Ghent University
Faculty of Economics and
Business Administration

Stijn Baert | Transitions in Youth: on Springboards, Waterfalls and Bottlenecks

Submitted to the Faculty of Economics and Business Administration in fulfillment of the requirements for the Degree of Doctor in Economics

Ghent University
Faculty of Economics and
Business Administration

Stijn Baert | Transitions in Youth: on Springboards, Waterfalls and Bottlenecks

Supervisor:
Prof. Dr. Bart Cockx
UGent, UCLouvain, CESifo, IZA

Examination Committee:
Prof. Dr. Marc De Clercq
UGent; Dean
Prof. Dr. Patrick Van Kenhove
UGent; Academic Secretary
Prof. Dr. Bart Cockx
UGent, UCLouvain, CESifo, IZA; Supervisor
Prof. Dr. Michèle Belot
University of Edinburgh
Prof. Dr. Sabien Dobbelaere
VU University Amsterdam, Tinbergen Institute, IZA
Prof. Dr. Jozef Konings
KU Leuven, CEPR, IZA
Prof. Dr. Eddy Omey
UGent
Prof. Dr. Dieter Verhaest
HUBrussel, UGent
Prof. Dr. Elsy Verhofstadt
HoGent, UGent

Contact modalities:
Faculty of Economics and Business Administration
Department of Social Economics
Tweekerkenstraat 2
B-9000 Gent, Belgium
T +32-9-264.34.81
E Stijn.Baert@UGent.be

Acknowledgements (in Dutch)

Dit doctoraat was nooit tot stand gekomen zonder de directe en indirecte steun die ik kreeg, zowel in de professionele als in de persoonlijke sfeer. In dit voorwoord richt ik me naar enkele personen die voor mij het verschil maakten. Een eerste persoon die ik wens te bedanken is mijn promotor, professor Bart Cockx, die vanzelfsprekend de grootste impact had op mijn werk. Toen ik, bij het begin van mijn onderzoek, één van zijn gewezen doctorandi, professor Matteo Picchio, hoorde praten over "my research father", vond ik dat eerder pathetisch klinken. Ondertussen begrijp ik wat hij bedoelde. Bart, enerzijds heb jij me werkelijk opgevoed van een leek tot een onderzoeker. Ik ben oprecht dankbaar dat iemand met dergelijke onderzoekscapaciteiten—ze blijven indruk maken—in mij heeft willen investeren. Meer dan 1100 e-mails heb je de afgelopen jaren naar me gestuurd, telkens met de bedoeling me als onderzoeker op een hoger niveau te tillen. Anderzijds bracht onze samenwerking de ontwikkeling van een warme band met zich mee. Dat laatste stond nochtans niet in de sterren geschreven. De man met principes tegenover de pragmatische jongen. Iemand die aan het begrip grondigheid diepgang gaf tegenover iemand die met zijn ongeduld vaak geen weg wist. Een promovendus van de KU Leuven tegenover een UGent'er pur sang. En toch . . . Mijn professionele leven zou veel minder aardig zijn zonder het samenwerken met jou. Dankjewel, Bart.

Verder ben ik een aantal professoren uit de examenjury bijzondere dank verschuldigd. Vooreerst professor Eddy Omey. Als ik onderwijs heb kunnen verstrekken dat gewaardeerd werd, dan is dat grotendeels aan hem te wijten. Zonder het vertrouwen dat ik vanaf dag 1 kreeg, afgekruid met af en toe wat wijze raad, had ik nooit kunnen doen wat ik deed. De kortetermijndeadlines inzake het onderwijs dat ik aan de Vakgroep Sociale Economie mocht verstrekken, vormden bovendien het ideale complement voor het langetermijnperspectief inherent aan doctoraal onderzoek. Dankjewel, Eddy. Ook in mijn onderzoek en dienstverlening heb je me altijd ten volle gesteund. Als je positief kon antwoorden op een vraag, heb je dat altijd gedaan. Professor De Clercq, Marc, ik heb je in het verleden al vaak geloofd als een goede vriend van vele jaren, maar ook professioneel ben je meermaals een bron van inspiratie geweest. Jij was het die me toonde dat een goed lesgever zijn persoonlijkheid niet

aan de kant schuift. Professor Verhaest, Dieter, ons eerste gezamenlijke onderzoeksproject, thans de laatste studie van dit doctoraat, is voor mij echt een belangrijke katalysator geweest. Het heeft mee de onderzoeksambitie in mij losgemaakt. Daarnaast kon ik ook altijd bij je terecht om in vertrouwen ervaringen te delen. Ik dank je voor de heel fijne samenwerking. Professor Verhofstadt, Elsy, bedankt voor de immer fijne omgang, de motiverende woorden en, op de valreep, de leuke samenwerking met ons veldonderzoek rond discriminatie van gewezen jeugddelinquenten. Professor Dobbelaere, Sabien, bedankt voor het harde werk in mijn begeleidingscommissie en om mee mijn onderzoeksverblijf in Amsterdam te realiseren. Ook dank aan de externe juryleden: professor Belot, Michèle—ik hoop dat we ooit een mooi experimenteel onderzoeksproject opstarten—en professor Konings, Joep—het was fijn een "idool" van vele jaren via mijn verdediging te leren kennen.

Ook buiten deze jury ben ik heel wat mensen, in de eerste plaats collega's van de Faculteit Economie en Bedrijfskunde en bij uitbreiding van de Universiteit Gent, dank verschuldigd voor de ondersteuning en samenwerking. Ik denk daarbij aan professor Freddy Heylen, een bron van inspiratie en positieve energie. Als ik van professor De Clercq heb proberen stelen dat je jezelf moet blijven tijdens een college, ben jij, Freddy, altijd een voorbeeld geweest in het oprecht "graag zien" van de studenten. Ik vond het echt geweldig dat ik met iemand die ik zo hard respecteer en apprecieer aan onderzoek mocht doen in een studie buiten dit doctoraat. Hopelijk meer van dat in de toekomst! Verder dank aan professor Van de gaer, Dirk, voor de aangename samenwerking in het kader van het opleidingsonderdeel "Micro-economie: Beslissingstheorie" én voor de uitstekende raad op de juiste momenten. Professor Gerdie Everaert ben ik dankbaar voor de transpiratie in mijn begeleidingscommissie en doctor Walter Van Trier ben ik erkentelijk voor de inspiratie daarbuiten. Dank ook aan professor Matteo Picchio van wie ik, tijdens de relatief korte periode dat hij aan de UGent verbleef, heel wat bijleerde. Ten slotte vermeld ik ook het secretariaat van de Vakgroep Sociale Economie, mevrouw Erna Haerens in het bijzonder, collega Daan Isebaert en gewezen Masterproefstudenten Niels Gheyle en Cora Vandamme voor de prima samenwerking en ondersteuning. In het algemeen ben ik de Universiteit Gent, mijn geliefde alma mater, dankbaar voor de ingrijpende positieve impact die zij sinds 2001, toen ik startte als student, en meer nog sinds 2004, toen ik me binnen de instelling actief ging engageren als studentenvertegenwoordiger, heeft gehad op mijn ontplooiing. Werken aan de UGent, is voor mij werken met fierheid en passie.

Buiten de Universiteit Gent ben ik in de eerste plaats dank verschuldigd aan professor Bas van der Klaauw die me ontving als gastonderzoeker aan de Vrije Universiteit Amsterdam. Zowel de samenwerking met Bas als, meer algemeen, het verblijven in de dynamische onderzoeksomgeving van de VU zijn zeer leerrijk voor mij geweest. Dank ook aan professor Michael

Rosholm, die me ontving aan de universiteit van Aarhus voor een kort onderzoeksverblijf.

In de persoonlijke sfeer gaat mijn eerste woord van dank uit naar Jens. Omdat hij een klankbord was voor gekke ideeën en kleine angsten. Omdat hij rust bracht als ik in "overdrive" dreigde te gaan. Verder dank aan alle vrienden, kennissen en familieleden die me hielpen, op de gepaste momenten, het hoofd leeg te maken. In het bijzonder denk ik aan mijn getrouwen buiten de humane wetenschappen en, bij uitbreiding, buiten de academisch wereld die ondanks hun eigen professionele invalshoek toch altijd het respect en de interesse toonden voor waar ik mee bezig was.

Ten slotte dank aan de persoon die verantwoordelijk is voor 100% van mijn gaven en 0% van mijn gebreken. Voor 100% van mijn goede opvoeding en 0% van mijn ontsporingen. De persoon die mij leerde wat pure liefde is, wat goed zijn is. De persoon aan wie mijn doctoraat is opgedragen. Dankjewel, mama.

Dutch summary

De vier studies die gepresenteerd worden in dit doctoraal proefschrift, zijn allen gericht op de identificatie van succesfactoren in de transities van jongeren op school en op de arbeidsmarkt. Voor elk van deze studies werd gebruik gemaakt van Vlaamse data. Enerzijds werd de SONAR-databank, longitudinale data als resultaat van een representatieve bevraging van drie geboortecohorten, verder ontgonnen. Anderzijds werd data verzameld door het opzetten van een veldexperiment in de Vlaamse arbeidsmarkt.

In een eerste studie delen we de waargenomen verschillen tussen autochtone en allochtone jongeren in school- en arbeidsmarktuitkomsten op in (i) een stuk dat verklaard kan worden door sociaal-economische kenmerken (zoals het onderwijsniveau van de ouders) en (ii) een resterend "zuiver etnisch verschil" (veroorzaakt door bijvoorbeeld discriminatie en etnische verschillen in voorkeuren en verwachtingen). We bouwen daartoe een dynamisch discretekeuzemodel waarin we de opeenvolging van schooluitkomsten en eerste arbeidsmarktuitkomsten verklaren. Onze bijdrage tot de literatuur ligt erin te focussen op (i) de vertraging waarmee schooluitkomsten gerealiseerd worden, (ii) de identificatie van de specifieke momenten waarop zuiver etnische verschillen opduiken, (iii) het afzonderlijke belang van het slagen voor een schooljaar en de beslissing om verder te studeren nadien en (iv) het belang van de taal die in het ouderlijke huis wordt gesproken. We vinden enerzijds, en in lijn met de literatuur, dat het zuiver etnische verschil in schooluitkomsten klein is wanneer geen rekening wordt gehouden met de schoolvertraging waarmee deze gerealiseerd worden. Anderzijds is het zuiver etnische verschil substantieel eens op schoolvertraging en op eerste arbeidsmarktuitkomsten wordt gefocust. Wat de rol van het gebruik van Nederlands in het ouderlijke huis betreft, vinden we enkel een significante rol in de overgang van school naar werk voor laaggeschoolden. In een tweede studie gaan we dieper in op één van de achterliggende mechanismen van het zuiver etnische verschil in eerste arbeidsmarktuitkomsten: discriminatie. We bekijken empirisch of aanwervingsdiscriminatie de overgang van school naar werk voor Turkse jongeren in Vlaanderen beïnvloedt. We dragen in deze tweede studie bij tot de internationale economische literatuur door, als eersten, de theoretische relatie tussen aanwervingsdiscriminatie en

arbeidsmarktkrapte te testen. Daartoe voeren we een correspondentietest uit op de Vlaamse arbeidsmarkt: fictieve sollicitaties, afwisselend gekoppeld aan een Vlaamse en Turkse naam, worden verzonden naar bestaande vacatures voor schoolverlaters. De resultaten bevestigen de theoretische verwachtingen. Enerzijds vinden we dat, in vergelijking met hun autochtone tegenhangers, Turkse kandidaten even vaak uitgenodigd worden voor een jobgesprek wanneer zij solliciteren voor jobs waarvoor de arbeidsmarktkrapte hoog is. Anderzijds dienen deze Turkse kandidaten dubbel zoveel sollicitaties uit te voeren om even vaak uitgenodigd te worden voor een jobgesprek als autochtone kandidaten wanneer zij solliciteren voor banen waarvoor de vacatures makkelijk in te vullen zijn.

Een derde studie is gericht op de korte- en langetermijneffecten van overzitten op verdere schooluitkomsten. Via een dynamisch discretekeuzemodel verklaren we de opeenvolgende schoolresultaten en schoolbeslissingen van jongeren tijdens het secundair onderwijs. We houden daarbij rekening met de typische "watervalkenmerken" van het Vlaams onderwijs door ook de transities tussen studierichtingen te modelleren. Bovendien wordt statistisch gecontroleerd voor kenmerken die niet waarneembaar zijn vanuit het standpunt van de onderzoeker. In contrast met eerdere bijdragen vinden we dat overzitten een positief effect heeft op de slaagkansen voor het volgende schooljaar. Op langere termijn is het effect heterogeen: terwijl meer getalenteerde studenten een eerder negatief langetermijneffect ondervinden van overzitten, kunnen minder getalenteerde studenten een blijvend voordeel hebben.

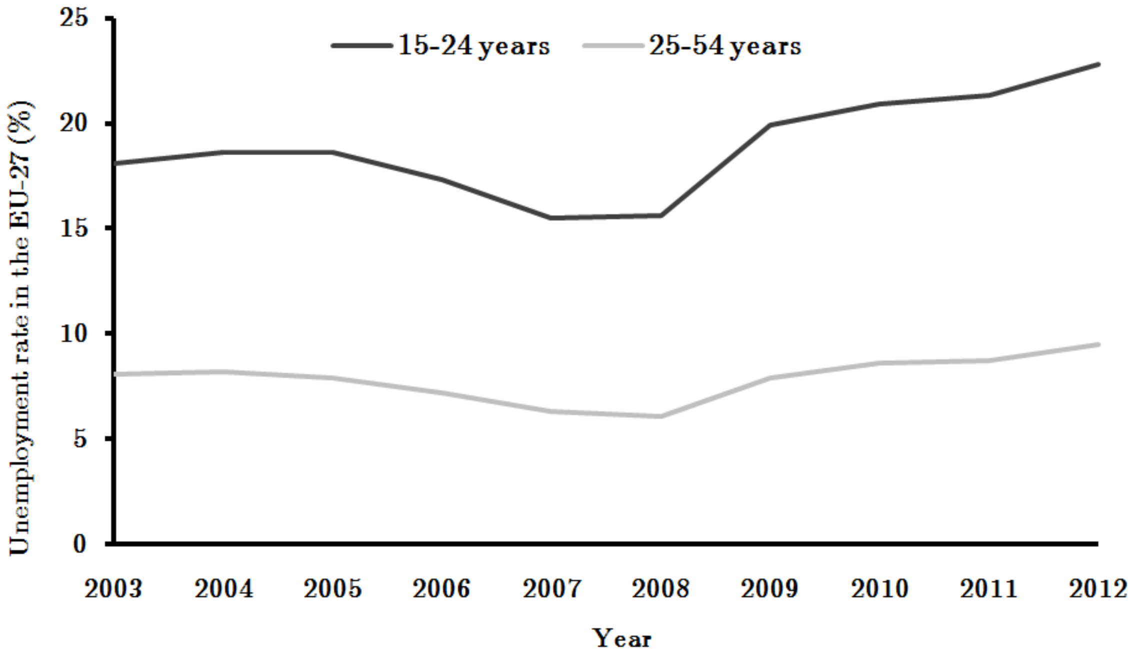
In een vierde studie, ten slotte, bestuderen we of werkloze schoolverlaters hun overgang naar een adequate job kunnen versnellen door een baan te aanvaarden onder hun scholingsniveau. Om deze onderzoeksvraag te beantwoorden, passen we de "Timing of Events"-methode toe. De timing van het instromen in een job waarin men overschoold is, wordt daarbij gebruikt om het oorzakelijke effect te identificeren van het aanvaarden van deze job op de duurtijd tot een adequate job. We vinden dat overscholing veeleer een val is waarin men blijft vastzitten dan een springplank naar een adequate job. In concreto wijzen onze onderzoeksresultaten uit dat jongeren die een job aanvaarden waarin zij overschoold zijn (in plaats van enkel adequate jobs te aanvaarden) de snelheid van hun overgang naar adequate arbeid verlaagd zien met 51 –98%. Hoe vroeger overscholing wordt aanvaard, hoe negatiever het effect.

