being overeducated, obtaining a college degree may still improve their chances of obtaining a medium-skilled job (cf. Verhaest et al., 2018) and, henceforth, generate a substantial wage return for these individuals. As an explanation for heterogeneous realised (ex-post) returns for education, meanwhile, overeducation was found to be much more important. By simulating a random matching process w.r.t. the first job based on the parameter estimates of our model, we found the wage returns conditional on this random matching to be negative for a substantial percentage of the graduates despite their unconditional (ex-ante) return being positive. This is consistent with overeducation being far more indicative of labour market frictions (cf. Gautier, 2002; Dolado et al., 2009) and, henceforth, investing in higher education being a risky venture (cf. Leuven and Oosterbeek, 2011).

These results have important policy implications. First, they suggest that reducing investments in higher education may not be the right answer to observations of widespread overeducation among young workers. On the contrary, widening access to bachelor's degree programmes may even be beneficial in this respect. Second, rather than viewing overeducation as being indicative of inefficient educational policies, our findings suggest that it would be much more fruitful to focus on labour market policies that reduce frictions. The reduction of these frictions may not only reduce one's risk of being overeducated at the initial stage of one's career but may also minimise the scarring effects that result from this initial labour market mismatch.

We end by indicating some directions for further research. First, our analysis was based on data covering the early nineties and the first years of the new century. Not only is this a period for which job polarisation has been well documented (Goos et al., 2009), but participation in higher education has only continued to increase since then. An analysis relying on more recent data would therefore be interesting. Second, the Belgian labour market is known to be relatively rigid. Besides being associated with stronger overeducation penalties (Levels et al., 2014), the context of a rigid labour market is also presumed to be associated with stronger scarring effects in the case of a bad labour market entry (Cockx and Ghirelli, 2016). Estimating a similar model to ours while relying on data from a more flexible labour market context would therefore provide another interesting avenue for further research. Finally, by focusing on obtaining a higher level of education, we only accounted for the quantitative dimension of additional investments in education. Several studies have shown overeducation to be correlated with the selectivity and prestige of the study programmes and institutions (Robst, 1995; Verhaest and van der Velden, 2013). It would therefore be interesting to extend our model by also accounting for this more qualitative dimension of investments in education.

References

Acemoglu, Daron (1998). “Why do new technologies complement skills? Directed technical change and wage inequality”. In: *The Quarterly Journal of Economics* 113 (4), pp. 1055–1089.

Adda, Jerome, Christian Dustmann, Costas Meghir, and Jean-Marc Robin (2010). “Career progression and formal versus on-the-job training”. In: *IFS Working Papers* 09 (06).

Agopsowicz, Andrew, Chris Robinson, Ralph Stinebrickner, and Todd Stinebrickner (2020). “Careers and mismatch for college graduates: college and noncollege jobs”. In: *Journal of Human Resources* 55 (4), pp. 1194–1221.

Allen, Jim and Rolf van der Velden (2001). “Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search”. In: *Oxford Economic Papers* 53 (3), pp. 434–452.

Arcidiacono, Peter (2004). “Ability sorting and the returns to college major”. In: *Journal of Econometrics* 121 (1-2), pp. 343–375.

— (2005). “Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?” In: *Econometrica* 73.5, pp. 1477–1524.

Arcidiacono, Peter and Paul B. Ellickson (2011). “Practical methods for estimation of dynamic discrete choice models”. In: *Annual Review of Economics* 3, pp. 363–394.

Arcidiacono, Peter and John Bailey Jones (2003). “Finite mixture distributions, sequential likelihood and the EM Algorithm”. In: *Econometrica* 71 (3), pp. 933–946.

Ashworth, Jared, V. Joseph Hotz, Arnaud Maurel, and Tyler Ransom (2021). “Changes across cohorts in wage returns to schooling and early work experiences”. In: *Journal of Labor Economics* 39 (4), pp. 931–964.

Autor, David H., Frank Levy, and Richard J. Murnane (2003). “The skill content of recent technological change: an empirical exploration”. In: *The Quarterly Journal of Economics* 118 (4), pp. 1279–1333.

Baert, Stijn, Bart Cockx, and Dieter Verhaest (2013). “Overeducation at the start of the career: Stepping stone or trap?” In: *Labour Economics* 25, pp. 123–140.

Baert, Stijn, Brecht Neyt, Eddy Omey, and Dieter Verhaest (2022). “Student work during secondary education, educational achievement, and later employment: a dynamic approach”. In: *Empirical Economics* 63 (3), pp. 1605–1635.

Barnichon, Regis and Yanos Zylberberg (2019). "Underemployment and the trickle-down of unemployment". In: *American Economic Journal: Macroeconomics* 11 (2), pp. 40–78.

Belzil, Christian and François Poinas (2010). "Education and early career outcomes of second-generation immigrants in France". In: *Labour Economics* 17 (1), pp. 101–110.

Cameron, Stephen V. and James J. Heckman (1998). "Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males". In: *Journal of Political Economy* 106 (2), pp. 262–333.

— (2001). "The dynamics of educational attainment for black, Hispanic, and white males". In: *Journal of Political Economy* 109 (3), pp. 455–499.

Carneiro, Pedro, James J. Heckman, and Edward J. Vytlacil (2011). "Estimating marginal returns to education". In: *American Economic Review* 101 (6), pp. 2754–2781.

Charalambidou, Christiana and Steven McIntosh (2021). "Over-education in Cyprus: micro and macro determinants, persistence and state dependence. A dynamic panel analysis". In: *The Manchester School* 89 (2), pp. 172–189.

Charlot, Olivier and Bruno Decreuse (2005). "Self-selection in education with matching frictions". In: *Labour Economics* 12 (2), pp. 251–267.

Chevalier, Arnaud (2003). "Measuring over-education". In: *Economica* 70 (279), pp. 509–531.

Chevalier, Arnaud and Joanne Lindley (2009). "Overeducation and the skills of UK graduates". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 172 (2), pp. 307–337.

Cockx, Bart and Corinna Ghirelli (2016). "Scars of recessions in a rigid labor market". In: *Labour Economics* 41, pp. 162–176.

Cockx, Bart, Matteo Picchio, and Stijn Baert (2019). "Modeling the effects of grade retention in high school". In: *Journal of Applied Econometrics* 34, pp. 403–424.

Colding, Bjorg (2006). "A dynamic analysis of educational progression of children of immigrants". In: *Labour Economics* 13 (4), pp. 479–492.

Davia, Maria A., Séamus McGuinness, and Philip J. O'Connell (2017). "Determinants of regional differences in rates of overeducation in Europe". In: *Social Science Research* 63, pp. 67–80.

De Grip, Andries, Hans Bosma, Dick Willems, and Martin van Boxtel (2008). "Job-worker mismatch and cognitive decline". In: *Oxford Economic Papers* 60 (2), pp. 237–253.

De Groote, Olivier (2022). "A dynamic model of effort choice in high school". In: *TSE Working Paper* 19-1002.

Declercq, Koen and Frank Verboven (2018). "Enrollment and degree completion in higher education without admission standards". In: *Economics of Education Review* 66, pp. 223–244.

Delaney, Judith, Séamus McGuinness, Konstantinos Pouliakas, and Paul Redmond (2020). "Educational expansion and overeducation of young graduates: A comparative analysis of 30 European countries". In: *Oxford Review of Education* 46 (1), pp. 10–29.

Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin (1977). "Maximum likelihood from incomplete data via the EM Algorithm". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 39 (1), pp. 1–38.

Di Cintio, Marco (2022). "Overeducation and R&D: Theoretical Aspects and Empirical Evidence". In: *Journal of Human Capital* 16 (2), pp. 183–232.

Dolado, Juan, Marcel Jansen, and Juan F. Jimeno (2009). "On-the-job search in a matching model with heterogeneous jobs and workers". In: *Economic Journal* 119 (534), pp. 200–228.