Table of Contents

1
General Introduction

The future will tell which name will be given to the deep economic crisis that coincided the years of this doctoral research. Candidates indicating its severity are "the Great Recession" and "the Long Recession". This crisis developed from a liquidity crisis in the financial markets into a global economic and sovereign debt crisis. The last years it also turned into a labour market crisis. Figure 1.1 describes the evolution of the unemployment rate in the EU-27 according to the ILO definition. While at the start of the crisis in 2008 the prime age adult (aged between 25 and 54 years) unemployment rate in the European Union attained just 6.1%, it has been rising from then onwards reaching 9.5% in 2012. This evolution is even more alarming among youth (under 25 years old): the youth unemployment rate grew from 15.6% in 2008 to 22.8% in 2012. Clearly, young people bear much of the brunt of the current economic crisis. Their higher incidence of unemployment at the start of the career is particularly worrisome, since it can induce long-lasting scars on the subsequent career development, a mechanism on which we elaborate further during the following chapters. The risks posed by a "scarred" generation have motivated many governments to take action, notably by scaling up funds for youth labour market programmes. Recently the European Commission launched the Youth Guarantee, a for the period 2014–2020 6 million euro worth action to help EU countries get young people into employment, further education or (re)training within four months of leaving school.

Figure 1.1: Unemployment rate in the EU-27 (2003-2012)



Source: Eurostat.

In order to develop adequate policy actions to fight youth unemployment in Europe and elsewhere, there is need for identifying success factors in first labour market transitions. In addition, since labour market success is closely related to school achievement,[1] success in the former cannot be independent from success in the latter. In the following four chapters of this thesis we present studies on three factors that influence success in school and in the transition from school to work: (i) ethnicity, (ii) school retention and (iii) overeducation[2] at the start of the career.

Chapter 2 is motivated by the fact that in Belgium, as in many other countries, school and labour market transitions are much more successful for native youth than for ethnic minority youth. The question is whether policy action targeted at the latter youth is the right response. It is if the observed gaps are induced by pure ethnic differences in preferences and expectations or by discrimination. However, if these gaps just mirror different family endowments that result in different levels of educational attainment and therefore in different labour market performances, then no specific measures for minority youth are required to eliminate these gaps. Therefore, in Chapter 2 the observed gaps in educational attainment and first labour

[1] While the youth unemployment among individuals without a secondary education degree was 30.3%, it was "only" 20.0% (17.9%) among the ones with only a secondary (tertiary) education degree (source: Eurostat).

[2] A worker is considered to be overeducated if her/his education level is higher than the level that is typically required to perform adequately (McGuinness, 2006).

market outcomes between native and immigrant youth in Flanders[3] are decomposed into (i) differences in observed family endowments and (ii) a residual "pure ethnic gap". This study innovates by explicitly taking delays in educational attainment into account, by identifying the moments at which the pure ethnic gaps arise, by disentangling the decision to continue schooling at the end of a school year from the achievement within a particular grade and by integrating the language spoken at home among the observed family endowments. In line with the literature, the pure ethnic gap in educational attainment is found to be small if educational delays are neglected. However, this pure ethnic gap is substantial if these delays are taken into account and for school-to-work transitions.

In order to test whether the pure ethnic gap in the transition from school to work, as outlined in Chapter 2, reflects discrimination in the hiring process, in Chapter 3 the results of a field experiment on unequal treatment based on ethnic origin are reported. This chapter contributes to the international economic literature in being the first to test the theoretical relationship between hiring discrimination and labour market tightness in an empirical way. To this end we sent out fictitious job applications of school-leavers, randomly assigned to individuals with a native and a Turkish sounding name, to vacancies for jobs requiring no work experience in Flanders. Classifying these jobs on two measures of labour market tightness, we verify to what extent our measure of discrimination, the differential callback rate, differs between types of jobs. In line with theoretical expectations, we find that, compared to natives, candidates with a Turkish sounding name are equally often invited to a job interview if they apply for occupations for which vacancies are difficult to fill, but they have to send twice as many applications for occupations for which labour market tightness is low. Our findings are robust against various sensitivity checks.

Grade retention is used in many countries as a tool to improve poor academic performances. The hypothesis is that, by resitting the same grade, low-achieving students have extra time to catch up to the grade-level requirements, in terms both of knowledge and emotional maturity. Moreover, the threat of retention might be an incentive device to work harder. However, retention might generate personal and academic costs with both short- and long-term effects. Most former empirical contributions on the effects of grade retention conclude that grade retention has a negative effect on subsequent performances. In Chapter 4, we add to the literature on short- and long-term effects of grade retention controlling for school track mobility. We model schooling attainment and track choices by means of a dynamic qualitative choice model which flexibly takes into account the presence of unobserved characteristics jointly determining the educational choices and performances. In this model, we allow the effect of past retention

[3]Flanders is the Northern and Dutch speaking region of Belgium.

episodes to vary across different levels of the unobserved determinants. By doing so we find that grade retention has a positive impact on the next evaluation and can permanently affect subsequent educational achievements. The direction of the permanent effect is heterogeneous: while more able students are permanently penalised by retention, less able students benefit from it.

Numerous studies have shown that many young workers are overeducated at the start of their career. Given that overeducated workers have lower earnings and job satisfaction, one might wonder why young job seekers actually accept jobs with requirements below their educational attainment. One potential answer is given by the career mobility theory. This theory states that overeducation is an investment in work experience which enhances promotion opportunities to higher level positions inside or outside the firm. In Chapter 5 we test this hypothesis. More concretely, we are the first to investigate whether graduates who accept a job below their level of education accelerate or delay the transition into a first job that matches their level of education. Contrary to many other contributions on the long-term effects of accepting an overeducated job at the start of the career, we handle selection (into overeducated and adequate employment) on both observables and unobservables. For this, we apply the Timing of Events approach. The research results show that overeducation is a trap. By accepting a job for which one is overeducated rather than only accepting adequate job matches, monthly transition rates into adequate employment fall by 51–98%, depending on the elapsed unemployment duration.

All studies presented in this PhD thesis are based on Flemish data. For Chapter 2, Chapter 4 and Chapter 5 we further explore the SONAR data. These data are based on a representative longitudinal survey conducted in Flanders on 9,000 individuals of the 1976, 1978 and 1980 cohorts and aimed at studying the transition from school to work. The SONAR data contain detailed information regarding both school and labour market careers, which makes them very suitable for investigating success factors in transitions in youth. Also the experimental data on which the analyses reported in Chapter 3 are based we gathered in Flanders. This means that our research results and derived policy recommendations are particularly relevant for the Flemish region. However, also policy makers outside Flanders may take an interest in our results. First, the causal mechanisms we identify are also relevant for other countries. There is no reason to believe that, for instance, the relation between labour market discrimination and labour market tightness we identify in Chapter 3 or the trap effect of accepting a job in which one is overeducated we identify in Chapter 5, would be specific to the Flemish context. Moreover, the results on pure ethnic gaps in school and in the labour market in Chapter 2 complement those of similar studies in the US, France and Denmark. Furthermore, although in Chapter 4 we integrate the particular features of the Flemish schooling system in our econometric model, we believe that the results based on this model are relevant for all countries adopting grade retention.[4] Second, in Chapter 3 and Chapter 5, we are the first to empirically test two economic theories that are widespread both among academics and policy makers. Third, the methodological workhorses we employ in the studies of this thesis can be replicated by (policy) researchers inside and outside Flanders to answer the policy questions they are interested in.

The four chapters in this thesis are completely self-contained[5] and can thereby be read in any order. However, although these studies focus on completely different key determinants of transitions in youth and follow different methodologies, there is bilateral cohesion between particular chapters in different respects both in terms of content and in terms of methodology. As regards content, two links are worth highlighting here. First, and as mentioned before, the study presented in Chapter 3 is motivated by the results of the one presented in Chapter 2. Second, the research results of both Chapter 2 and Chapter 4 stress the importance of grade retention in school. In Chapter 2 the number of years of grade retention at a particular level of education is an important *outcome* in which ethnic groups, even after controlling for family endowments, differ. In the latter chapter grade retention is investigated as a key (short- and long-term) *determinant* of later education outcomes.

From a methodological point of view, all studies in this thesis focus on (and contribute to the literature by) solving selection problems. Based on survey or administrative data, individuals who are identical in observable characteristics may differ in characteristics that are unobservable to the researcher (for instance motivation and preference for leisure) but affect the outcomes of interest. In Chapter 3 we control all the employers' decision making information by our experimental design so that selection on unobservable characteristics is not an issue. In Chapter 2, Chapter 4 and Chapter 5 we use non-experimental statistical methods to account for "selection on unobservables". In particular, we solve for a dynamic selection problem in education in Chapter 2 and Chapter 4. This problem is brought about by the progressively growing negative correlation between observed endowments (such as parental educational attainment) and unobserved endowments because pupils with adverse observed endowments pass the final evaluation at the end of a particular grade and continue schooling

[4] Grade retention is possible in most European countries. Data from the 2009 PISA survey show that the lowest retention rates (less than 3%) are found in Slovenia, the United Kingdom, Iceland and Finland. The highest rates (more than 30%) are found in Spain, France, Luxembourg and Portugal (OECD, 2011).

[5] Since according to the current international quality standards a PhD aims at demonstrating that the candidate can conduct research at a level of that is published in international peer reviewed journals, a PhD thesis nowadays consists more and more of a collection of self-contained articles. At the moment this PhD thesis went to print, Chapter 2 and Chapter 5 were accepted for publication in, respectively, the Economics of Education Review and Labour Economics.

only if their unobserved endowments are sufficiently favourable. This biases the coefficients of the observed endowments downwards, and more so as one proceeds to higher grades. To that end we model subsequent schooling outcomes from the start of primary school (Chapter 2) or from the start of secondary school (Chapter 4), explicitly accounting for the initial conditions problem.

Besides the highlighted links between the chapters of this PhD thesis, the careful reader might spot that while in Chapter 4 we argue that the modelling of educational tracks in (secondary) education (and therefore of horizontal transitions within a school year) is important when investigating educational outcomes, at the same time this dimension is completely absent in the model in Chapter 2. On the one hand, this has to do with research focus. While in Chapter 4 the focus lies directly on (vertical and horizontal) transition dynamics in secondary education, in Chapter 2 we focus on ethnic gaps in particular school and labour market outcomes. On the other hand, we were forced to make abstraction of the horizontal dimension to keep the programming and the estimation of the model in Chapter 2 feasible. Two reflections are relevant in this context. First, as mentioned before, in Chapter 2 we find only little evidence for pure ethnic gaps in educational outcomes without specifying the potential delay with which these are attained, but do find important pure ethnic gaps if we take delays into account. When discussing this result we suggest that similar conclusions may arise with respect to other measures of educational achievement within a particular level of educational attainment. This might be the case for the educational track in which pupils realise their educational attainment. In other words: there might be a pure ethnic gap (caused by, for instance, ethnic differences in parental expectations) in the level of this track. Results by Colding et al. (2009) indicate, however, that, at least in Denmark, observed ethnic gaps in the choice for prestigious tracks in secondary education can be to a large extent explained by family endowments. Second, the reader might worry that the results presented in Chapter 2 are biased by making abstraction of the horizontal dimension. Indeed, if due to pure ethnic differences immigrant youth are overrepresented in particular tracks and these tracks lead, ceteris paribus, to better (or worse) schooling or labour market outcomes, the pure ethnic gaps presented in Chapter 2 could be biased downwards (upwards) by not controlling for this horizontal dimension. In this respect Chapter 4 shows that school drop-out is, ceteris paribus, higher in the least prestigious track of Flemish secondary education, the vocational track. If, therefore, immigrant youth are, after controlling for family endowments, overrepresented in this track, the pure ethnic gaps in educational outcomes *without specifying the delay* with which these are attained would be even smaller after controlling for educational tracks. For the educational outcomes that *take schooling delay into account* the direction of the bias is, however, less clear-cut. The likelihood of grade retention is, ceteris paribus, the lowest in both the most and the least prestigious tracks of secondary education. Therefore, it is unclear in which direction the pure ethnic gap in grade retention would be affected if we were able to control for educational tracks in Chapter 2.

After presenting the announced studies in Chapter 2, Chapter 3, Chapter 4 and Chapter 5, in a General Conclusion we highlight three key results to take away for policy makers and three key directions for future research.

2

Pure Ethnic Gaps in Educational Attainment and School to Work Transitions. When Do They Arise?

This chapter is joint work with Prof. Dr. Bart Cockx (Ghent University, Université catholique de Louvain, CESifo and IZA).

2.1 Introduction

In Europe school-to-work transitions are much more successful for native youth than for ethnic minority youth. In 2011, the youth unemployment rate of non-EU-15 residents in the EU-15 was as high as 29% compared to 20% for natives.[1] In Belgium, the country of analysis, these figures attained 32% and 18%, resulting in a gap of fourteen percentage points, which is reported to be one of the largest in the OECD (OECD, 2008; Nonneman, 2012). This gap is particularly worrisome, since the higher incidence of unemployment at the start of the career can induce long-lasting scars on the subsequent career development (Arulampalam, 2001; Gregg and Tominey, 2005; Mroz and Savage, 2006). Therefore, not surprisingly, the OECD

[1]Source: Eurostat (Labor Force Study: Unemployment rates by sex, age groups and nationality). Youth is defined as individuals between 15 and 24 years old.

(2010) calls ethnic minority youth a target group for intensive assistance. The question is whether this is the right response. It is if the observed unemployment gaps are induced by pure ethnic differences in behaviour or by discrimination. However, if these gaps just mirror different family endowments that result in different levels of educational attainment and therefore in different labour market performances, then no specific measures for minority youth are required to eliminate this gap. Heckman (2011), for instance, argues that in contemporary American society the racial gap in achievement is primarily due to gaps in skills and that, consequently, by closing the gaps in skills, the racial gap disappears. According to this view discrimination in the educational system and in the labour market are not an issue and policies need not be targeted to ethnic minorities but rather to providing support to disadvantaged families of all racial and ethnic backgrounds as early as possible as to enhance the skills of their children.

School attainment and early labour market outcomes of immigrant youth have been studied amply in the literature. Researchers have mostly focused on a single or a couple of educational or labour market transitions in isolation from related transitions, such as the decision to enrol in tertiary education (see, e.g., Hagy and Staniec, 2002), the probability of succeeding the first year at university (see, e.g., Ortiz and Dehon, 2008) or the probability of a successful transition to work (see, e.g., Eckstein and Wolpin, 1998; Ryan, 2001; Pozzoli, 2009). A problem with this literature is that analyses that ignore the dynamic sorting that takes place in the educational progression are biased. Cameron and Heckman (1998) show this formally. Intuitively, this bias is brought about by the progressively growing negative correlation between observed endowments, such as the parental educational attainment, and unobserved endowments because pupils with adverse observed endowments pass the final evaluation at the end of a particular grade and continue schooling only if their unobserved endowments are sufficiently favourable. This biases the coefficients of the observed endowments downwards and more so as one proceeds to higher grades.

Cameron and Heckman (2001) explicitly address this selectivity problem by modelling, beyond the maximum compulsory school age, the decision to drop out in each school year as a dynamic discrete choice model that explicitly takes into account unobserved determinants of this decision that can generate the aforementioned sorting. Based on this model they investigate the sources of racial and ethnic disparity in college attendance. They find that the racial gap in educational attainment is eliminated or even reversed once they adjust for differences in parental background and family environment.

These conclusions are not only relevant for the US. For instance, based on a version of the model of Cameron and Heckman (2001) that disregards the age dimension, Belzil and Poinas

(2010) report that the gap in higher educational attainment between second generation immigrants and natives in France is mainly explained by family background. In addition, these authors study the gap in the school to work transition and find that the gap in access to permanent employment is nearly completely closed once both family background and educational attainment are conditioned upon. Colding (2006) and Colding et al. (2009) also disregard the age dimension but extend the model of Cameron and Heckman (2001) by taking into account that students need not only to decide whether they continue education beyond the current grade level, but also, if they proceed, in which branch (academic or vocational). They estimate this model on Danish data. Their results corroborate previous findings that family background is an important determinant of educational outcomes, but also demonstrate that differences in endowments alone do not explain the observed gap in educational attainment between natives and ethnic minorities in Denmark.

In this study we follow this line of research to study to what extent the ethnic gap in educational attainment and in school-to-work transitions in Belgium can be explained by observed family endowments or whether a residual pure ethnic gap, reflecting differences in behaviour and unobserved endowments, or discrimination, remains present. Our analysis is based on a retrospective survey taken at age 23 of a representative sample of three cohorts born in 1976, 1978 and 1980, living in Flanders, the Northern Dutch speaking region of Belgium. It contrasts natives to grandchildren of women of "non-Western" nationality, born in Belgium or immigrated prior to age three. The latter selection avoids that the pure ethnic gap partly captures the effects of additional barriers that recent immigrants face (see, e.g., Colding et al., 2009).

We contribute to the literature in a number of ways. First, in the past researchers have studied ethnic gaps in the attainment of particular levels of education, such as secondary school completion or college entry, irrespectively of the age at which these levels are attained. However, since, depending on the educational system, youths can be retained at various points in the educational career, youths may attain these levels at different ages. This matters. Even if retention may improve educational achievement (see Chapter 4), it is costly if it eventually induces pupils to enter the labour market with delay and if employers use it as a negative signal of productivity in their hiring decision. In this study we therefore explicitly take these delays into account both when measuring educational attainment, and by explicitly modelling them as outcomes and determinants of schooling progression. We show that conclusions crucially depend on whether or not delays are considered in the measure of educational achievement.[2]

[2]Cameron and Heckman (2001) also implicitly model schooling delay since they allow schooling choices at particular grades to depend on age. However, they only take these delays into account for one particular outcome: the probability of being in grade nine or higher at age 15. For all other outcomes they consider the

Second, we propose a method to decompose the residual pure ethnic gap according to the moment at which this gap is generated. We do this by studying to what extent this gap diminishes by sequentially conditioning on prior levels of educational attainment. As such, critical grades of study can be identified in which the pure ethnic gap arises more prominently than in other grades. This can be a useful tool to get a better understanding where this gap originates from. In this study the data allow us to identify whether the gap originates in primary school or in any grade beyond age twelve. This is much earlier than in the literature so far that usually starts analysing these gaps from around grade nine in secondary school.

Third, we move a step forward by disentangling the educational outcomes within a particular grade. Sociologists Boudon (1974) and Erikson et al. (2010) argue that observed social class schooling differentials result both from the "primary effects" of differing levels of academic performance, that is passing or failing, and from the "secondary effects" in the educational choices, that is continuing school or dropping out, that one makes at given levels of performance. The pure ethnic gaps may emerge within these steps of the educational progression with a different magnitude or even in the opposite direction, which may call for different policy actions. We distinguish in our empirical analysis between the educational achievement (passing or failing) realised at the end of each grade and the decision to continue schooling (rather than stopping) at the end of each school year and we allow the outcomes of each of these components to depend on past decisions and achievements.

A final innovation is that we integrate the language spoken at home among the observed family endowments. Language is reported to be an important determinant of school and labour market success. van Ours and Veenman (2003) conclude that language proficiency of migrants in the Netherlands has a positive effect on the educational attainment of their sons but no effect on the educational attainment of their daughters. Dustmann et al. (2010) indicate language as the key factor for minority youth in the UK to catch up with white pupils throughout compulsory schooling. Moreover, Dustmann and Fabbri (2003), Chiswick (2008) and Aldashev et al. (2009) conclude that migrants who speak the language of their destination country have better labour market prospects. Therefore, it may matter to control for language usage in an analysis of the determinants of ethnic gaps in schooling and labour market outcomes.

This chapter is structured in the following way. Section 2.2 summarises the institutional setting: the educational system and the youth labour market in Belgium. The next section describes the dataset and provides descriptive statistics that motivate our analysis. Section

schooling outcomes at age 24, an age at which most schooling must be completed, irrespectively of schooling delays.

2.4 presents the econometric model and the simulation and decomposition methodology. The empirical findings are reported subsequently, starting with an assessment of the model in terms of within-sample fit and followed by a series of counterfactual simulations that aim at answering the main research questions. A final section concludes.

2.2 The Institutional Setting: Education and Youth Labour Market

In Belgium the language communities (Flemish and French) are in charge of the organisation of the educational system, while labour market regulation is in the period of analysis mostly organised at the national level. Since the data we analyse concern only inhabitants of Flanders, we restrict the description to the Flemish educational system. School choice is free at all levels and schools are mixed in that children cannot be refused on grounds of gender or ethnicity. Education is compulsory from the first of September of the year in which a child reaches age six and lasts until his/her eighteenth anniversary or the 30th of June of the year in which (s)he reaches age eighteen. Even though a regular student graduates from (the sixth[3] grade of) secondary school at age eighteen, this is not the case for an important share (40%), since students who do not attain a certain competency level are retained and thus required to repeat the school year. This retention may already take place in primary school. Talented pupils can skip grades in nursery and primary school. In our dataset 107 (89) of the 7,256 native children start primary (secondary) school at age five (eleven) instead of six (twelve). None of the immigrant children skip a grade. Special (nursery, primary or secondary) education is aimed at children who need special help, temporarily or permanently. This may be due to physical or mental disability, serious behavioural or emotional problems, or serious learning difficulties. In our research project, these pupils (1% of the total number) are dropped from the sample.

Children can enter nursery school when they are two and a half to three years old. Although nursery education is not compulsory, in Flanders 98% of the kids attend it. A child usually starts primary education at age six, but if the child is not school ready entry can be delayed. Primary education comprises six consecutive years of study. When graduating from primary school, students enter secondary education. Without grade retention (or grade skipping) at primary school pupils enter secondary education in the year in which they reach age twelve. At this point pupils choose between four tracks: general, technical, arts or vocational. A pupil is granted the diploma of general, technical or arts secondary education after successfully

[3]This corresponds to twelfth grade in the US. In the sequel of this chapter we reset, in accordance with the Flemish system, the counter of grades to zero at the start of secondary school.

completing six years ("grades"). Without grade retention (or grade skipping), this occurs in the last compulsory schooling year, at age eighteen. Students in the vocational track are granted a secondary school diploma only after completing a seventh grade, but, since this seventh grade involves quite some specialisation, we assimilate it in this study as part of higher (tertiary) education. Students with a secondary school diploma can enrol directly, without any entry exam,[4] into higher (tertiary) education, that is college or university. Our observation period of education registrations is prior to adoption of the Bologna process. Three sorts of higher education degrees could be obtained: (i) non-university of the "short type" (typically vocationally oriented and lasting three years), (ii) non-university of the "long type" (typically four years mixing a vocational and a more academic curriculum) and (iii) academic university education (typically four or five years). No tuition fee has to be paid at nursery, primary and secondary school and very low and stable tuition fees (from 80 euro to 600 euro in 2012, depending on the parents' income) in higher education. Twenty-two colleges and seven universities are spread over less than 14,000 km^2 resulting in a high regional diffusion of providers of tertiary education. For more details on the educational system, see Chapter 4 and De Ro (2008).

There is no compulsory military service in Flanders and school-leavers enter the labour market directly after school leaving. Moreover (and different from other countries and regions) school-leavers can claim unemployment benefits after a "waiting period" of nine months. This period starts with the registration at the employment office after leaving school. Labour regulation distinguishes between two types of labour contracts: with time stipulation (temporary employment contracts) and without time stipulation (permanent employment contracts). A finite number of successive temporary employment contracts between the same employee and employer, are permitted for a maximum of three years.

2.3 Data and Some Facts

2.3.1 The Data: Retrospective Survey of a Representative Sample of Three Birth Cohorts

The data source (the so called "SONAR" data) consists of representative samples of 3,000 individuals each of three cohorts born in 1976, 1978 and 1980 and living in Flanders when they were 23 years old, the moment of interview. Follow-up interviews were conducted at age 26 and/or 29. Data of these follow-up interviews were, however, not used in the main analysis to avoid drop-out selectivity. They are only used in a sensitivity analysis discussed below.

[4]The only exception is the entry exam for students who want to study medicines.

This database contains exceptionally rich information on both the educational career and the start of the labour market career. It contains, apart from a range of socio-economic variables, monthly information on the educational choices and progression as well as on the labour market status[5] from the moment secondary school is entered,[6] until the moment of the last interview. In addition, the age at which primary school is started, is reported. This information was collected by trained interviewers conducting oral interviews at the interviewees' home address.

2.3.2 Motivating Gaps

Throughout this chapter, two sub-populations of the SONAR cohorts are indexed by the nationality of their grandmother on mother's side. On the one hand, we identify "natives", that are youths whose grandmother on mother's side possesses the Belgian nationality (8,091 individuals). On the other hand, we consider "immigrants", that are youth whose grandmother on mother's side neither has the Belgian nationality nor any other Western[7] nationality (545 individuals). This group is heterogeneous since the nationality of the grandmother does not determine the moment at which the immigration occurred. Card (2005) and Chiswick and DebBurman (2004) find that the educational attainments of the immigrants of the second generation are better than the ones of the first generation. We therefore make the group more homogeneous by only retaining grandchildren who resided and went to school in Flanders from the start of nursery school onwards. This means that we essentially exclude first generation immigrants from the sample.[8] In a robustness check we further enhance the homogeneity of the sample by restricting our sample to respondents with the Belgian nationality (at age 23). Consequently, if the findings of the aforementioned authors apply for Belgium, the gaps in the educational attainment found in this research are a lower bound for the gap of first generation immigrants.

Dropping individuals with (i) missing explanatory variables; (ii) inconsistent school registrations and (iii) years of special education (see Section 2.2) we obtain a sample of 7,256 native respondents and 359 immigrant respondents. Among the immigrants those with a Turkish (122 individuals) and Moroccan (87 individuals) origin are highly represented. 316 of

[5] An individual is employed when holding a job of at least one hour a week and during at least one month. Part-time jobs held by students in the vocational track are not considered as employment, but as part of the educational career.

[6] As indicated before, in principle, secondary school is started in the year of one's twelfth anniversary. In case of grade skipping or retention, this can be at an earlier or later age.

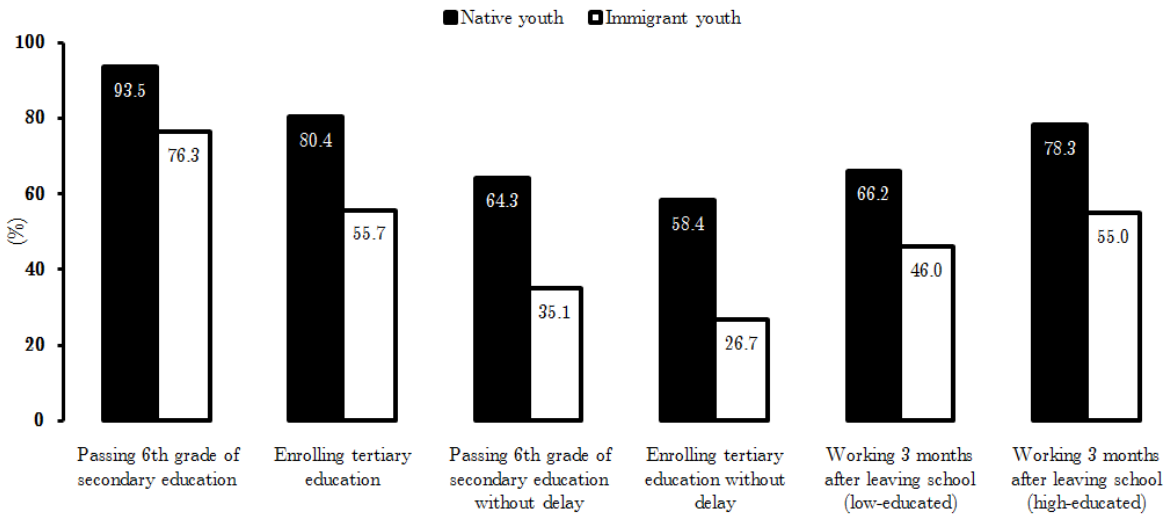
[7] In particular, by "Western" nationality we refer to a North American, British, Scandinavian, Western European or Australian nationality.