Dolton, Peter and Mary Silles (2008). "The effects of over-education on earnings in the graduate labour market". In: *Economics of Education Review* 27 (2), pp. 125–139.

Duncan, Greg J. and Saul D. Hoffman (1981). "The incidence and wage effects of overeducation". In: *Economics of Education Review* 1 (1), pp. 75–86.

Frenette, Marc (2004). "The overqualified Canadian graduate: the role of the academic program in the incidence, persistence, and economic returns to overqualification". In: *Economics of Education Review* 23 (1), pp. 29–45.

Gautier, Pieter A. (2002). "Unemployment and Search Externalities in a Model with Heterogeneous Jobs and Workers". In: *Economica* 69 (273), pp. 21–40.

Goldin, Claudia and Lawrence F. Katz (n.d.). *The Race between Education and Technology*. Harvard University Press.

Goos, Maarten, Alan Manning, and Anna Salomons (2009). "Job polarization in Europe". In: *American Economic Review* 99 (2), pp. 58–63.

Green, Francis and Steven McIntosh (2007). "Is there a genuine under-utilization of skills amongst the over-qualified?" In: *Applied Economics* 39 (4), pp. 427–439.

Green, Francis, Steven McIntosh, and Anna Vignoles (2002). "The utilization of education and skills: Evidence from Britain". In: *Manchester School* 70 (6), pp. 792–811.

Gregg, Paul (2001). "The impact of youth unemployment on adult unemployment in the NCDS". In: *The Economic Journal* 111 (475), pp. 626–653.

Groot, Wim and Henriëtte Maassen Van Den Brink (2000). "Overeducation in the labor market: a meta-analysis". In: *Economics of Education Review* 19 (2), pp. 149–158.

Gunderson, M. and P. Oreopoulos (2010). "Returns to Education in Developed Countries". In: *International Encyclopedia of Education (Third Edition)*. Ed. by Penelope Peterson, Eva Baker, and Barry McGaw. Third Edition. Oxford: Elsevier, pp. 298–304.

Hartog, Joop (2000). "Over-education and earnings: where are we, where should we go?" In: *Economics of Education Review* 19 (2), pp. 131–147.

Heckman, James J., John Eric Humphries, and Gregory Veramendi (2016). "Dynamic treatment effects". In: *Journal of Econometrics* 191 (2), pp. 276–292.

— (2018a). "Returns to education: The causal effects of education on earnings, health, and smoking". In: *Journal of Political Economy* 126, S197–S246.

— (2018b). "The Nonmarket Benefits of Education and Ability". In: *Journal of Human Capital* 12 (2), pp. 282–304.

Heckman, James J. and Salvador Navarro (2007). "Dynamic discrete choice and dynamic treatment effects". In: *Journal of Econometrics* 136 (2), pp. 341–396.

Heckman, James J. and Burton Singer (1984). "A method for minimizing the impact of distributional assumptions in econometric models for duration data". In: *Econometrica* 52 (2), pp. 271–320.

Heckman, James J., Jora Stixrud, and Sergio Urzua (2006). "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior". In: *Journal of Labor Economics* 24 (3), pp. 411–482.

Holzer, Harry J. (1987). "Job search by employed and unemployed youth". In: *Industrial and Labor Relations Review* 40 (4), pp. 601–611.

Hotz, V. Joseph, Lixin Colin Xu, Marta Tienda, and Avner Ahituv (2002). "Are there returns to the wages of young men from working while in school?" In: *The Review of Economics and Statistics* 84 (2), pp. 221–236.

Keane, Michael P., Petra E. Todd, and Kenneth I. Wolpin (2011). “The structural estimation of behavioral models: Discrete choice dynamic programming methods and applications”. In: *Handbook of Labor Economics, Vol 4A* 4, pp. 331–461.

Korpi, Tomas and Michael Tåhlin (2009). “Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974-2000”. In: *Labour Economics* 16 (2), pp. 183–193.

Lessaer, Bartek, Paolo Pasimeni, Konstantinos Pouliakas, and Mantas Sekmokas (2015). “Chapter III.1 Supporting skills development and matching in the EU”. In: *Employment and Social Developments in Europe*, pp. 229–273.

Leuven, Edwin and Hessel Oosterbeek (2011). “Overeducation and mismatch in the labor market”. In: *Handbook of the Economics of Education* , Chapter 3, vol. 4, pp. 283–326.

Levels, Mark, Rolf van der Velden, and Jim Allen (2014). “Educational mismatches and skills: new empirical tests of old hypotheses”. In: *Oxford Economic Papers* 66 (4), pp. 959–982.

Mavromaras, Kostas, Séamus Mcguinness, Nigel O’Leary, Peter Sloane, and Zhang Wei (2013). “Job mismatches and labour market outcomes: Panel evidence on university graduates”. In: *Economic Record* 89 (286), pp. 382–395.

Mavromaras, Kostas G., Séamus McGuinness, Nigel C. O’Leary, Peter J. Sloane, and Zhang Wei (2013). “Job mismatches and labour market outcomes: panel evidence on university graduates”. In: *Economic Record* 89 (286), pp. 382–395.

McCormick, Barry (1990). “A theory of signalling during job search, employment efficiency, and ”stigmatised” jobs”. In: *Review of Economic Studies* 57 (2), pp. 299–313.

McGuinness, Séamus (2006). “Overeducation in the labour market”. In: *Journal of Economic Surveys* 20 (3), pp. 387–418.

— (2008). “How biased are the estimated wage impacts of overeducation? A propensity score matching approach”. In: *Applied Economics Letters* 15 (2), pp. 145–149.

McGuinness, Séamus, Konstantinos Pouliakas, and Paul Redmond (2018). “Skills mismatch: concepts, measurement and policy approaches”. In: *Journal of Economic Surveys* 32 (4), pp. 985–1015.

McGuinness, Séamus and Peter Sloane (2011). “Labour market mismatch among UK graduates: An analysis using REFLEX data”. In: *Economics of Education Review* 30 (1), pp. 130–145.

Meroni, Elena Claudia and Esperanza Vera-Toscano (2017). "The persistence of overeducation among recent graduates". In: *Labour Economics* 48, pp. 120–143.

Neyt, Brecht, Dieter Verhaest, Lorenzo Navarini, and Stijn Baert (2022). "The impact of internship experience on schooling and labour market outcomes". In: *CESifo Economic Studies* 68 (2), pp. 127–154.

Ordine, Patrizia and Giuseppe Rose (2017). "Too Many Graduates? A Matching Theory of Educational Mismatch". In: *Journal of Human Capital* 11 (4), pp. 423–446.

Oreopoulos, Philip and Kjell G. Salvanes (2011). "Priceless: the nonpecuniary benefits of schooling". In: *Journal of Economic Perspectives* 25 (1), pp. 159–84.

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz (2012). "The Short- and Long-Term Career Effects of Graduating in a Recession". In: *American Economic Journal: Applied Economics* 4 (1), pp. 1–29.

Pissarides, Christopher A. (1994). "Search Unemployment with On-the-job Search". In: *The Review of Economic Studies* 61 (3), pp. 457–475.

Robst, John (1995). "College quality and overeducation". In: *Economics of Education Review* 14 (3), pp. 221–228.

Rodríguez, Jorge, Sergio Urzúa, and Loreto Reyes (2016). "Heterogeneous economic returns to post-secondary degrees: evidence from Chile". In: *The Journal of Human Resources* 51 (2), pp. 416–460.

Roller, Christiane, Christian Rulff, and Michael M. Tamminga (2020). "It's a mismatch! Overeudcation and career mobility in Germany". In: *German Economic Review* 21 (4), pp. 493–514.

Rubb, Stephen (2006). "Educational Mismatches and Earnings: Extensions of Occupational Mobility Theory and Evidence of Human Capital Depreciation". In: *Education Economics* 14 (2), pp. 135–154.