[8] First-generation immigrants are only included to the extent that they immigrated to Belgium between birth and the start of nursery school at age three. We chose to not exclude them completely as to avoid a too small immigrant sample, but, as mentioned in the main text, performed a sensitivity analysis that makes this group more homogeneous by only retaining immigrants with the Belgian nationality.

these 359 immigrant respondents have the Belgian nationality at age 23. In the benchmark analysis all 359 immigrants are considered as one group. However, in sensitivity analyses, we restrict once the immigrant sample to those of Turkish and Moroccan origin and once to those of Belgian nationality at age 23. In what follows, we refer to "natives" and "immigrants"[9] according to the definitions in this section.

Figure 2.1 presents some relevant observed gaps in school attainment and successful transitions to work between the native and immigrant groups in our dataset. First, we present the gaps for two key schooling outcomes: (i) passing sixth grade of (and thereby graduating from) secondary education and (ii) enrolling in tertiary education. Concerning these outcomes, we distinguish between realising them (without specifying any potential delay) and realising them without schooling delay. Those in which schooling delay is left unspecified are usually considered in the literature. However, as argued in Section 2.1, it makes sense to also consider educational outcomes specifying the delay with which they are attained, since eventual schooling delays are costly. They translate in postponed labour market entry and therefore in substantial foregone earnings. Moreover, these schooling delays are commonly experienced in the Flemish educational system: 40% of the pupils graduate from secondary school with delay.

Figure 2.1: Schooling and School-to-Work Transition Outcomes by Ethnic Group



Source: own calculations based on the SONAR database. Low-educated is defined as holding a secondary education degree or lower. High-educated is defined as leaving school with one to four successful years of tertiary education and at most one year of schooling delay.

Second, we report, conditional on observed school attainment, the gap for being employed

[9]The label "immigrant" is used somewhat loosely, since it essentially comprises immigrants beyond the first generation besides a minority of first-generation immigrants (see previous footnote).

three months after leaving school as an indicator of successful school-to-work transition. We report this gap for low educated (defined as holding a secondary educational degree or lower) and high educated (defined as leaving school with one to four successful years of tertiary education and at most one year of schooling delay).[10] We chose to condition this indicator on school attainment, since the observed gap in school-to-work transitions unconditional on school attainment is biased downwards, because some youth is still in education at the time that the survey is conducted at age 23: this is more likely to be the case for natives and, since this group is more likely highly educated, its employment propensity is higher.

The first two statistics in Figure 2.1 show that the observed ethnic gaps in school attainment are substantial, both in absolute and in relative terms. Native youth is 17 percentage points more likely than immigrant youth to graduate from secondary education, while they are 25 percentage points more likely to enrol into higher education. Proportionally, these gaps amount to 23% and 44%. These differences are even more outspoken if we consider the fractions of natives and immigrants who attain these educational levels without any schooling delay: 29 and 32 percentage points in absolute terms, or 83% and 119% in relative terms. Finally, the last two statistics illustrate that the observed gaps in the school-to-work transitions are also important, even if we condition on attained educational level. Observe that these gaps do not differ much between the low and the high educated: in absolute terms the difference is 20 to 23 percentage points while in relative terms this varies between 44% and 42%.

2.3.3 Explanatory Variables

In this subsection, we describe the explanatory variables used for each modelled outcome. The choice of covariates is restricted by their availability, their required strict exogeneity, and by their relevance according to the existing research. Cameron and Heckman (2001) find that long-run factors associated with parental background and family environment are strong predictors of the educational disparity between natives and ethnic minorities in that once they control for these long-term factors the gap in educational attainment is completely eliminated or even reversed. This is confirmed in the research of Belzil and Poinas (2010) and partly in that of Colding (2006) and Colding et al. (2009). We aim at verifying to what extent similar conclusions can be drawn for Belgium.

We therefore include the following family endowments as explanatory variables: the gender, the educational attainment of father and that of mother, the number of siblings, the day of

[10] The latter definition ensures that these high educated individuals have stopped studying at the moment of the interview (see the subsequent discussion in the main text), so that we can unambiguously define their employment status.

birth within a year, and an indicator whether or not Dutch (possibly among other languages) was spoken at home. The first four variables are standard ones that are also included by the other researchers.[11] The day of birth is included as to control for age effects within a birth cohort for a given educational delay, since relative age within a birth cohort is found to positively affect educational achievements (Angrist and Krueger, 1991; Bedard and Dhuey, 2006). Finally, we control for the language spoken at home, since this is arguably a key determinant of educational progression and labour market success for minority youth.[12]

Table 2.1 reports descriptive statistics of these variables by ethnicity. These statistics confirm that immigrant youths generally are characterised by more unfavourable family endowments than natives. First and most importantly, both fathers and mothers of immigrants have successfully completed on average more than three and a half years of education less than natives. Second, in the sample immigrants are slightly (nine days on average) younger than natives. Third, in only 79% of the immigrant households Dutch (possibly among other languages) is spoken, whereas this fraction attains 98% among the natives.[13] The table indicates furthermore that immigrants have on average twice as many siblings as natives do and that the immigrant sample contains slightly more girls than that of the natives. The impact of the latter two variables on educational achievement and labour market outcomes is, however, not clearly established (Cameron and Heckman, 2001; Ryan, 2001; Pozzoli, 2009; van der Klaauw and van Vuuren, 2010).

In the literature one sometimes also controls in addition for family income, neighbourhood characteristics, indicators of regional labour market conditions, the regional level of tuition fees and grants for college enrolment. Most of these controls are not included in our analysis. First, since the analysis is restricted to one region with a homogenous and stable schooling system, there is no need to control for regional variation in the features of the educational system. Second, we cannot take family income into account, since we do not have any information on it. However, this might not be problematic, since Cameron and Heckman (2001) find that family income plays only a minor role in explaining ethnic gaps in educational attainment in the US. However, we do include the annual regional unemployment rate in Flanders as a time-varying indicator of labour market conditions. The unemployment rate of the 24 to 64 year old male population proxies the labour market conditions of the (usually) male breadwinner

[11] Belzil and Poinas (2010) include information on the occupation of father and mother instead of their level of education and they do not condition on the number of siblings.

[12] See the references to the literature in Section 2.1.

[13] Recall that the native and immigrant populations are determined on the basis of the Belgian or "non-Western" nationality of the grandmother on mother's side. The fact that in a relatively high fraction of immigrant families Dutch is spoken at home can be explained by this definition and by the exclusion from the sample of immigrants who immigrated after age three. Since Belgium consists of an important French speaking community some of the natives may only speak French at home.

Table 2.1: Summary Statistics of the Exogenous Individual Explanatory Variables by Ethnic Groups

	Flemish youth		**Immigrant youth**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
A. Female gender	0.49	0.50	0.54	0.50
B. Mother's education level	5.54	3.13	1.83	2.79
C. Father's education level	5.98	3.44	2.36	3.24
D. Number of siblings	1.52	1.18	3.58	2.38
E. Day of birth within calendar year	171.16	100.35	180.50	98.34
F. Dutch at parental home	0.98	0.12	0.79	0.41

Variables B and C measure the number of successful schooling years beyond secondary school. For instance, it is equal to 6 if the parent has successfully completed secondary education, but did not successfully complete any year of tertiary education. Variable F captures the respondent's answer to the question whether Dutch was spoken (possibly among other languages) at the parental home.

during the period that his child is in education. It is therefore included as an explanatory variable in the logit models explaining the educational outcomes. By contrast for the logit model that explains the transition from school to work, we include the youth (aged 15 to 24) unemployment rate as time-varying covariate. The evolution of these unemployment rates are described in Table 2.3 and Table 2.4 reported in the Appendix of this chapter.

2.4 Methodology

2.4.1 Econometric model

Schooling outcomes (choices and results) at any age are the outcome of previous schooling outcomes (see, e.g., Keane and Wolpin, 1997; Cameron and Heckman, 1998). The probability that a young person enrols into college or university depends on secondary school graduation which in turn depends on successively passing each secondary school grade and afterwards deciding to continue schooling. To capture this sequential aspect of economic decisions and attainments, we extend the dynamic logit model of Cameron and Heckman (2001) by explicitly distinguishing between achievements (success or failure) within each school year and the subsequent decision to continue or stop schooling. Adding these achievements to the set of educational outcomes makes it possible to study ethnic gaps in school attainment before the end of compulsory education, point before which the decision to continue schooling is irrelevant. We do this by starting modelling schooling outcomes as from the start of primary school instead of from the end of compulsory education as researchers did in past.

We propose to evaluate the relative educational performance of immigrants relative to natives based on a cumulative measure of this educational achievement: the relative fraction that passes a particular educational grade without delay, that is without having failed in any past schooling year or without having started primary school with delay (unless this delay is undone by skipping a grade during primary education). By considering this new measure of educational achievement we introduce a finer measure than in the existing literature that considers school attainment irrespectively of delay. Moreover, since ethnic gaps according to this measure may arise at much earlier ages and are dynamically linked over time, a dynamic decomposition of this measure that allows identifying when the gap arises is a valuable tool. We propose a method to realise this decomposition in Section 2.4.2. We first present the econometric model.

We model the school progression as a sequence of discrete outcomes and choices. This sequence starts at the beginning of primary school. For most pupils this occurs at age six. However, as mentioned in Section 2.2 pupils can start primary school one year earlier[14] or one year later. The starting point of our model is therefore an initial condition that models the number of years of delay (negative in case of an early start) at the start of primary schooling. Subsequently, since we only observe the grade by grade educational progression as from the start of secondary school, we group the progression made during primary school in a single stage in which we model the number of years of delay at the start of secondary education conditionally on the number of years of delay at the start of primary school. Figure 2.2 shows a graphical representation of our modelling strategy from the first grade of secondary school onwards.[15] We model for each (secondary and tertiary) schooling year, conditional on starting it, the probability of passing (P) respectively not passing (NP) and, conditional on this event, the probability to continue schooling (at a higher grade when passing or at the same grade when not passing). Finally, when leaving school, we model the probability of being employed three months later (W/NW).[16]

A couple of points should be noticed. First, as a consequence of mandatory schooling until age eighteen, the probability of continuing school is below one only from the fourth grade of secondary school onwards. This is the point from which the dynamic sorting as induced by drop-out starts playing a role (see Section 2.1). Second, each of the grade specific outcomes and choices are allowed to depend on the past history through the accumulated number of years of schooling delay and in the employment outcome in addition through the attained number of years of schooling. This introduces a second source of dynamic sorting,

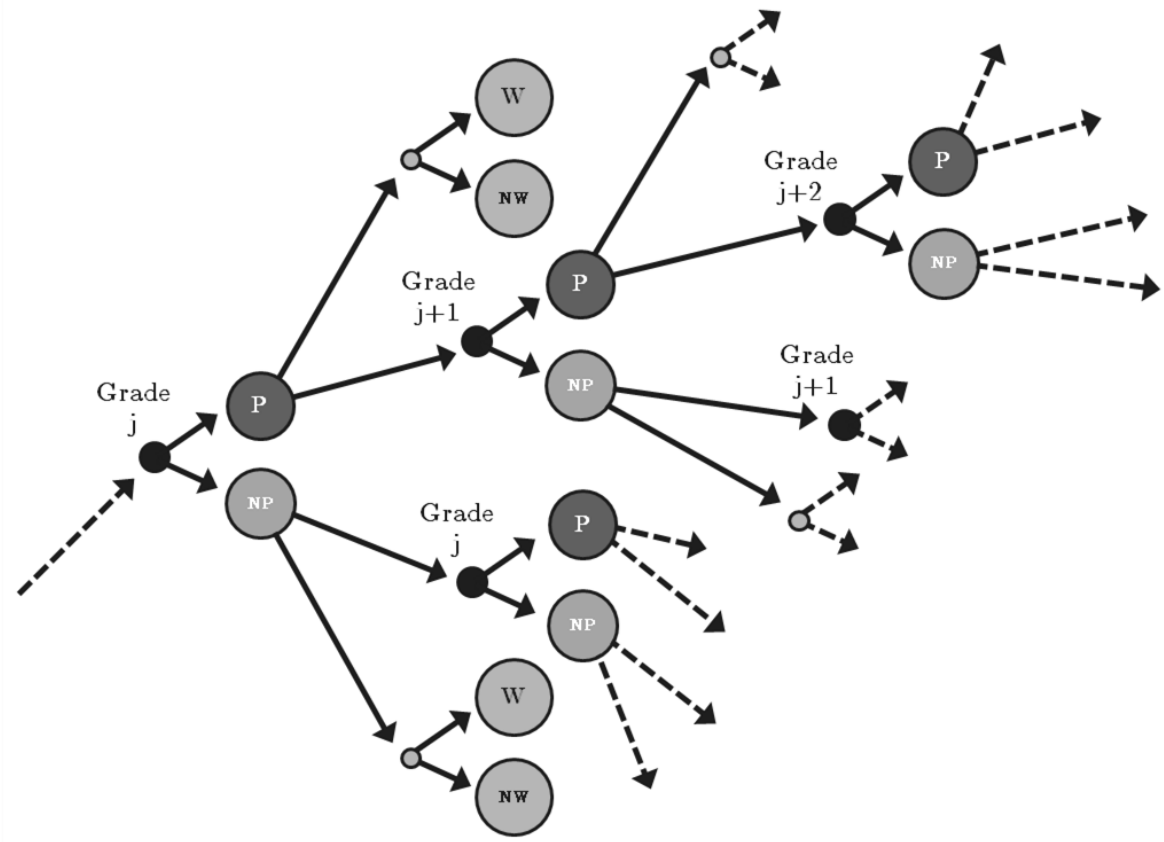
[14] This is not observed for immigrants in the data.

[15] If one has no delay, the first grade starts in September of the year that one becomes 12 years old. We continue counting when one completes mainstream secondary school after the sixth grade (without delay, this is at the school leaving age of 18) and pursues tertiary education.

[16] In a sensitivity analysis (see further) we adopt employment with a permanent contract two years after leaving school as the labour market outcome.

since students with successful schooling achievements possess more favourable unobserved endowments than those who have encountered schooling failures in the past. As to avoid selection bias induced by these sorting processes, we explicitly allow the choices and outcomes to depend on unobservable characteristics of individuals.

Figure 2.2: Transition Model



Some abbreviations are used: P (passing the grade), NP (not passing the grade), W (being employed 3 months after leaving school) and NW (being not employed 3 months after leaving school).

Econometrically, our model is specified as a sequence of (ordered and binary) logistic probabilities. Rational and forward looking agents with a schooling status determined at each time period t by their obtained schooling level, that is grade g, and their accumulated years of school delay V_t, make their "choices" from a feasible choice set.[17] We define $t \equiv -1$ and $g \equiv -1$ at the start of primary school and $t \equiv 0$ and $g \equiv 0$ at the start of secondary school. Subsequently, t increases by one unit for each year that passes since the start of secondary school and g increases by one unit for each successful schooling year that passes. A conse-

[17] We use quotation marks around the word "choices" as, properly speaking, (not) passing a grade and being employed three months after leaving school are not outcomes under full control of the modelled youth.

quence is that, for $t > 0$, $V_t = V_0 + t - g$. The dependence on the grade g respectively on the schooling delay V_t can be thought of as the memory of our model, increasing in each grade respectively at each year of grade retention.

We distinguish between seven types of outcomes O_g, depending on the considered (if still in education) or realised (if left education) grade g: (i) the years of delay at the start of primary education ($O_{-1} = 1$), (ii) the years of delay at the start of secondary education ($O_0 = 2$), (iii) the school attainment (passing or not passing) at the end of each of the six grades of secondary education ($O_g = 3$ for $g = 1, 2, \ldots, 6$), (iv) the subsequent school decisions (continuing or stopping) at the end of grades four to six of secondary education ($O_g = 4$ for $g = 4, 5, 6$), (v) the school attainments at the end of each grade of tertiary education ($O_g = 5$ for $g = 7, 8, \ldots, 12$), (vi) the subsequent school decisions at the end of each of the grades of tertiary education ($O_g = 6$ for $g = 7, 8, \ldots, 12$) and (vii) the employment status three months after leaving school ($O_g = 7$ for $g = 3, 4, \ldots, 12$). For each type of outcome O_g that we consider here, the outcomes are ordered or binary. The choice set, denoted by C^{O_g}, can therefore be given by a set of ordinal numbers: $C^{O_g} = \{0, 1, \ldots, n^{O_g}\}$, where n^{O_g} defines the number of ordered choices minus one that can be made for outcome O_g. In fact $n^{O_g} = 1$ except for $O_{-1} = 1$ and for $O_0 = 2$: $n^1 = 2$ ($n^1 = 1$) for natives (immigrants), since the number of years of delay at the start of primary school varies between -1 and 1 (0 and 1) and $n^2 = 3$ ($n^2 = 2$) for natives (immigrants), since the number of years of delay at the start of secondary school varies between -1 and 2 (0 and 2).

The optimal choice $\hat{c}^{O_g}_{g,t}$ of an individual with respect to outcome type O_g at time t in grade g (or after completing grade g in case that school is left) is then:

$$\hat{c}^{O_g}_{g,t} = c \in C^{O_g} \quad \text{if} \quad \omega_c^{O_g} < U^{O_g}_{g,t,c} \leq \omega^{O_g}_{c+1}, \tag{2.4.1}$$

where $U^{O_g}_{g,t,c}$ is the latent utility of choice c for outcome type O_g in (after) grade g at time t, and $\omega_c^{O_g}$ and $\omega^{O_g}_{c+1}$ are threshold utilities that determine the ordered choice ($\omega_0^{O_g} \equiv -\infty$ and $\omega^{O_g}_{n^{O_g}+1} \equiv +\infty$).[18] As advocated by, for instance, Heckman (1981b) and adopted by other authors, we approximate this $U^{O_g}_{g,t,c}$ by a linear index:

$$U^{O_g}_{g,t,c} = \alpha_g^{O_g} + \mathbf{Z_t'}\boldsymbol{\beta}^{\mathbf{O_g}} + \gamma^{O_g} V_t + \nu^{O_g}_{g,t,c}, \tag{2.4.2}$$

where $\alpha_g^{O_g}$ is a parameter that depends on the grade in which the outcome type O_g is consi-

[18] In the case of a binary choice the threshold $\omega_1^{O_g}$ is thus set to (minus) the constant term instead of to zero, since the constant term of the latent utility is normalised to zero. This leads to the standard logit model.

dered,[19][20] $\mathbf{Z_t}$ is a $M \times 1$ vector representing the M number of (possibly time-varying) observed strictly exogenous variables, $\boldsymbol{\beta}^{\mathbf{O_g}}$ is the vector of associated parameters, γ^{O_g} is a parameter measuring the effect of accumulated years of school delay and $\nu_{g,t,c}^{O_g}$ is unobservable from the point of view of the researcher.

We follow Cameron and Heckman (2001) by assuming that $\nu_{g,t,c}^{O_g}$ is characterised by a factor structure. However, in line with the more recent literature (Carneiro et al., 2003; Heckman and Navarro, 2007; Fruehwirth et al., 2011), we generalise by allowing that the factor "loadings" depend on the treatment status, which in our case is the number of years of schooling delay V_t:

$$\nu_{g,t,c}^{O_g} = \delta^{O_g}\eta + \phi^{O_g} V_t \eta + \epsilon_{g,t,c}^{O_g}, \tag{2.4.3}$$

in which δ^{O_g} and ϕ^{O_g} are outcome type specific coefficients and $\epsilon_{g,t,c}^{O_g}$ is the i.i.d. error term, and η is a random individual specific effect that is independent across people and that captures unobserved "abilities" affecting all outcomes considered in the model. Assuming that the unobserved determinants are common to all outcomes is restrictive, but, as shown in the aforementioned literature, the advantage of doing so is that it allows that the effect of schooling delay depends on unobserved heterogeneity,[21] as it does by the introduction of the second term in Equation 2.4.3, and that this treatment heterogeneity can be identified non-parametrically. Fruehwirth et al. (2011) argue that this may be important and indeed find evidence of heterogeneous reactions to grade retention.

Identification of treatment heterogeneity in the effect of schooling delay does not require an exclusion restriction if the outcome in the first period, that is the number of years of delay at the start of primary school, is free of selection. This means that $\mathbf{Z_t}$ should be independent of η for all g, t and choice sets C^{O_g}. Note that this does not mean that conditional on past choices beyond the start of primary school $\mathbf{Z_t}$ is independent of η, since, as mentioned in Section 2.1, dynamic sorting will induce negative correlation between favourable observed determinants of the educational outcomes that we consider, and the unobserved η. This is because pupils with unfavourable observed endowments experience successful educational outcomes only if these unfavourable endowments are compensated for by favourable unobserved endowments (Cameron and Heckman, 1998). The independence assumption rather means that the unobserved abilities capture factors that are independent of observed family endowments. We

[19]The parameter corresponding to the first grade that can be observed within the outcome type is taken as the reference grade. It is normalised to zero, since it cannot be separately identified from the threshold utilities $\omega_c^{O_g}$.

[20]For the school outcomes in tertiary education ($O_g = 5$ and $O_g = 6$) and for the employment decision ($O_g = 7$) we restrict the dependence to be linear in g.

[21]This is labeled "essential heterogeneity" by Heckman et al. (2006).

improve in this respect on the existing literature by starting modelling the schooling progression from a much earlier point: at the start of primary school rather than at the end of mandatory schooling, usually around age 16. Consequently, in our approach the effect of observed family endowments on the educational outcomes is purged from the bias induced by the negative correlation with the unobserved determinants of successful schooling outcomes (no delay) during the period of mandatory schooling.[22]

We assume that $\epsilon_{g,t,c}^{O_g}$ is logistically distributed, independent of η for all O_g, g, t and c, and therefore we can write the probability of an outcome as:

$$\begin{aligned}\Pr(\hat{c}_{g,t}^{O_g} = c|\mathbf{Z_t}, V_t, g, O_g, \eta; \boldsymbol{\theta}) = \\ &\frac{\exp(\omega_{c+1}^{O_g} - \alpha_g^{O_g} - \mathbf{Z_t'}\boldsymbol{\beta}^{\mathbf{O_g}} - \delta^{O_g}\eta - \phi^{O_g}V_t\eta)}{1+\exp(\omega_{c+1}^{O_g} - \alpha_g^{O_g} - \mathbf{Z_t'}\boldsymbol{\beta}^{\mathbf{O_g}} - \delta^{O_g}\eta - \phi^{O_g}V_t\eta)} \\ &- \frac{\exp(\omega_{c}^{O_g} - \alpha_g^{O_g} - \mathbf{Z_t'}\boldsymbol{\beta}^{\mathbf{O_g}} - \delta^{O_g}\eta - \phi^{O_g}V_t\eta)}{1+\exp(\omega_{c}^{O_g} - \alpha_g^{O_g} - \mathbf{Z_t'}\boldsymbol{\beta}^{\mathbf{O_g}} - \delta^{O_g}\eta - \phi^{O_g}V_t\eta)},\end{aligned} \tag{2.4.4}$$

in which we denote the vector of unknown parameters by $\boldsymbol{\theta}$. The likelihood contribution $l_i(\mathbf{Z_t}, V_t, \eta; \boldsymbol{\theta})$ for any sampled individual, conditional on the unobservable η, is then constructed by the product of the probabilities of the school and labour market outcomes as expressed by (2.4.4) realised in each time period t between the start of primary school, and the labour market entry or the highest grade that the respondent has attained at the interview date at age 23.

Following Heckman and Singer (1984), we adopt a non-parametric discrete distribution for the unobserved random variable η. We assume that this distribution is characterised by an a priori unknown number of K points of support η_k to which are assigned probabilities $p_k(\lambda)$ specified as logistic transforms:

$$p_k(\lambda) = \frac{\exp(\lambda_k)}{\sum_{j=1}^{K}\exp(\lambda_j)} \quad \text{with} \quad k = 1, 2, \ldots, K; \lambda \equiv [\lambda_1, \lambda_2, \ldots, \lambda_K]\prime; \lambda_1 = 0. \tag{2.4.5}$$

Hence, the unconditional individual likelihood contribution for an agent i is:

$$l_i(\mathbf{Z_{it}}, V_{it}; \boldsymbol{\theta}, \boldsymbol{\lambda}) = \sum_{k=1}^{K} p_k(\boldsymbol{\lambda}) l_i(\mathbf{Z_{it}}, V_{it}, \eta_k; \boldsymbol{\theta}). \tag{2.4.6}$$

[22] The outcome scholastic ability test (AFQT) that Cameron and Heckman (2001) add as control in part of their models may capture these unobserved determinants of early schooling outcomes. Belzil and Poinas (2010) add an indicator for grade repetition in primary school to proxy for these unobservables, but do not take the endogeneity of this variable into account.

Since the estimation is conducted separately on the native and immigrant sample, the log-likelihood function is the logarithm of these unconditional likelihood contributions summed over all N_j ($j = N, I$) sampled individuals, where N_N (N_I) stands for the number of sampled native (immigrant) individuals. This is maximised with respect to the unknown parameters. In order to determine the number of points of support we follow common practice (see, e.g., Belzil and Poinas, 2010) and select the number of mass points by choosing the model that minimises an information criterion. In our case (see Section 2.5) the optimal choice minimises both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

2.4.2 Goodness-of-Fit and Decomposition Strategy

In order to answer our main research questions, we develop a "counterfactual" decomposition strategy in the spirit of Machado and Mata (2005) aimed at disentangling the respective importance of pure ethnic differences versus differences in observed endowments between natives and immigrants in explaining the observed ethnic gap in educational attainment (conditioning on educational delay or not) and in school-to-work transitions. Moreover, we propose a method to decompose the residual pure ethnic gap according to the moment at which it is generated. We first propose a simulation method to test the model's capacity to fit the ethnic gaps of a particular outcome. Subsequently, we discuss how, based on this simulation method, we can realise the aforementioned decompositions.

The method simulates the model on random samples each of size R ($R = 5,000$ in the application) of the native and immigrant samples that were used for estimation. Let $\mathbf{Z_N}$ and $\mathbf{Z_I}$ be $R \times M$ matrices storing the R random draws from the native respectively immigrant youth observed exogenous endowment distributions and from the time-varying strictly exogenous variables. Let $\hat{\boldsymbol{\theta}}_\mathbf{N}$ and $\hat{\boldsymbol{\theta}}_\mathbf{I}$ denote the native respectively immigrant parameter estimates including the ones that refer to the endogenous variables (grade g and schooling delay V_t) and the unobserved heterogeneity distribution. In this simulation, the endogenous variables take on all possible values weighted by their predicted probability of occurrence according to the parameter estimates. The observed gap as represented by the log expected odds ratio between natives and immigrants of a particular outcome (for instance, the probability of passing sixth grade of secondary education) can then be predicted by simulation as follows:

$$\log\left(\frac{E_{\mathbf{Z_N}} \Pr[\hat{c}_{g,t}^{O_g} = c|\mathbf{Z_N}; \hat{\boldsymbol{\theta}}_N]}{E_{\mathbf{Z_I}} \Pr[\hat{c}_{g,t}^{O_g} = c|\mathbf{Z_I}; \hat{\boldsymbol{\theta}}_I]}\right), \tag{2.4.7}$$

where $\Pr[\hat{c}_{g,t}^{O_g} = c|.;.]$ is the probability that the particular outcome $\hat{c}_{g,t}^{O_g} = c$[23] is realised

[23]We may consider outcomes that are not conditioned on any particular time period t or grade g. In that

according to the model simulation and $E_{\mathbf{Z_N}}$ and $E_{\mathbf{Z_I}}$ the expectations over the distributions of $\mathbf{Z_N}$ respectively $\mathbf{Z_I}$. Note that the existing literature expresses the gap in terms of the absolute difference in the probabilities of realisation of a particular outcome instead of in terms of the log odds ratio. Expressing the ethnic gap in terms of the log odds ratio will prove to be useful in decomposing the residual pure ethnic gap according to the moments that it is generated. The 95% confidence intervals of these (and subsequent) log odds ratios are constructed by simulation, the steps of which are given in the Appendix of this chapter.