Sloane, Peter J., Harminder Battu, and Paul T. Seaman (1999). "Overeducation, undereducation and the British labour market". In: *Applied Economics* 31 (11), pp. 1437–1453.

Thurow, Lester C. (1975). *Generating inequality : mechanisms of distribution in the U. S. economy*. New York (N.Y.) : Basic books, 1975., p. 258.

Verhaest, Dieter, Elene Bogaert, Jeroen Dereymaeker, Laura Mestdagh, and Stijn Baert (2018). "Do employers prefer overqualified graduates? a field experiment". In: *Industrial Relations* 57 (3), pp. 361–388.

Verhaest, Dieter and Eddy Omey (2006). "The impact of overeducation and its measurement". In: *Social Indicators Research* 77 (3), pp. 419–448.

— (2009). "Objective over-education and worker well-being: A shadow price approach". In: *Journal of Economic Psychology* 30 (3), pp. 469–481.

— (2012). "Overeducation, undereducation and earnings: further evidence on the importance of ability and measurement error bias". In: *Journal of Labor Research* 33 (1), pp. 76–90.

Verhaest, Dieter and Rolf Van Der Velden (2013). "Cross-country differences in graduate overeducation". In: *European Sociological Review* 29 (3), pp. 642–653.

A Data

Table A1: Missing values breakdown

Total number of individuals in SONAR	**9000**
Individuals with > 2 years delay prior to primary education	76
Individuals in special needs schools	124
Inconsistent, erroneous or incomplete data on exogenous variables and educational career	638
Final sample educational outcomes	**8162**
No information on first job	701
No information on overeducation	250
Final sample overeducation start first job	**7211**
Still in education or no job at age 23	1519
Surveyed, but no wage questions at age 23	1145
Non-response or outliers wage age 23	333
Final sample wages at age 23	**4214**
Not surveyed at age 26	3686
Still in education or no job at age 26	84
Surveyed, but no wage questions at age 26	79
Non-response or outliers wage age 26	116
Final sample wages at age 26	**3246**
Not surveyed at age 29	4030
Still in education or no job at age 29	42
Surveyed, but no wage questions at age 29	45
Non-response or outliers wage age 29	38
Final sample wages at age 29	**3056**

B Treatment effects tables

Table A1: Wage returns treatment effects: Direct effects

		Direct effects											
		ATE†			ATE			ATT			ATNT		
		HSE	LTE	HTE	HSE	LTE	HTE	HSE	LTE	HTE	HSE	LTE	HTE
Wage 23	AM-AM	0.047*	0.067***	0.103***	0.034	0.074***	0.103***	0.034	0.073***	0.123***	0.031	0.074***	0.090***
		(0.027)	(0.012)	(0.020)	(0.021)	(0.011)	(0.018)	(0.021)	(0.014)	(0.017)	(0.022)	(0.011)	(0.021)
	AM-OE	0.014	0.038**	-0.010	0.001	0.045***	-0.010	0.001	0.044***	0.010	-0.002	0.045***	-0.022
		(0.027)	(0.015)	(0.024)	(0.021)	(0.014)	(0.023)	(0.021)	(0.017)	(0.024)	(0.022)	(0.014)	(0.024)
	OE-AM	0.042	0.100***	0.132***	0.028	0.106***	0.132***	0.028	0.106***	0.151***	0.025	0.107***	0.119***
		(0.032)	(0.012)	(0.023)	(0.026)	(0.011)	(0.021)	(0.026)	(0.014)	(0.020)	(0.027)	(0.010)	(0.024)
	OE-OE	0.009	0.071***	0.019	-0.005	0.078***	0.019	-0.005	0.077***	0.039	-0.008	0.078***	0.006
		(0.032)	(0.015)	(0.026)	(0.026)	(0.014)	(0.024)	(0.026)	(0.017)	(0.025)	(0.026)	(0.014)	(0.026)
	Unconditional WR	0.029	0.077***	0.063***	0.015	0.083***	0.066***	0.015	0.083***	0.091***	0.012	0.083***	0.050***
		(0.027)	(0.011)	(0.018)	(0.020)	(0.010)	(0.015)	(0.020)	(0.013)	(0.015)	(0.021)	(0.009)	(0.018)
	Unconditional WR (Dir.)	0.029	0.077***	0.063***	0.015	0.083***	0.066***	0.015	0.083***	0.091***	0.012	0.083***	0.050***
		(0.027)	(0.011)	(0.018)	(0.020)	(0.010)	(0.015)	(0.020)	(0.013)	(0.015)	(0.021)	(0.009)	(0.018)
Wage 26	AM-AM	0.073**	0.087***	0.119***	0.072***	0.092***	0.126***	0.073***	0.085***	0.138***	0.067**	0.099***	0.116***
		(0.033)	(0.012)	(0.014)	(0.026)	(0.011)	(0.013)	(0.026)	(0.014)	(0.013)	(0.028)	(0.010)	(0.015)
	AM-OE	0.057*	0.051***	0.048***	0.057**	0.056***	0.054***	0.057**	0.049***	0.067***	0.051*	0.064***	0.045**
		(0.033)	(0.015)	(0.017)	(0.026)	(0.014)	(0.017)	(0.026)	(0.017)	(0.019)	(0.028)	(0.014)	(0.018)
	OE-AM	0.014	0.103***	0.155***	0.013	0.108***	0.161***	0.014	0.101***	0.174***	0.008	0.115***	0.152***
		(0.038)	(0.012)	(0.017)	(0.032)	(0.011)	(0.016)	(0.032)	(0.014)	(0.016)	(0.034)	(0.010)	(0.018)
	OE-OE	-0.002	0.067***	0.084***	-0.002	0.072***	0.090***	-0.002	0.065***	0.103***	-0.008	0.080***	0.080***
		(0.038)	(0.015)	(0.019)	(0.032)	(0.014)	(0.019)	(0.032)	(0.017)	(0.021)	(0.034)	(0.014)	(0.020)
	Unconditional WR	0.051	0.087***	0.101***	0.050**	0.091***	0.109***	0.050**	0.085***	0.125***	0.044	0.098***	0.098***
		(0.032)	(0.011)	(0.012)	(0.025)	(0.010)	(0.011)	(0.025)	(0.013)	(0.012)	(0.027)	(0.009)	(0.013)
	Unconditional WR (Dir.)	0.051	0.087***	0.101***	0.050**	0.091***	0.109***	0.050**	0.085***	0.125***	0.044	0.098***	0.098***
		(0.032)	(0.011)	(0.012)	(0.025)	(0.010)	(0.011)	(0.025)	(0.013)	(0.012)	(0.027)	(0.009)	(0.013)
Wage 29	AM-AM	0.044	0.053***	0.112***	0.017	0.064***	0.121***	0.018	0.041***	0.137***	0.014	0.087***	0.110***
		(0.033)	(0.012)	(0.012)	(0.026)	(0.011)	(0.011)	(0.026)	(0.015)	(0.012)	(0.028)	(0.009)	(0.012)
	AM-OE	0.008	0.018	0.027**	-0.019	0.028**	0.037***	-0.019	0.006	0.053***	-0.022	0.052***	0.025*
		(0.033)	(0.015)	(0.013)	(0.026)	(0.014)	(0.013)	(0.026)	(0.017)	(0.014)	(0.028)	(0.013)	(0.014)
	OE-AM	0.069*	0.090***	0.147***	0.042	0.100***	0.157***	0.042	0.078***	0.173***	0.039	0.124***	0.145***
		(0.037)	(0.013)	(0.014)	(0.030)	(0.011)	(0.014)	(0.030)	(0.016)	(0.014)	(0.031)	(0.010)	(0.015)
	OE-OE	0.032	0.054***	0.063***	0.005	0.065***	0.072***	0.005	0.043**	0.089***	0.002	0.088***	0.061***
		(0.036)	(0.015)	(0.015)	(0.029)	(0.014)	(0.014)	(0.029)	(0.018)	(0.015)	(0.030)	(0.013)	(0.015)
	Unconditional WR	0.033	0.063***	0.088***	0.006	0.074***	0.099***	0.006	0.052***	0.119***	0.003	0.096***	0.085***
		(0.032)	(0.011)	(0.010)	(0.024)	(0.010)	(0.009)	(0.024)	(0.015)	(0.010)	(0.026)	(0.008)	(0.011)
	Unconditional WR (Dir.)	0.033	0.063***	0.088***	0.006	0.074***	0.099***	0.006	0.052***	0.119***	0.003	0.096***	0.085***
		(0.032)	(0.011)	(0.010)	(0.024)	(0.010)	(0.009)	(0.024)	(0.015)	(0.010)	(0.026)	(0.008)	(0.011)

Notes: we simplify the notation and we refer to: (i) Wage returns AM-AM as $\Omega^{M,M}_{a,j}$, (ii) Wage returns AM-OE as $\Omega^{M,O}_{a,j}$, (iii) Wage returns OE-AM as $\Omega^{O,M}_{a,j}$ and (iv) Wage returns OE-OE as $\Omega^{O,O}_{a,j}$. Moreover, Unconditional WR refers to Unconditional Wage Returns and Unconditional WR (Dir.) refers to the Unconditional Wage Return computed as the direct effect.

Table A2: Wage returns treatment effects: Total effects

		Total effects											
		ATE†			ATE			ATT			ATNT		
		HSE	LTE	HTE	HSE	LTE	HTE	HSE	LTE	HTE	HSE	LTE	HTE
Wage 23	AM-AM	0.094***	0.093***	0.103***	0.038*	0.079***	0.103***	0.037*	0.073***	0.123***	0.051**	0.086***	0.090***
		(0.025)	(0.012)	(0.020)	(0.021)	(0.011)	(0.018)	(0.021)	(0.014)	(0.017)	(0.022)	(0.011)	(0.021)
	AM-OE	0.052**	0.048***	-0.010	0.004	0.047***	-0.010	0.003	0.044***	0.010	0.017	0.049***	-0.022
		(0.026)	(0.014)	(0.024)	(0.021)	(0.014)	(0.023)	(0.021)	(0.017)	(0.024)	(0.022)	(0.013)	(0.024)
	OE-AM	0.089***	0.126***	0.132***	0.032	0.112***	0.132***	0.031	0.106***	0.151***	0.046*	0.118***	0.119***
		(0.030)	(0.012)	(0.023)	(0.026)	(0.011)	(0.021)	(0.026)	(0.014)	(0.020)	(0.027)	(0.010)	(0.024)
	OE-OE	0.047	0.081***	0.019	-0.001	0.080***	0.019	-0.002	0.077***	0.039	0.011	0.082***	0.006
		(0.030)	(0.014)	(0.026)	(0.026)	(0.014)	(0.024)	(0.026)	(0.017)	(0.025)	(0.027)	(0.013)	(0.026)
	Unconditional WR	0.077***	0.096***	0.063***	0.019	0.087***	0.066***	0.018	0.083***	0.091***	0.033	0.091***	0.050***
		(0.025)	(0.011)	(0.018)	(0.020)	(0.010)	(0.015)	(0.020)	(0.013)	(0.015)	(0.022)	(0.009)	(0.018)
	Unconditional WR (Dir.)	0.029	0.077***	0.063***	0.015	0.083***	0.066***	0.015	0.083***	0.091***	0.012	0.083***	0.050***
		(0.027)	(0.011)	(0.018)	(0.020)	(0.010)	(0.015)	(0.020)	(0.013)	(0.015)	(0.021)	(0.009)	(0.018)
Wage 26	AM-AM	0.147***	0.127***	0.119***	0.080***	0.101***	0.126***	0.078***	0.085***	0.138***	0.097***	0.117***	0.116***
		(0.031)	(0.011)	(0.014)	(0.026)	(0.011)	(0.013)	(0.026)	(0.014)	(0.013)	(0.029)	(0.010)	(0.015)
	AM-OE	0.113***	0.081***	0.048***	0.062**	0.063***	0.054***	0.061**	0.049***	0.067***	0.075**	0.077***	0.045**
		(0.032)	(0.014)	(0.017)	(0.026)	(0.014)	(0.017)	(0.026)	(0.017)	(0.019)	(0.029)	(0.013)	(0.018)
	OE-AM	0.088**	0.143***	0.155***	0.020	0.116***	0.161***	0.019	0.101***	0.174***	0.038	0.133***	0.152***
		(0.036)	(0.012)	(0.017)	(0.032)	(0.011)	(0.016)	(0.032)	(0.014)	(0.016)	(0.034)	(0.010)	(0.018)
	OE-OE	0.053	0.097***	0.084***	0.003	0.079***	0.090***	0.002	0.065***	0.103***	0.016	0.093***	0.080***
		(0.036)	(0.014)	(0.019)	(0.032)	(0.014)	(0.019)	(0.032)	(0.017)	(0.021)	(0.034)	(0.013)	(0.020)
	Unconditional WR	0.122***	0.122***	0.101***	0.057**	0.099***	0.109***	0.056**	0.085***	0.125***	0.073**	0.114***	0.098***
		(0.031)	(0.011)	(0.012)	(0.025)	(0.010)	(0.011)	(0.025)	(0.013)	(0.012)	(0.029)	(0.009)	(0.013)
	Unconditional WR (Dir.)	0.051	0.087***	0.101***	0.050**	0.091***	0.109***	0.050**	0.085***	0.125***	0.044	0.098***	0.098***
		(0.032)	(0.011)	(0.012)	(0.025)	(0.010)	(0.011)	(0.025)	(0.013)	(0.012)	(0.027)	(0.009)	(0.013)
Wage 29	AM-AM	0.086***	0.096***	0.112***	0.020	0.073***	0.121***	0.018	0.041***	0.137***	0.040	0.106***	0.110***
		(0.031)	(0.012)	(0.012)	(0.026)	(0.011)	(0.011)	(0.026)	(0.015)	(0.012)	(0.028)	(0.009)	(0.012)
	AM-OE	0.040	0.047***	0.027**	-0.017	0.034**	0.037***	-0.019	0.006	0.053***	0.001	0.064***	0.025*
		(0.031)	(0.014)	(0.013)	(0.026)	(0.013)	(0.013)	(0.026)	(0.017)	(0.014)	(0.027)	(0.012)	(0.014)
	OE-AM	0.111***	0.133***	0.147***	0.044	0.110***	0.157***	0.043	0.078***	0.173***	0.064**	0.142***	0.145***
		(0.034)	(0.012)	(0.014)	(0.029)	(0.011)	(0.014)	(0.029)	(0.016)	(0.014)	(0.031)	(0.010)	(0.015)
	OE-OE	0.064*	0.084***	0.063***	0.007	0.071***	0.072***	0.005	0.043**	0.089***	0.026	0.100***	0.061***
		(0.034)	(0.014)	(0.015)	(0.029)	(0.014)	(0.014)	(0.029)	(0.018)	(0.015)	(0.030)	(0.012)	(0.015)
	Unconditional WR	0.077***	0.101***	0.088***	0.008	0.081***	0.099***	0.007	0.052***	0.119***	0.029	0.112***	0.085***
		(0.030)	(0.011)	(0.010)	(0.024)	(0.010)	(0.009)	(0.024)	(0.015)	(0.010)	(0.026)	(0.008)	(0.011)
	Unconditional WR (Dir.)	0.033	0.063***	0.088***	0.006	0.074***	0.099***	0.006	0.052***	0.119***	0.003	0.096***	0.085***
		(0.032)	(0.011)	(0.010)	(0.024)	(0.010)	(0.009)	(0.024)	(0.015)	(0.010)	(0.026)	(0.008)	(0.011)

Notes: we simplify the notation and we refer to: (i) Wage returns AM-AM as $\Omega_{a,j}^{M,M}$, (ii) Wage returns AM-OE as $\Omega_{a,j}^{M,O}$, (iii) Wage returns OE-AM as $\Omega_{a,j}^{O,M}$ and (iv) Wage returns OE-OE as $\Omega_{a,j}^{O,O}$. Moreover, Unconditional WR refers to Unconditional Wage Returns and Unconditional WR (Dir.) refers to the Unconditional Wage Return computed as the direct effect.