We now propose the following decomposition of the predicted ethnic gap, as expressed by the log odds ratio in Equation (2.4.7) into the sum of an "explained" and a "residual pure ethnic gap":

$$\log\left(\frac{E_{\mathbf{Z_N}}\Pr[\hat{c}_{g,t}^{O_g}=c|\mathbf{Z_N};\hat{\boldsymbol{\theta}}_N]}{E_{\mathbf{Z_I}}\Pr[\hat{c}_{g,t}^{O_g}=c|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_I]}\right)= \log\left(\frac{E_{\mathbf{Z_N}}\Pr[\hat{c}_{g,t}^{O_g}=c|\mathbf{Z_N};\hat{\boldsymbol{\theta}}_N]}{E_{\mathbf{Z_I}}\Pr[\hat{c}_{g,t}^{O_g}=c|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_N]}\right)+\log\left(\frac{E_{\mathbf{Z_I}}\Pr[\hat{c}_{g,t}^{O_g}=c|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_N]}{E_{\mathbf{Z_I}}\Pr[\hat{c}_{g,t}^{O_g}=c|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_I]}\right), \tag{2.4.8}$$

The first term on the right-hand side of (2.4.8) is the gap that can be explained by differences in the observed endowments $\mathbf{Z_N}$ and $\mathbf{Z_I}$ evaluated by using the parameters as estimated on the native sample, $\hat{\boldsymbol{\theta}}_N$. The last term in Equation (2.4.8) defines the residual "pure ethnic gap". It reflects the gap induced by differences in the parameter estimates, including the ones that relate to the unobservables, between native and immigrant youth.[24] It is the latter gap that has been found in the literature to be negligible (Cameron and Heckman, 2001; Belzil and Poinas, 2010) or reduced substantially (Colding, 2006; Colding et al., 2009) as compared to the observed gap, both in terms of educational outcomes as in indicators of successful school-to-work transitions and which has led researchers to conclude that the ethnic gap in outcomes is not due to discrimination, but rather to a shortfall in skills, natives and immigrants alike.

In the empirical analysis below, we will show that, in line with the existing literature, the pure ethnic gap in educational outcomes (leaving schooling delay unspecified) is indeed relatively small or even disappears if we consider the gap at the enrolment in higher education.

case one would take the expectation of the probability over this dimension. Alternatively, we may consider outcomes in which the number of years of schooling delay is specified. For instance, in the empirical analysis we consider schooling outcomes at particular grades (passing a grade or continuing education after passing that grade) that are attained without schooling delay. Then for some $g > 0$ the probability of interest is given by $\Pr[\hat{c}_{g,g}^{O_g} = 1|.;.]$, since after starting secondary school a schooling outcome can only be attained without delay if the outcome is successful at all $t = g$.

[24]An alternative decomposition strategy consists in evaluating the endowment gap at the immigrant parameter estimates and the pure ethnic gap at the values of the native covariates. By conditioning on the endowments of the immigrant youth, as we do in Equation (2.4.8) and in the benchmark empirical analysis, we focus on the gap for youth with typical immigrant characteristics, so at the lower end of the socioeconomic scale. We implement the alternative decomposition as a sensitivity analysis (see further).

However, if we consider the realisation of schooling outcomes without delay or success in the school-to-work transition, this is no longer the case. Then, in order to identify the cause of this gap, it is useful to determine the moment at which it originates. We therefore propose a procedure that decomposes the pure ethnic gap into parts that depend on the moments that it is generated. It uses the fact that a particular educational attainment can only be realised if at earlier stages educational outcomes were successful: educational attainments realise sequentially. This means that we can write the probability of a successful educational outcome as a product of conditional probabilities in which the conditioning is each time related to a successful educational outcome at an earlier stage. If we write the ethnic gaps in terms of log odds ratios, we can therefore decompose a successful educational outcome at a particular stage in a sum of log odds ratios of the conditional probabilities of educational success in earlier stages.

We explain the decomposition procedure on the basis of an example. Suppose that we are interested in identifying when the pure ethnic gap in the fraction that passes fifth grade of secondary school ($\hat{c}_5^3 = 1$) originates. We therefore aim at decomposing the pure ethnic gap of this outcome, as defined on the left-hand side of the equality in Equation (2.4.9):[25]

$$
\begin{aligned}
&\log\left(\frac{E_{\mathbf{Z_I}}\Pr[\hat{c}_5^3 = 1|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_N]}{E_{\mathbf{Z_I}}\Pr[\hat{c}_5^3 = 1|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_I]}\right) = \\
&\quad \log\left(\frac{E_{\mathbf{Z_I}}\Pr[\hat{c}_4^3 = 1|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_N]}{E_{\mathbf{Z_I}}\Pr[\hat{c}_4^3 = 1|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_I]}\right) \\
&\quad + \log\left(\frac{E_{\mathbf{Z_I}}\Pr[\hat{c}_4^4 = 1|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_N]/E_{\mathbf{Z_I}}\Pr[\hat{c}_4^3 = 1|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_N]}{E_{\mathbf{Z_I}}\Pr[\hat{c}_4^4 = 1|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_I]/E_{\mathbf{Z_I}}\Pr[\hat{c}_4^3 = 1|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_I]}\right) \\
&\quad + \log\left(\frac{E_{\mathbf{Z_I}}\Pr[\hat{c}_5^3 = 1|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_N]/E_{\mathbf{Z_I}}\Pr[\hat{c}_4^4 = 1|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_N]}{E_{\mathbf{Z_I}}\Pr[\hat{c}_5^3 = 1|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_I]/E_{\mathbf{Z_I}}\Pr[\hat{c}_4^4 = 1|\mathbf{Z_I};\hat{\boldsymbol{\theta}}_I]}\right).
\end{aligned}
\tag{2.4.9}
$$

Notice first that this gap cannot realise before the start of fourth grade of secondary school, since by compulsory schooling until age 18 nobody leaves school before this moment. This means that, if we ignore schooling delays for the moment, the first moment at which the ethnic gap can differ from zero is by not passing fourth grade of secondary school ($\hat{c}_4^3 = 0$). The gap that is generated at that moment is expressed by the first term on the right-hand side of the equality in Equation (2.4.9). Subsequently, the gap can further originate from deciding not to start fifth grade ($\hat{c}_4^4 = 0$), conditional on having passed fourth grade ($\hat{c}_4^3 = 1$). This source of the gap is quantified by the second term on the right-hand side of the equality sign in Equation (2.4.9). Finally, the source of the gap can originate from not having passed fifth

[25]Note that we do not condition the choice on t, meaning that we implicitly average over t.

grade of secondary school ($\hat{c}_5^3 = 0$), conditional on deciding to start fifth grade ($\hat{c}_4^4 = 1$). The sum of the terms on the right-hand side of the equality in (2.4.9) is by construction equal to the term on the left-hand side. This means that we can determine the relative importance of the moments at which the gap originates. It is not difficult to generalise this procedure for other outcomes and longer sequences of outcomes. This is what we do in the empirical application.

2.5 Results

We estimate the econometric model separately for native and immigrant youth. As mentioned in Section 2.4.1, we did this by gradually adding points of support until the log-likelihood value of the model failed to increase. Subsequently, we chose the best fitting model according to two information criteria. Table 2.5 in the Appendix of this chapter reports the log-likelihood, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) values of the model according to the estimated number of points of support of the heterogeneity distribution. All information criteria are lower for the models that control for unobserved heterogeneity than for a simpler scheme that ignores unobserved heterogeneity. The lowest AIC and BIC values are obtained with five points of support for both the native and the immigrant youth. Since the main aim of this study is to decompose the ethnic gap in educational attainment and in school-to-work transitions, and since we have estimated a large number of parameters (nearly one hundred for each ethnic group),[26] we do not report the estimated parameters. These are available on request. Instead, we first report a goodness-of-fit analysis with respect to the main outcomes of interest. Subsequently, we present the decomposition along the lines of our presentation in Section 2.4.2. In Section 2.5.3 we specifically focus on the role that language plays in this composition. Finally, we briefly present some sensitivity analyses in Section 2.5.4.

2.5.1 Goodness of Fit

We simulate the fraction of both native and immigrant youth who realise a variety of schooling and labour market outcomes. The difference between the first two columns of Table 2.2 (or the two first panels of Figure 2.3 for the main outcomes of interest) measures the goodness of fit of our model. Column A describes the observed gaps between native and immigrant youth in our data by means of log odds ratios of the native respectively the immigrant schooling and labour market outcome probabilities. A positive number means that native youth is more

[26] When comparing these numbers of parameters to the size of our native and immigrant research sample one should take in mind that we observe (and model) multiple observations for each individual.

likely to achieve the considered outcome. Since $\log(1+x) \cong x$, these log odds ratios mirror the proportional gaps between native and immigrant youth as reported in the descriptive analysis in Section 2.3.2. Column B describes the corresponding gaps based on the simulations that were described in Section 2.4.2. The main outcomes of interest, introduced as motivating gaps in Section 2.3.2, are denoted in bold and shaded in grey. The decomposition of the gaps according to the moments at which they originate are reported in the lines below the main outcomes of interest (neither in bold nor shaded in grey). They sum to the main outcome of which they are components. The fit is very good, since in all cases, the actual gap lies within the 95% confidence interval of the simulated outcome.

2.5.2 The Role of Family Endowments in Explaining the Gaps

Column C of Table 2.2 presents the pure ethnic gap for a range of outcomes as obtained by conducting the counterfactual simulations outlined in Section 2.4.2, by equating the observed family endowments of both ethnic groups to the immigrant level. We first discuss the findings with respect to schooling outcomes without specifying the potential delay with which these are attained. This is common in the literature. Subsequently, we contrast these results to those obtained for the same schooling outcomes, but restricting that these outcomes should be realised without schooling delay. Finally, we consider the pure ethnic gap in the school-to-work transition. We decompose the pure ethnic gaps of all considered outcomes as to determine the key moments at which these pure ethnic gaps are generated. First, we focus on the probability of passing the last (sixth) grade of secondary school. Equating observed endowments reduces the log odds ratio of this (predicted) probability from 0.21 to 0.07. This means that if a native and an immigrant child are equal in terms of individual and household characteristics the native child is about 7% more likely to complete secondary education. Second, we consider the probability of enrolling in tertiary education. In this case conditioning on observed endowments completely eliminates the 35 points wide predicted ethnic gap. These results are completely in line with the literature mentioned in Section 2.1. Differences in family background explain the gap in educational attainment to a large extent and especially so for higher levels of education.

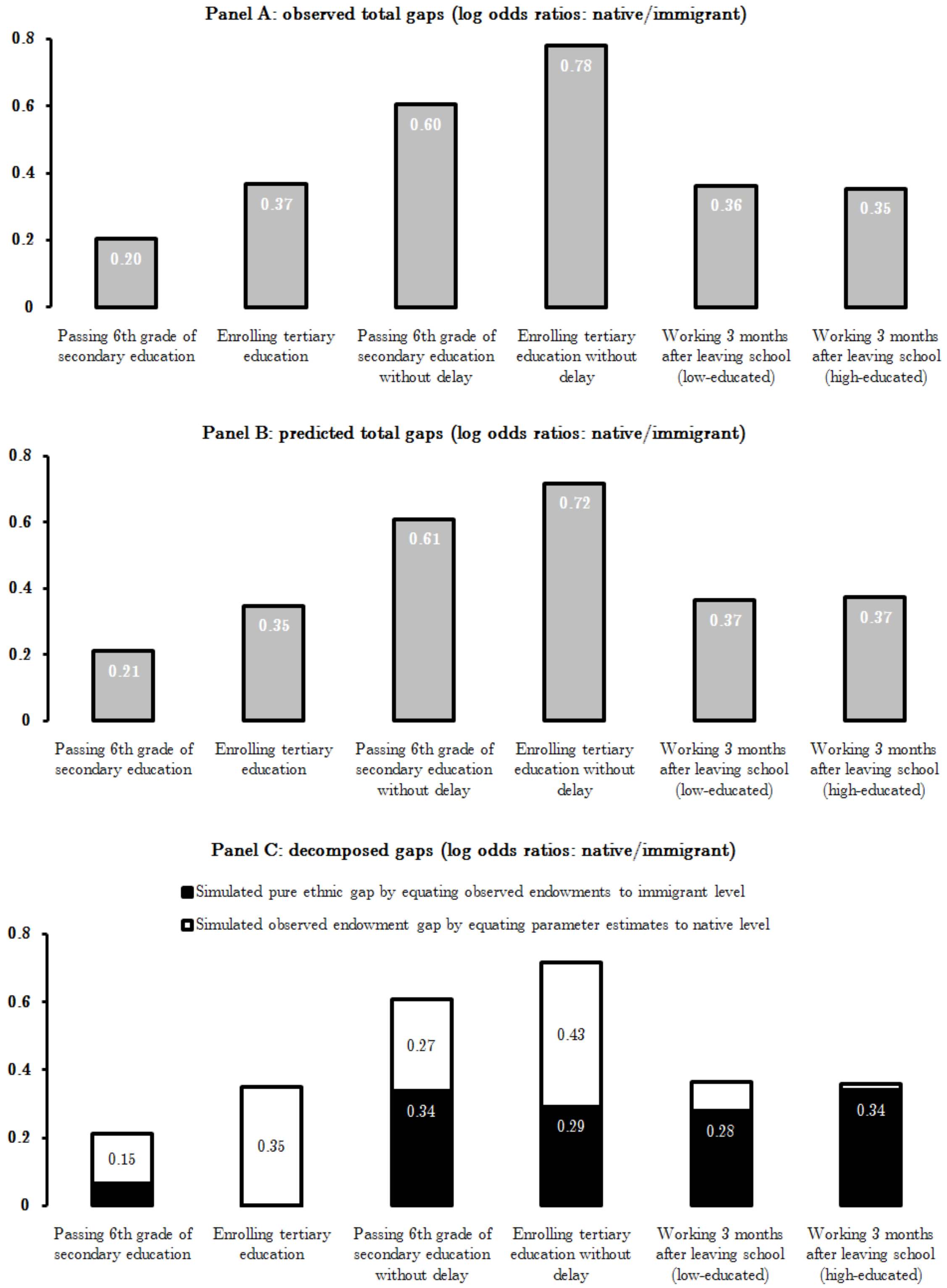
In the lines below these main outcomes denoted in bold in Table 2.2 the pure ethnic gaps are further decomposed. First, consider the 7% pure ethnic gap of successfully completing secondary school. Decomposing this gap by analogy with the lines of Equation (2.4.9) in Section 2.4.2, we find that the major part of this gap is generated by a higher dropout rate for immigrants after successfully completing fourth grade of secondary school and by a higher fraction