Table A3: Overeducation wage penalty

	Direct effects											
	ATE†			ATE			ATT			ATNT		
	Wage 23	Wage 26	Wage 29	Wage 23	Wage 26	Wage 29	Wage 23	Wage 26	Wage 29	Wage 23	Wage 26	Wage 29
HSE	-0.033***	-0.016*	-0.037***	-0.033***	-0.016*	-0.037***	-0.033***	-0.016*	-0.037***	-0.033***	-0.016	-0.037***
	(0.008)	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)	(0.008)	(0.011)	(0.010)
LTE	-0.029**	-0.035***	-0.035***	-0.029**	-0.035***	-0.035***	-0.029**	-0.036***	-0.035***	-0.029**	-0.035***	-0.036***
	(0.013)	(0.013)	(0.011)	(0.013)	(0.013)	(0.011)	(0.013)	(0.013)	(0.011)	(0.013)	(0.013)	(0.011)
HTE	-0.113***	-0.071***	-0.084***	-0.113***	-0.071***	-0.084***	-0.113***	-0.071***	-0.084***	-0.113***	-0.071***	-0.084***
	(0.026)	(0.019)	(0.013)	(0.026)	(0.019)	(0.013)	(0.026)	(0.019)	(0.013)	(0.026)	(0.019)	(0.013)

	Total effects											
	ATE†			ATE			ATT			ATNT		
	Wage 23	Wage 26	Wage 29	Wage 23	Wage 26	Wage 29	Wage 23	Wage 26	Wage 29	Wage 23	Wage 26	Wage 29
HSE	-0.042***	-0.034***	-0.047***	-0.033***	-0.017**	-0.037***	-0.033***	-0.017**	-0.037***	-0.034***	-0.022**	-0.039***
	(0.007)	(0.007)	(0.006)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.010)	(0.008)
LTE	-0.045***	-0.046***	-0.049***	-0.032***	-0.038***	-0.039***	-0.029**	-0.036***	-0.035***	-0.036***	-0.040***	-0.042***
	(0.012)	(0.011)	(0.009)	(0.012)	(0.012)	(0.011)	(0.013)	(0.013)	(0.011)	(0.012)	(0.011)	(0.010)
HTE	-0.113***	-0.071***	-0.084***	-0.113***	-0.071***	-0.084***	-0.113***	-0.071***	-0.084***	-0.113***	-0.071***	-0.084***
	(0.026)	(0.019)	(0.013)	(0.026)	(0.019)	(0.013)	(0.026)	(0.019)	(0.013)	(0.026)	(0.019)	(0.013)