Table 2.2: Simulation Results

	A. Observed gap	B. Predicted gap		C. Pure ethnic gap by equating observed endowments to immigrant level		D. Pure ethnic gap by equating observed endowments to native level		E. Observed endowment gap: language	
Passing SE6	**0.20**	**0.21**	**[0.18, 0.31]**	**0.07****	**[0.02, 0.17]**	**0.07*****	**[0.04, 0.17]**	**0.00**	**[-0.03, 0.03]**
Passing SE4	0.05	0.03	[0.02, 0.06]	0.01	[-0.01, 0.04]	0.01***	[0.00, 0.03]	-0.00	[-0.01, 0.00]
Starting SE5 | passing SE4	0.04	0.03	[0.02, 0.07]	0.02**	[0.00, 0.05]	0.01***	[0.00, 0.03]	-0.00	[-0.01, 0.00]
Passing SE5 | starting SE5	0.03	0.04	[0.02, 0.06]	0.00	[-0.02, 0.03]	0.01*	[-0.00, 0.03]	-0.00	[-0.01, 0.00]
Starting SE6 | passing SE5	0.04	0.05	[0.03, 0.10]	0.01	[-0.02, 0.06]	0.01**	[0.00, 0.05]	0.01	[-0.00, 0.02]
Passing SE6 | starting SE6	0.04	0.06	[0.03, 0.10]	0.03	[-0.01, 0.07]	0.03**	[0.01, 0.07]	-0.00	[0.01, 0.00]
Starting TE	**0.37**	**0.35**	**[0.28, 0.48]**	**-0.00**	**[-0.08, 0.15]**	**0.10****	**[0.02, 0.26]**	**0.03**	**[-0.03, 0.07]**
Passing SE6	0.20	0.21	[0.18, 0.31]	0.07**	[0.02, 0.17]	0.07***	[0.04, 0.17]	0.00	[-0.03, 0.03]
Starting TE | passing SE6	0.16	0.14	[0.08, 0.21]	-0.07*	[-0.13, 0.01]	0.02	[-0.03, 0.11]	0.02*	[-0.00, 0.02]
Passing SE6 without D	**0.60**	**0.61**	**[0.51, 0.81]**	**0.34*****	**[0.22, 0.55]**	**0.32*****	**[0.20, 0.58]**	**-0.04**	**[-0.10, 0.02]**
Starting PE without D	0.04	0.04	[0.03, 0.16]	0.03***	[0.01, 0.15]	0.00	[-0.01, 0.32]	0.01	[-0.01, 0.05]
Starting SE1 without D | starting PE without D	0.24	0.23	[0.14, 0.32]	0.12**	[0.02, 0.21]	0.10	[-0.11, 0.23]	-0.01	[-0.05, 0.03]
Starting SE2 without D | starting SE1 without D	0.06	0.05	[0.02, 0.10]	0.04**	[0.01, 0.09]	0.03**	[0.00, 0.07]	-0.01**	[-0.02, -0.00]
Starting SE3 without D | starting SE2 without D	0.02	0.02	[-0.00, 0.06]	0.01	[-0.02, 0.05]	0.01	[-0.02, 0.06]	-0.01**	[-0.01, -0.00]
Starting SE4 without D | starting SE3 without D	0.10	0.07	[0.03, 0.12]	0.04**	[0.00, 0.09]	0.04**	[0.01, 0.10]	-0.01**	[-0.02, 0.00]
Starting SE5 without D | starting SE4 without D	0.06	0.10	[0.05, 0.17]	0.07***	[0.02, 0.14]	0.07***	[0.02, 0.14]	-0.01	[-0.02, 0.01]
Starting SE6 without D | starting SE5 without D	0.06	0.05	[0.01, 0.11]	0.00	[-0.04, 0.06]	0.02	[-0.02, 0.08]	0.00**	[-0.01, -0.00]
Passing SE6 without D | starting SE6 without D	0.03	0.05	[0.01, 0.10]	0.03	[-0.01, 0.08]	0.04**	[0.00, 0.08]	-0.01**	[-0.01, -0.00]
Starting TE without D	**0.78**	**0.72**	**[0.61, 0.94]**	**0.29*****	**[0.17, 0.53]**	**0.35*****	**[0.21, 0.63]**	**-0.02**	**[-0.09, 0.05]**
Passing SE6 without D	0.60	0.61	[0.51, 0.81]	0.34***	[0.22, 0.55]	0.32***	[0.20, 0.58]	-0.04	[-0.10, 0.02]
Starting TE | passing SE6 without D	0.18	0.11*	[0.06, 0.19]	-0.05	[-0.11, 0.05]	0.03	[-0.01, 0.11]	0.03 **	[0.00, 0.05]
Empl. 3m. after LS | LS before passing SE6	0.38	0.43	[0.22, 0.67]	0.33***	[0.10, 0.58]	0.19*	[-0.04, 0.50]	0.06	[-0.02, 0.13]
Empl. 3m. after LS | LS with only SE6 d.	0.30	0.29	[0.17, 0.43]	0.23***	[0.09, 0.37]	0.09	[-0.06, 0.31]	0.06*	[-0.01, 0.10]
Empl. 3m. after LS | LS with only SE6 d. or lower	**0.36**	**0.37**	**[0.22, 0.52]**	**0.28*****	**[0.11, 0.45]**	**0.12**	**[-0.04, 0.37]**	**0.06**	**[-0.01, 0.11]**
Empl. 3m. after LS | LS with TE1 - TE4 d. and ⩽ 1 year of D	**0.35**	**0.37**	**[0.20, 0.57]**	**0.34*****	**[0.16, 0.54]**	**0.20****	**[0.04, 0.44]**	**0.03**	**[-0.03, 0.09]**
Empl. 3m. after LS | LS with TE3 or TE4 d. and ⩽ 1 year of D	0.34	0.36	[0.15, 0.64]	0.33***	[0.11, 0.61]	0.21**	[0.02, 0.49]	0.03	[-0.03, 0.11]

The described gaps are log odds ratios of the observed and simulated fractions of the native respectively the immigrant youth. Some abbreviations are used: PE (primary education), SE1 (1st grade of secondary education), SE2 (2nd grade of secondary education), SE3 (3rd grade of secondary education), SE4 (4th grade of secondary education), SE5 (5th grade of secondary education), SE6 (6th grade of secondary education), TE (tertiary education), TE1 (1th grade of tertiary education), TE2 (2th grade of tertiary education), TE3 (3th grade of tertiary education), TE4 (4th grade of tertiary education), D (delay), LS (leaving school), m. (months), Empl. (employed) and d. (degree). 95% confidence intervals, in brackets, were calculated using 999 random draws from the distributions of the underlying model's estimated parameters. In Column B ***(**)((*)) indicates significantly difference from the value in column A at the 1% (5%) ((10%)) level. In column C, column D and column E ***(**)((*)) indicates significantly difference from 0 at the 1% (5%) ((10%)) level.

Figure 2.3: Decomposition of Observed Schooling and Labour Market Gaps



that does not successfully pass sixth grade after starting it. At these moments 2% respectively 3% of the 7% total pure ethnic gap originates. However, these interpretations are hazardous, since these subcomponents are small and not very precisely estimated. We conclude that it is difficult to assign a precise moment at which this pure ethnic gap emerges. But this is not so problematic given that the total pure ethnic gap is small anyway. Second, the decomposition of the zero pure ethnic gap with regards to enrolment in tertiary education learns that it arises by a pure ethnic advantage of 7% that arises for immigrants in this enrolment decision conditional on secondary school completion. This advantage erases the aforementioned 7% gap in secondary school completion. Cameron and Heckman (2001) report similar findings for the US. This ethnic advantage may reflect that for a given level of socioeconomic background immigrant youths possess "better" unobserved characteristics (ability or motivation) than native youth, since parents of immigrants may, as a consequence of fewer opportunities, have attained a lower level of educational attainment than parents of natives. In other words, for a same level of educational attainment of their parents immigrants may be more able or more motivated than natives, and may therefore more likely enrol in higher education. However, further research is required to confirm this interpretation.

We now consider the same two educational outcomes, but restrict these outcomes to be realised without schooling delay. As already mentioned in Section 2.3.2, this restriction substantially increases the total ethnic gaps for these outcomes. The log odds ratio for completing secondary education increases from 0.21 to 0.61, and for enrolling in tertiary education from 0.35 to 0.72. More importantly, even if these gaps are substantially reduced if observed family endowments are controlled for, in contrast to the case in which no schooling delay is specified, the pure ethnic gaps remain substantial: 0.34 respectively 0.29. This is an important finding, since it means that ethnic schooling gaps, in particular gaps in schooling delay, cannot be eliminated by focussing policy to disadvantaged groups irrespectively of their ethnic background. It also suggests that similar conclusions might arise with respect to other measures of educational achievement within a particular level of educational attainment, such as scores on standardised tests of achievement, implying that our findings may also be relevant for countries in which grade retention is less wide spread than in Belgium.

In the lines below these log odd ratios are decomposed according to the grade in which they originate. First, observe that the lower pure ethnic gap for enrolment in higher education without any delay reflects a pure ethnic advantage for immigrants that was also detected in case we did not specify the schooling delay. This 5% advantage is, however, no longer significantly different from zero. More interesting is to get an insight into the grades at which the 0.34 gap in the log odds ratio in secondary school completion is generated. Since the pure ethnic gap matters predominantly if schooling delay is taken into account, we know that it is retention and not drop out that is the main driver of the pure ethnic gap at each grade. We therefore in this decomposition make no distinction between passing and the decision to continue schooling within each grade. From Table 2.2 we deduce that, even if the pure ethnic gap seems to emerge relatively gradually throughout the educational progression, the major part originates in secondary school. The components of the total odds ratio assigned to secondary school sum to 0.19, while those generated during or at the start of primary school sum to 0.15 only. This is, however, not unexpected, since retention is in Belgium more frequently used in secondary school than in primary school. The data did not allow determining at which particular grades of primary school the pure gap emerges. We can only conclude that the ratio attained already 0.03 at the start of primary school, so that the remaining 0.12 is generated during the first six compulsory schooling years. By contrast, within secondary school we can identify the evolution of the pure ethnic gap by grade. There we can (again) clearly identify fourth grade as a major source of the pure gap: 0.07 of the total 0.34 originates in that grade. This means that more than 20% of the total pure gap that is generated between the start of primary school and the end of secondary school can be assigned to this grade. This is an important finding, since it informs to which grade analysts should target attention to get a better understanding of where the pure ethnic gap originates from. The analysis learns in addition that the first, third and last year of secondary school are critical as well, but to a lesser extent.

Finally, we consider the pure ethnic gap in being employed three months after leaving school given a particular level of school attainment. As can be deduced from Column C, and in contrast with the findings of Belzil and Poinas (2010), for all levels of education equating observed endowments between natives and immigrants hardly reduces the ethnic gaps in the transition to work. Independently of the level of education, a native school-leaver is about 30% more likely to be employed three months after leaving school compared with an immigrant school-leaver with the same observed endowments. This suggests that, in contrast to France, discrimination of ethnic minorities may affect labour market outcomes of ethnic minorities in Belgium (Flanders). Notwithstanding that, as discussed below, alternative and complementary interpretations are possible, these results square with the findings outlined in Chapter 3.

Contrary to the existing literature, we thus find evidence for important pure ethnic gaps in educational outcomes and in the transition from school to work. For the educational outcomes this is a consequence of explicitly taking schooling delays into account. These pure ethnic gaps need to be interpreted with caution and may not be simply identified with proof of discrimination. Discrimination is just one explanation among others. We mention a number of alternative explanations without aiming to be comprehensive. First, the pure ethnic gap may partly be caused by ethnic differences in preferences or expectations. Constant et al. (2010) provide evidence on divergence in economic preferences and attitudes between natives and second generation migrants in Germany. Migrants are found to be, for instance, less risk-averse. Moreover, they conclude that these differentials matter in terms of employment probabilities two months after unemployment entry. More evidence on the importance of preferences and expectations in explaining school attainment and labour market outcome gaps is provided by, for instance, Hennessey et al. (2008), Filippin (2009) and Zaiceva and Zimmermann (2010). Second, a recent literature deals with the role of ethnic networks in explaining labour market outcome gaps (Winters et al., 2001; Mahuteau and Juanankar, 2008; Yamauchi and Tanabe, 2008; Zenou, 2011) and diverging school outcomes can be related to class and school segregation of migrants as a consequence of the concentration of immigrants in certain neighbourhoods (Colding, 2006; Colding et al., 2009; Dustmann et al., 2010; Agirdag et al., 2011). Third, in the absence of specific teaching incentive programs for disadvantaged or immigrant groups (Dustmann et al., 2010), teachers may pay more attention to native (advantaged) groups. Fourth, part of the pure ethnic gap can be related to differentials in the unobserved "ability" distributions between natives and immigrants (Cameron and Heckman, 2001). Finally, part of the pure ethnic gap may be induced by differences in language proficiency that are not captured by the language usage variable that was controlled for in the analysis. We turn to a discussion of this point in the next section.

2.5.3 Gap Closing Role for Language?

Column E of Table 2.2 presents evidence on language spoken at parental home as a source of schooling and first labour market gaps between native and immigrant youth. In the spirit of Cameron and Heckman (2001) these gaps are obtained by estimating the following ratio:

$$\log\left(\frac{E_{\mathbf{Z}_\mathbf{I}^*}\Pr[\hat{c}_{g,t}^{O_g}=c|\mathbf{Z}_\mathbf{I}^*;\hat{\boldsymbol{\theta}}_I]}{E_{\mathbf{Z}_\mathbf{I}}\Pr[\hat{c}_{g,t}^{O_g}=c|\mathbf{Z}_\mathbf{I};\hat{\boldsymbol{\theta}}_I]}\right), \tag{2.5.1}$$

in which $\mathbf{Z}_\mathbf{I}^*$ differs from $\mathbf{Z}_\mathbf{I}$ by the value of the variable capturing usage of Dutch at the parental home. This value is set in $\mathbf{Z}_\mathbf{I}^*$ for all draws to the mean native level.

From Column E of Table 2.2 we deduce that speaking Dutch at home plays hardly any role in closing the observed ethnic gap in educational attainment. The contribution of language is always very small and mostly not significantly different from zero. Conditional on graduation from secondary school, it explains 2 of the 14 points predicted log odds in enrolling in tertiary education (without delay). Noticing in addition that by controlling for other observed family endowments this predicted gap even turns into a 7 points pure ethnic advantage for immigrants, this contribution is small. Our estimates also indicate that not speaking Dutch at home is rather an advantage than a disadvantage for immigrants to continue schooling without delay, since it decreases the predicted ethnic gap significantly by one percentage point in all but fourth and fifth grade of secondary school. This is consistent with the findings of Dustmann et al. (2010) indicating that in the UK during secondary school the educational achievement of ethnic minority pupils for whom English is not the mother tongue improves more relative to White British pupils than that of ethnic minority pupils for whom English is the mother tongue. However, globally these ethnic advantages are no longer significantly different from zero if the unconditional gap in graduating from secondary school without delay is considered.