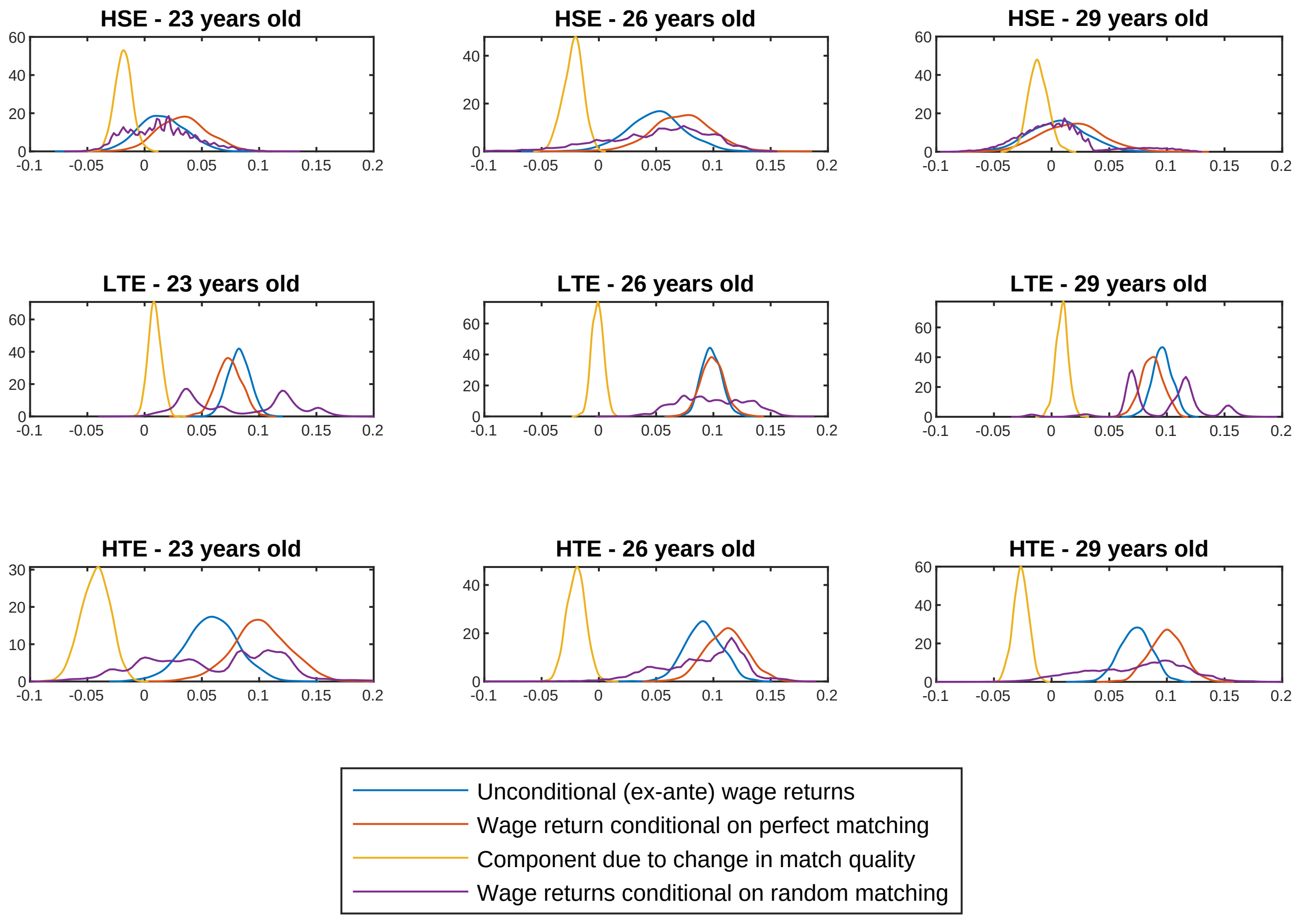
Table A4: Decomposition wage returns

		Direct effects						Total effects					
		ATE† HSE	LTE	HTE	ATE HSE	LTE	HTE	ATE† HSE	LTE	HTE	ATE HSE	LTE	HTE
	Unconditional WR	0.029	0.077***	0.063***	0.015	0.083***	0.066***	0.077***	0.096***	0.063***	0.019	0.087***	0.066***
		(0.027)	(0.011)	(0.018)	(0.020)	(0.010)	(0.015)	(0.025)	(0.011)	(0.018)	(0.020)	(0.010)	(0.015)
	AM-AM	0.047*	0.067***	0.103***	0.034	0.074***	0.103***	0.094***	0.093***	0.103***	0.038*	0.079***	0.103***
		(0.027)	(0.012)	(0.020)	(0.021)	(0.011)	(0.018)	(0.025)	(0.012)	(0.020)	(0.021)	(0.011)	(0.018)
	Difference	-0.018**	0.010*	-0.040***	-0.019**	0.010*	-0.037***	-0.017***	0.003	-0.040***	-0.018**	0.008	-0.037***
		(0.007)	(0.006)	(0.012)	(0.007)	(0.006)	(0.011)	(0.007)	(0.006)	(0.012)	(0.007)	(0.006)	(0.011)
Wage 23	Change in match quality	-0.019**	0.010*	-0.040***	-0.018**	0.009*	-0.036***	-0.017***	0.002	-0.040***	-0.018**	0.007	-0.036***
		(0.007)	(0.006)	(0.012)	(0.007)	(0.005)	(0.011)	(0.007)	(0.003)	(0.012)	(0.007)	(0.005)	(0.011)
	Change in overeducation penalty	-0.011	0.002	-0.022***	-0.011	0.002	-0.022***	-0.013*	-0.001	-0.022***	-0.011	0.000	-0.022***
		(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	(0.008)	(0.007)	(0.003)	(0.007)	(0.007)	(0.007)	(0.008)
	Change in overeducation risk	-0.008***	0.008**	-0.018***	-0.008**	0.007**	-0.014***	-0.004	0.003***	-0.018***	-0.007**	0.007**	-0.014***
		(0.003)	(0.003)	(0.005)	(0.003)	(0.003)	(0.004)	(0.003)	(0.001)	(0.005)	(0.003)	(0.003)	(0.004)
	Unconditional WR	0.051	0.087***	0.101***	0.050**	0.091***	0.109***	0.122***	0.122***	0.101***	0.057**	0.099***	0.109***
		(0.032)	(0.011)	(0.012)	(0.025)	(0.010)	(0.011)	(0.031)	(0.011)	(0.012)	(0.025)	(0.010)	(0.011)
	AM-AM	0.073**	0.087***	0.119***	0.072***	0.092***	0.126***	0.147***	0.127***	0.119***	0.080***	0.101***	0.126***
		(0.033)	(0.012)	(0.014)	(0.026)	(0.011)	(0.013)	(0.031)	(0.011)	(0.014)	(0.026)	(0.011)	(0.013)
	Diff	-0.022**	-0.000	-0.018**	-0.022***	-0.001	-0.016**	-0.025***	-0.004	-0.018**	-0.023***	-0.002	-0.016**
		(0.009)	(0.005)	(0.007)	(0.009)	(0.005)	(0.007)	(0.008)	(0.005)	(0.007)	(0.008)	(0.005)	(0.007)
Wage 26	Change in match quality	-0.022**	-0.000	-0.018**	-0.022***	-0.001	-0.016**	-0.025***	-0.002	-0.018**	-0.023***	-0.002	-0.016**
		(0.009)	(0.005)	(0.007)	(0.008)	(0.005)	(0.007)	(0.008)	(0.002)	(0.007)	(0.008)	(0.005)	(0.007)
	Change in overeducation penalty	-0.019**	-0.009	-0.008	-0.019**	-0.015	-0.008	-0.023***	-0.003*	-0.008	-0.019**	-0.009	-0.008
		(0.008)	(0.016)	(0.006)	(0.008)	(0.017)	(0.006)	(0.009)	(0.002)	(0.006)	(0.008)	(0.023)	(0.006)
	Change in overeducation risk	-0.004	0.009	-0.010***	-0.003	0.014	-0.008**	-0.002	0.001	-0.010***	-0.003	0.008	-0.008**
		(0.003)	(0.015)	(0.004)	(0.002)	(0.017)	(0.004)	(0.003)	(0.001)	(0.004)	(0.002)	(0.023)	(0.004)
	Unconditional WR	0.033	0.063***	0.088***	0.006	0.074***	0.099***	0.077***	0.101***	0.088***	0.008	0.081***	0.099***
		(0.032)	(0.011)	(0.010)	(0.024)	(0.010)	(0.009)	(0.030)	(0.011)	(0.010)	(0.024)	(0.010)	(0.009)
	AM-AM	0.044	0.053***	0.112***	0.017	0.064***	0.121***	0.086***	0.096***	0.112***	0.020	0.073***	0.121***
		(0.033)	(0.012)	(0.012)	(0.026)	(0.011)	(0.011)	(0.031)	(0.012)	(0.012)	(0.026)	(0.011)	(0.011)
	Diff	-0.012	0.010*	-0.024***	-0.012	0.010*	-0.022***	-0.010	0.004	-0.024***	-0.011	0.008	-0.022***
		(0.008)	(0.005)	(0.006)	(0.008)	(0.005)	(0.006)	(0.007)	(0.005)	(0.006)	(0.008)	(0.005)	(0.006)
Wage 29	Change in match quality	-0.012	0.010*	-0.024***	-0.012	0.009*	-0.022***	-0.010	0.002	-0.024***	-0.011	0.008	-0.022***
		(0.008)	(0.005)	(0.006)	(0.008)	(0.005)	(0.006)	(0.007)	(0.002)	(0.006)	(0.008)	(0.005)	(0.006)
	Change in overeducation penalty	-0.003	0.001	-0.012***	-0.003	0.001	-0.012***	-0.006	-0.001	-0.012***	-0.003	-0.001	-0.012***
		(0.008)	(0.007)	(0.004)	(0.009)	(0.007)	(0.004)	(0.008)	(0.002)	(0.004)	(0.009)	(0.006)	(0.004)
	Change in overeducation risk	-0.009**	0.009***	-0.012***	-0.008	0.009***	-0.010***	-0.004	0.003***	-0.012***	-0.008	0.008***	-0.010***
		(0.004)	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)	(0.004)	(0.001)	(0.003)	(0.005)	(0.003)	(0.003)

Notes: The measures used in the decomposition are the following: unconditional (ex-ante) wage returns (Unconditional WR), wage return conditional on perfect matching (WR conditional on PM).

B.1 Decomposition of change in match quality graphs

Figure B1: Decomposition of change in match quality



C Model estimates

Table C1: Model estimates

		BM - K3		
			S.E.	p
Delay Primary Education	Female	-0.006	0.182	0.975
	Siblings	0.059	0.052	0.256
	Foreign origin	1.551	0.268	0.000
	Education Father	-0.008	0.034	0.808
	Education Mother	0.014	0.031	0.655
	Birth day / 100	0.443	0.094	0.000
	Unemployment Delay	0.023	0.032	0.478
	Cohort 1978	-0.010	0.275	0.972
	Cohort 1980	0.294	0.246	0.232
	_cons	-5.473	0.524	0.000
	Het par 1	-0.481	0.194	0.013
	Het par 2	-0.161	0.360	0.655
Delay Secondary Education	Female	-0.274	0.079	0.001
	Siblings	0.094	0.025	0.000
	Foreign origin	0.822	0.135	0.000
	Education Father	-0.113	0.015	0.000
	Education Mother	-0.070	0.014	0.000
	Birth day / 100	0.298	0.040	0.000
	Unemployment Delay	-0.001	0.020	0.969
	Delay	3.048	0.215	0.000
	Cohort 1978	0.258	0.120	0.031
	Cohort 1980	0.236	0.114	0.038
	_cons	-1.927	0.268	0.000
	Het par 1	-0.358	0.085	0.000
	Het par 2	-0.340	0.172	0.048

Table C1: Model estimates

	Female	0.410	0.050	0.000
	Siblings	-0.087	0.019	0.000
	Foreign origin	0.171	0.125	0.171
	Education Father	0.125	0.009	0.000
	Education Mother	0.145	0.008	0.000
Start	Birth day / 100	-0.077	0.025	0.002
and track choice	Unemployment Delay	-0.025	0.011	0.027
Secondary	Delay	0.960	0.224	0.000
Education	Delay SE	-1.975	0.109	0.000
3rd year	Cohort 1978	-0.168	0.070	0.017
	Cohort 1980	-0.287	0.079	0.000
	Het par 1	0.423	0.057	0.000
	Het par 2	0.190	0.108	0.078
	cut 1	-4.295	0.181	0.000
	cut 2	1.174	0.130	0.000
	Female	0.660	0.130	0.000
	Siblings	-0.114	0.034	0.001
	Foreign origin	-0.037	0.211	0.862
	Education Father	0.068	0.025	0.008
	Education Mother	0.116	0.025	0.000
	Birth day / 100	0.004	0.062	0.946
Lower	Unemployment Delay	-0.038	0.028	0.183
Secondary	Delay	-0.420	0.339	0.215
Education	Delay SE	-0.713	0.144	0.000
	Track Choice LSE	1.905	0.219	0.000
	Cohort 1978	0.198	0.190	0.299
	Cohort 1980	0.141	0.178	0.429
	_cons	2.011	0.326	0.000
	Het par 1	0.418	0.131	0.001
	Het par 2	0.016	0.233	0.944

Table C1: Model estimates

Start and track choice Secondary Education 5th year	Female	0.347	0.076	0.000
	Siblings	-0.039	0.029	0.179
	Foreign origin	-0.072	0.185	0.696
	Education Father	0.070	0.014	0.000
	Education Mother	0.065	0.013	0.000
	Birth day / 100	0.018	0.038	0.644
	Unemployment Delay	-0.057	0.018	0.001
	Delay	-0.348	0.285	0.222
	Delay SE	-0.591	0.156	0.000
	Track Choice LSE	6.115	0.181	0.000
	Cohort 1978	0.244	0.097	0.012
	Cohort 1980	0.104	0.094	0.269
	Het par 1	0.389	0.086	0.000
	Het par 2	0.181	0.162	0.263
	cut 1	-2.978	0.248	0.000
	cut 2	5.218	0.292	0.000
Higher Secondary Education	Female	0.706	0.111	0.000
	Siblings	-0.081	0.033	0.015
	Foreign origin	-0.600	0.190	0.002
	Education Father	0.046	0.020	0.025
	Education Mother	0.065	0.020	0.001
	Birth day / 100	0.115	0.053	0.030
	Unemployment Delay	-0.035	0.024	0.144
	Delay	-0.266	0.340	0.434
	Delay SE	-0.594	0.139	0.000
	Track Choice LSE	0.147	0.177	0.407
	Track Choice HSE	1.311	0.217	0.000
	Cohort 1978	-0.161	0.139	0.248
	Cohort 1980	-0.157	0.133	0.240
	_cons	1.707	0.319	0.000
	Het par 1	0.727	0.111	0.000

Table C1: Model estimates

	Het par 2	0.302	0.209	0.149
	Female	0.129	0.049	0.008
	Siblings	-0.036	0.020	0.079
	Foreign origin	-0.120	0.140	0.391
	Education Father	0.062	0.009	0.000
	Education Mother	0.090	0.008	0.000
	Birth day / 100	0.021	0.025	0.391
Start	Unemployment Delay	-0.009	0.011	0.404
and track choice	Delay	0.404	0.231	0.080
Lower	Delay SE	-0.721	0.105	0.000
Tertiary	Track Choice LSE	1.285	0.095	0.000
Education	Track Choice HSE	2.238	0.106	0.000
	Cohort 1978	0.150	0.071	0.035
	Cohort 1980	0.156	0.069	0.024
	Het par 1	0.494	0.058	0.000
	Het par 2	0.032	0.109	0.767
	cut 1	1.354	0.136	0.000
	cut 2	5.212	0.158	0.000
	Female	0.550	0.071	0.000
	Siblings	-0.015	0.030	0.625
	Foreign origin	-0.808	0.202	0.000
	Education Father	0.038	0.013	0.003
	Education Mother	0.032	0.012	0.007
	Birth day / 100	-0.007	0.036	0.845
	Unemployment Delay	-0.057	0.016	0.000
Lower	Delay	-0.257	0.323	0.426
Tertiary	Delay SE	-0.296	0.165	0.074
Education	Track Choice LSE	0.075	0.112	0.506
	Track Choice HSE	1.117	0.112	0.000
	Track Choice LTE	0.650	0.098	0.000
	Cohort 1978	0.382	0.106	0.000
	Cohort 1980	0.259	0.102	0.011

Table C1: Model estimates

	_cons	-0.352	0.200	0.079
	Het par 1	0.581	0.082	0.000
	Het par 2	0.035	0.153	0.817
Start and track choice Higher Tertiary Education	Female	-0.518	0.076	0.000
	Siblings	-0.025	0.036	0.490
	Foreign origin	0.147	0.288	0.609
	Education Father	0.051	0.013	0.000
	Education Mother	0.049	0.012	0.000
	Birth day / 100	0.059	0.038	0.126
	Unemployment Delay	-0.031	0.017	0.071
	Delay	0.708	0.411	0.085
	Delay SE	-0.346	0.280	0.217
	Track Choice LSE	-0.185	0.206	0.369
	Track Choice HSE	1.083	0.185	0.000
	Track Choice LTE	2.854	0.086	0.000
	Cohort 1978	0.101	0.096	0.297
	Cohort 1980	0.019	0.096	0.843
	Het par 1	0.500	0.100	0.000
	Het par 2	1.030	0.173	0.000
	cut 1	2.544	0.258	0.000
	cut 2	3.647	0.263	0.000
Higher Tertiary Education	Female	0.296	0.167	0.077
	Siblings	-0.100	0.075	0.183
	Foreign origin	-0.710	0.543	0.192
	Education Father	-0.032	0.030	0.281
	Education Mother	-0.060	0.029	0.036
	Birth day / 100	0.166	0.085	0.051
	Unemployment Delay	0.052	0.037	0.160
	Delay	-0.090	0.791	0.909
	Delay SE	-1.481	0.617	0.016
	Track Choice LSE	-0.283	0.857	0.741
	Track Choice HSE	-0.715	0.730	0.328

Table C1: Model estimates

	Track Choice LTE	-0.219	0.305	0.473
	Track Choice HTE	0.058	0.296	0.845
	Cohort 1978	0.097	0.203	0.634
	Cohort 1980	0.323	0.211	0.126
	_cons	10.302	6.187	0.096
	Het par 1	-7.334	6.144	0.233
	Het par 2	-7.825	6.149	0.203
Overeducation at the start of the career	Female	-0.061	0.054	0.258
	Siblings	0.014	0.020	0.488
	Foreign origin	0.008	0.134	0.952
	Education Father	0.000	0.010	0.959
	Education Mother	-0.046	0.009	0.000
	Birth day / 100	0.004	0.026	0.887
	Unemployment Delay	0.036	0.012	0.004
	Delay SE	0.016	0.094	0.865
	Track Choice LSE	0.089	0.576	0.877
	Start HSE	-0.569	0.251	0.023
	Track Choice HSE	0.373	0.183	0.041
	HSE	1.580	0.142	0.000
	Start LTE	0.074	0.081	0.361
	Track Choice LTE	0.062	0.117	0.598
	LTE	-1.113	0.083	0.000
	Start THE	0.426	0.446	0.339
	Track Choice HTE	-0.460	0.124	0.000
	HTE	0.362	0.438	0.408
	Cohort 1978	0.304	0.068	0.000
	Cohort 1980	0.442	0.070	0.000
	_cons	-1.336	0.280	0.000
	Het par 1	-0.156	0.061	0.010
	Het par 2	0.135	0.113	0.232

Table C1: Model estimates

	Female	0.029	0.062	0.642
	Siblings	-0.050	0.022	0.022
	Foreign origin	-0.576	0.145	0.000
	Education Father	-0.035	0.011	0.001
	Education Mother	-0.033	0.010	0.002
	Birth day / 100	-0.040	0.030	0.195
	Unemployment Delay	-0.001	0.015	0.967
	Delay SE	-0.104	0.106	0.325
	Track Choice LSE	-0.113	0.391	0.773
	LSE	-0.102	0.255	0.688
	Start HSE	0.141	0.259	0.587
Selection equation wage at 23	Track Choice HSE	-0.200	0.209	0.338
	HSE	0.120	0.145	0.408
	Start LTE	-0.369	0.097	0.000
	Track Choice LTE	-0.157	0.130	0.225
	LTE	-0.178	0.094	0.059
	Start THE	-0.846	0.543	0.119
	Track Choice HTE	-0.169	0.167	0.310
	HTE	-0.189	0.535	0.723
	Overeducation	0.050	0.067	0.452
	Cohort 1978	1.337	0.081	0.000
	Cohort 1980	1.694	0.084	0.000
	_cons	0.713	0.237	0.003
	Het par 1	-0.054	0.068	0.427
	Het par 2	0.083	0.130	0.524

Table C1: Model estimates

Log-wage at 23	Female	-0.075	0.006	0.000
	Siblings	-0.003	0.002	0.127
	Foreign origin	0.024	0.014	0.088
	Education Father	-0.001	0.001	0.405
	Education Mother	0.001	0.001	0.397
	Birth day / 100	0.000	0.003	0.924
	Unemployment Delay	0.001	0.001	0.442
	Delay SE	-0.015	0.009	0.118
	Track Choice LSE	-0.008	0.039	0.827
	LSE	0.016	0.024	0.501
	Start HSE	0.013	0.024	0.574
	Track Choice HSE	0.013	0.020	0.506
	HSE	0.013	0.014	0.338
	Start LTE	0.019	0.009	0.031
	Track Choice LTE	-0.007	0.014	0.588
	LTE	0.068	0.010	0.000
	Start THE	0.052	0.073	0.473
	Track Choice HTE	0.082	0.023	0.000
	HTE	-0.002	0.072	0.978
	Overeducation	0.007	0.022	0.736
	Overeducation*HSE	-0.040	0.023	0.080
	Overeducation*LTE	-0.001	0.015	0.971
	Overeducation*HTE	-0.071	0.027	0.010
	Cohort 1978	0.023	0.008	0.006
	Cohort 1980	0.031	0.008	0.000
	_cons	1.913	0.022	0.000
	Het par 1	-0.007	0.006	0.265
	Het par 2	0.186	0.012	0.000
	sigma	0.187	0.002	0.000

Table C1: Model estimates

Selection equation wage at 26	Female	-0.278	0.147	0.060
	Siblings	0.033	0.050	0.513
	Foreign origin	-1.322	0.283	0.000
	Education Father	0.037	0.027	0.170
	Education Mother	0.018	0.026	0.472
	Birth day / 100	-0.005	0.073	0.945
	Unemployment Delay	-0.052	0.035	0.134
	Delay SE	0.193	0.250	0.440
	Track Choice LSE	0.715	1.009	0.479
	LSE	1.036	0.717	0.148
	Start HSE	-0.199	0.748	0.790
	Track Choice HSE	0.111	0.498	0.824
	HSE	-0.627	0.387	0.105
	Start LTE	-0.297	0.217	0.171
	Track Choice LTE	-0.423	0.282	0.134
	LTE	0.136	0.218	0.532
	Start THE	-1.506	0.691	0.029
	Track Choice HTE	-0.478	0.462	0.301
	HTE	2.020	0.644	0.002
	Overeducation	-0.456	0.157	0.004
	Cohort 1978	-0.439	0.164	0.008
	_cons	-3.626	0.574	0.000
	Het par 1	6.929	0.248	0.000
	Het par 2	6.637	0.317	0.000

Table C1: Model estimates

	Female	-0.070	0.006	0.000
	Siblings	-0.003	0.002	0.116
	Foreign origin	0.036	0.016	0.027
	Education Father	0.001	0.001	0.486
	Education Mother	0.001	0.001	0.399
	Birth day / 100	-0.009	0.003	0.003
	Unemployment Delay	0.001	0.001	0.369
	Delay SE	-0.022	0.010	0.034
	Track Choice LSE	0.034	0.044	0.434
	LSE	0.018	0.027	0.510
	Start HSE	0.025	0.027	0.352
	Track Choice HSE	0.024	0.019	0.199
	HSE	0.023	0.015	0.124
Log-wage	Start LTE	0.017	0.009	0.061
at 26	Track Choice LTE	0.001	0.013	0.939
	LTE	0.085	0.010	0.000
	Start THE	-0.020	0.045	0.659
	Track Choice HTE	0.050	0.016	0.002
	HTE	0.121	0.044	0.006
	Overeducation	0.058	0.027	0.030
	Overeducation*HSE	-0.077	0.028	0.006
	Overeducation*LTE	-0.021	0.015	0.176
	Overeducation*HTE	-0.030	0.022	0.171
	Cohort 1978	0.016	0.007	0.011
	_cons	1.370	0.038	0.000
	Het par 1	0.567	0.031	0.000
	Het par 2	0.961	0.032	0.000
	sigma	0.144	0.002	0.000

Table C1: Model estimates

	Female	-0.021	0.101	0.833
	Siblings	0.049	0.035	0.162
	Foreign origin	-1.006	0.206	0.000
	Education Father	0.022	0.018	0.230
	Education Mother	0.018	0.017	0.287
	Birth day / 100	-0.086	0.051	0.090
	Unemployment Delay	-0.005	0.024	0.822
	Delay SE	-0.112	0.169	0.508
	Track Choice LSE	0.557	0.737	0.450
	LSE	0.362	0.427	0.396
	Start HSE	0.027	0.440	0.951
Selection equation wage at 29	Track Choice HSE	0.108	0.375	0.775
	HSE	0.009	0.243	0.970
	Start LTE	-0.254	0.157	0.106
	Track Choice LTE	0.347	0.229	0.129
	LTE	0.299	0.155	0.053
	Start THE	-0.575	0.686	0.402
	Track Choice HTE	-0.168	0.266	0.528
	HTE	0.911	0.674	0.176
	Overeducation	-0.052	0.109	0.636
	Cohort 1978	1.018	0.106	0.000
	_cons	-4.575	0.378	0.000
	Het par 1	5.600	0.181	0.000
	Het par 2	5.264	0.229	0.000

Table C1: Model estimates

	Female	-0.060	0.006	0.000
	Siblings	-0.003	0.002	0.184
	Foreign origin	0.033	0.014	0.019
	Education Father	0.000	0.001	0.899
	Education Mother	0.000	0.001	0.713
	Birth day / 100	-0.001	0.003	0.660
	Unemployment Delay	0.000	0.001	0.722
	Delay SE	-0.024	0.010	0.019
	Track Choice LSE	0.011	0.041	0.779
	LSE	0.022	0.028	0.424
	Start HSE	-0.053	0.028	0.053
	Track Choice HSE	0.091	0.022	0.000
	HSE	0.052	0.015	0.001
Log-wage	Start LTE	0.013	0.009	0.155
at 29	Track Choice LTE	0.004	0.011	0.751
	LTE	0.076	0.010	0.000
	Start THE	0.116	0.045	0.010
	Track Choice HTE	0.042	0.012	0.001
	HTE	-0.010	0.045	0.827
	Overeducation	-0.028	0.025	0.258
	Overeducation*HSE	-0.009	0.027	0.738
	Overeducation*LTE	-0.001	0.014	0.928
	Overeducation*HTE	-0.047	0.017	0.005
	Cohort 1978	0.030	0.006	0.000
	_cons	1.476	0.031	0.000
	Het par 1	0.559	0.024	0.000
	Het par 2	0.903	0.026	0.000
	sigma	0.147	0.002	0.000
	P(K=1)	0.274		
	P(K=2)	0.064		
	P(K=3)	0.662		

Table C1: Model estimates

Log-likelihood	-29770.49