These findings suggest that policies encouraging immigrant families to speak the native language at home are not effective in reducing the ethnic gap in school attainment. The fact that we could not discriminate between those families who speak Dutch at home among other languages from those that just speak the native language could be an explanation. In addition, we could not control for the quality of the spoken language. It may well be that speaking the native language matters only if the communication partners have native speaker proficiency. Other researchers finding some positive evidence of language indeed included some measure of proficiency as control variable (van Ours and Veenman, 2003; Dustmann et al., 2010).

Column E of Table 2.2 reports more, but still not very strong, evidence that speaking Dutch at home enhances the likelihood of transiting from school to work. It explains as much as six (= 16%) points of the 37 points predicted gap for the low educated, but only three points (8%) for the high educated who left school (before age 23). The finding that usage of native language at home helps more the low than high educated immigrant youth in finding a job is in line with Aldashev et al. (2009) and is consistent with the hypothesis that language usage is helpful in basic communications as required in low skilled job, but that it is no guarantee for proficiency, which is essential for high skilled jobs.

2.5.4 Robustness Checks

We performed several robustness checks to test the sensitivity of the results. First, we investigated the alternative decomposition strategy evaluating the pure ethnic gap at the native covariate registrations and thereby focusing on the higher end of the socioeconomic scale. In general, the results, as presented in Column D of Table 2.2, are in line with our main results. However, the pure ethnic gap in the probability of enrolling in higher education is now significant. This is because the pure immigrant advantage in enrolling in higher education conditional on secondary school completion disappears in this case: in comparison with Column C the pure ethnic gap for this outcome is 0.02 and insignificant instead of -0.07 and significant at the 10% level. This suggests that the immigrant advantage in terms of unobservables, as argued in Section 2.5.2, disappears if parents of immigrants acquire similar levels of educational attainment as parents from natives. The obtained pure ethnic gap concerning the probability of being employed three months after leaving school is somewhat lower following this alternative decomposition strategy: 0.12 and insignificant for the low educated and 0.20 and significant at the 5% level for the high educated. Immigrants with high socioeconomic background may be less penalised in terms of labour market networks relative to natives than immigrants with a low socioeconomic background.

Second, we re-estimated our model replacing our indicator of labour market success by an alternative one, that is being employed with a permanent contract two years after leaving school. Since at age 23 relatively few individuals have left education for two years or more, we had to use the data gathered in the follow-up interviews at ages 26 and 29. This means that these results are subject to sample attrition, especially for the higher educated group, since this group is the most likely to have left school less than two years ago at age 23. We report the goodness-of-fit and the decomposition results of this model in Table 2.6 in the Appendix of this chapter. Even if the predictions for the alternative labour market outcome deviate more from the observed ones than those in the benchmark model and are less stable between the different considered educational levels, the fit is satisfactory in that all the observed ethnic gaps lie in the 95% confidence interval of the simulation.

For the low educated both the total observed (total simulated) and the pure ethnic gap in being employed with a permanent contract two years after leaving school are much higher than in the benchmark model. The log odds ratio is respectively 0.74 (0.63) and 0.61 compared to 0.36 (0.37) and 0.28 in the benchmark. We therefore conclude that for the low educated the observed family endowments and prior school attainment seem to explain little of the ethnic gap in this alternative indicator of labour market success. The residual pure ethnic gap increases even substantially compared to the benchmark.

For the high educated the findings are less stable. This may be related to the attrition problem for this group, so more care should be taken in the conclusions. If we consider the labour market outcome of highest educated group, we observe that both the observed (simulated) and pure ethnic gap are close to zero: respectively -0.09 (0.03) and 0.02. However, if we consider the highest educated group but one, which was one of the main considered outcomes (denoted in bold), then the findings are much closer to the benchmark results. The aforementioned log odds ratios are then 0.27 (0.28) and 0.27 compared to 0.35 (0.37) and 0.34

in the benchmark. We conclude that the finding of the benchmark model that the pure ethnic gap of the labour market outcome is substantial is relatively robust, except for the highest level of education. However, due to drop out selectivity, the sensitivity analysis is less reliable for the latter group.

Third, we narrowed down our immigrant population definition. First, we restricted the immigrant population to the immigrant respondents with a Belgian nationality. In contrast to Euwals et al. (2010) who find a significantly positive relation between citizenship on the one hand and employment, tenured employment and job prestige on the other hand in the Netherlands, we obtain, for both the total and the pure ethnic gaps, very similar simulation results to our benchmark. This can be a consequence of the quasi exclusion of first generation immigrants from our sample (see Section 2.3.2). Second, we restricted the immigrant population to the more homogeneous population of youth with a Turkish or Moroccan origin. Using this definition leads to slightly larger predicted and simulated pure ethnic gaps. However, even if, as a consequence of the smaller sample sizes, the estimates are less precise, the empirical pattern remains very similar. The simulation results for the latter two sensitivity analyses are available on request.

2.6 Conclusions

Recently, researchers, among whom Heckman (2011), have claimed that ethnic gaps are primarily due to a lack of skills and that, consequently, by closing the gaps in skills, reflecting gaps in observed family endowments, the racial gaps disappears. In this research we investigated whether this claim upholds for Belgian society. To that purpose we built a dynamic schooling progression model that includes the school-to-work transition as a labour market outcome and estimated this model, separately on natives and immigrants, on a random sample of three birth cohorts living in Flanders, the Dutch speaking region in the North of Belgium. We then used this model to decompose, free of dynamic selection bias, the observed gap in both educational attainment and successful school-to-work transitions between native and immigrant youth into a part that can be explained by observed family endowments and a part that is inherent to ethnicity, the so called "pure ethnic gap". In this analysis natives are contrasted to grandchildren of women of "non-Western" nationality. We essentially excluded first generation immigrants from the analysis since only grandchildren born in Belgium or immigrated prior to age three were retained. Consequently, since the educational attainment of second generation immigrants is found to be better than that of first generation immigrants, the gap in the educational attainment found in this research is to be considered as a lower bound for the gap of first generation immigrants.

We contributed to the literature in essentially four dimensions. First, we incorporated years of schooling delay in the measures of educational attainment that are usually considered, such as completing secondary education or enrolling in higher education. We argued that it is important to incorporate this dimension, since arriving on the labour market with delay is very costly. Moreover, we pointed out that in countries where schooling delays are not important other scholastic achievement could play a similar role. Second, based on the insight that schooling outcomes realise sequentially, we proposed a method that allows identifying the moments at which the pure ethnic gaps emerge most prominently and therefore offer a tool that helps targeting research that tries to understand the origins of these gaps. Third, we moved a step forward by disentangling the educational outcomes within a particular grade. We distinguished between the educational achievement (passing or failing) realised at the end of each grade and the decision to continue schooling (rather than stopping) at the end of each school year. Finally, we innovated by including an indicator of whether the native language is spoken at home among the observed family endowments to investigate the role this factor plays in closing the observed ethnic gaps.

Our findings are the following. First, consistent with the existing literature, we find that observed family endowments alone explain the major part of the observed gap in secondary school completion and all of it with regards to the enrolment in higher education. This seems to suggest therefore that no specific policy for ethnic minorities is warranted to eliminate the existing schooling gaps in Belgium. However, once we take schooling delays into account this conclusion is no longer valid. The pure log odds ratio between natives and immigrants of the probability of completing secondary education without delay and of enrolling in higher education without delay is respectively 0.34 and 0.29.

Second, if we decompose the latter log odds ratios we find that the pure ethnic gap grows gradually throughout the educational progression. However, since retention is mainly used during secondary school, the major part of the pure ethnic gap in secondary school completion (0.19) emerges during this period. Fourth grade of secondary school has been identified as one of the key moments at which this gap originates: more than 20% of the total pure ethnic gap in the fraction that graduates from secondary school without delay is generated in this grade.

Third, in contrast to the finding of Belzil and Poinas (2010) for France, we find that family endowments and school attainment explain little of the ethnic gap in school-to-work transitions in Belgium. Independently of the level of education, a native school-leaver is about 30% more likely to be employed three months after leaving school compared with an immigrant school-leaver with the same observed endowments. For low educated school-leavers this conclusion is reinforced if an alternative labour market indicator that measures permanent employment two years after graduation is used. In that case the native school-leaver is even 80%[27] more likely to be employed in such a contract.

Finally, we find that speaking Dutch at home plays hardly any role in closing the ethnic gap in educational attainment, but that it does matter to some extent in explaining the different rate of transition from school to work between natives and immigrants, especially for the low educated. The fact that language is found to be less important than in other studies may be a consequence of our measure not capturing the proficiency of the native language sufficiently precisely. At the same time these findings are valuable in that they demonstrate that policies encouraging immigrant families to speak the native language at home are not very effective in reducing gaps in schooling, and labour market outcomes, especially among the high educated. Ensuring a higher degree of proficiency may be more effective, but more difficult to achieve.

Based on our analysis we disagree with the earlier evidence that observed ethnic gaps in educational achievement could be eliminated by targeting policy to socially disadvantaged groups irrespectively of their ethnic origin. We found that important ethnic gaps unrelated to family background may remain important if finer measurements of educational outcomes, such as in this study by specifying whether an educational level is attained with delay or not, are used. Therefore we believe that policies aimed at specific ethnic groups are still warranted to eliminate the ethnic gap in educational achievement. In addition, even if alternative explanations are possible, we believe that discrimination is a major candidate in explaining the important pure ethnic gap that we found in the transition from school to work in Belgium.

[27]This corresponds to a log odds ratio of 0.61.

2.7 Appendix: Additional Tables

Table 2.3: Unemployment Rate (UR) in Flanders for Males Aged 15 to 64

1982	**1983**	**1984**	**1985**	**1986**	**1987**	**1988**	**1989**
6.9%	6.8%	6.8%	5.7%	5.3%	5.1%	4.3%	2.8%
1990	**1991**	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**
2.3%	2.6%	2.6%	3.8%	4.8%	4.0%	4.1%	3.8%
1998	**1999**	**2000**	**2001**	**2002**	**2003**		
4.4%	4.4%	3.2%	3.6%	4.3%	5.2%		

Source: Steunpunt WSE of the Flemish government (based on Labor Force Study: Unemployment rates by sex).

Table 2.4: Unemployment Rate (UR) in Flanders for the Aged 15 to 24

1982	**1983**	**1984**	**1985**	**1986**	**1987**	**1988**	**1989**
18.9%	21.1%	23.4%	20.1%	16.5%	16.3%	12.0%	9.7%
1990	**1991**	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**
8.8%	8.7%	7.5%	11.7%	13.8%	12.5%	11.6%	11.7%
1998	**1999**	**2000**	**2001**	**2002**	**2003**		
11.0%	13.5%	11.3%	10.0%	11.6%	15.5%		

Source: Steunpunt WSE of the Flemish government (based on Labor Force Study: Unemployment rates by sex).

Table 2.5: Model Selection (Information Criteria Values)

K	**Flemish youth**				**Immigrant youth**			
	# parameters	**Log(L)**	**AIC**	**BIC**	**# parameters**	**Log(L)**	**AIC**	**BIC**
1	77	-40186.56	80527.12	80709.60	75	-2298.68	4747.37	4863.19
2	91	-39915.85	80013.70	80229.35	89	-2267.17	4712.34	4849.78
3	93	-39763.65	79713.31	79933.69	91	-2262.74	4707.49	4848.01
4	95	-39932.11	80054.23	80279.35	93	-2260.09	4706.18	4849.80
5	97	-39727.88	79649.76	79879.62	95	-2245.96	4681.92	4828.62
6	99	-39727.07	79652.14	79886.75	97	-2245.96	4685.92	4835.71
7	101	-39726.57	79655.15	79894.49	99	-2245.96	4689.92	4842.80
8	103	-39726.01	79658.02	79902.11	101	-2245.96	4693.92	4849.88

AIC: Akaike Information Criterion. BIC: Bayesian Information Criterion.

Table 2.6: Simulation Results (Model with Alternative Labour Market Outcome)

	A. Observed gap	B. Predicted gap		C. Pure ethnic gap by equating observed endowments to immigrant level		D. Pure ethnic gap by equating observed endowments to native level		E. Observed endowment gap: language	
Passing SE6	**0.20**	**0.21**	**[0.17, 0.30]**	**0.06****	**[0.02, 0.16]**	**0.07*****	**[0.04, 0.17]**	**0.00**	**[-0.03, 0.04]**
Passing SE4	0.05	0.03	[0.02, 0.06]	0.01	[-0.00, 0.04]	0.01***	[0.00, 0.03]	-0.00	[-0.01, 0.00]
Starting SE5 | passing SE4	0.04	0.03	[0.02, 0.07]	0.02**	[0.00, 0.05]	0.01***	[0.00, 0.03]	0.00	[-0.00, -0.01]
Passing SE5 | starting SE5	0.03	0.04	[0.02, 0.06]	0.00	[-0.02, 0.03]	0.01*	[-0.00, 0.03]	-0.00	[-0.01, 0.00]
Starting SE6 | passing SE5	0.04	0.05	[0.03, 0.10]	0.00	[-0.02, 0.06]	0.01*	[-0.00, 0.05]	0.01	[-0.00, 0.02]
Passing SE6 | starting SE6	0.04	0.06	[0.03, 0.10]	0.03*	[-0.00, 0.07]	0.03	[-0.02, 0.11]	-0.00	[-0.01, 0.00]
Starting TE	**0.37**	**0.34**	**[0.28, 0.47]**	**-0.01**	**[-0.09, 0.13]**	**0.10****	**[0.02, 0.27]**	**0.03**	**[-0.03, 0.08]**
Passing SE6	0.20	0.21	[0.15, 0.30]	0.06**	[0.02, 0.16]	0.07***	[0.04, 0.17]	0.00	[-0.03, 0.04]
Starting TE | passing SE6	0.16	0.13	[0.08, 0.21]	-0.07*	[-0.13, 0.01]	0.02	[-0.02, 0.11]	0.02 *	[-0.00, 0.05]
Passing SE6 without D	**0.60**	**0.62**	**[0.52, 0.81]**	**0.34*****	**[0.23, 0.54]**	**0.32*****	**[0.20, 0.59]**	**-0.04**	**[-0.11, 0.03]**
Starting PE without D	0.04	0.04	[0.03, 0.16]	0.03***	[0.01, 0.15]	0.00	[-0.01, 0.34]	0.01	[-0.01, 0.05]
Starting SE1 without D | starting PE without D	0.24	0.23	[0.14, 0.31]	0.12**	[0.02, 0.21]	0.11	[-0.09, 0.24]	-0.01	[-0.05, 0.03]
Starting SE2 without D | starting SE1 without D	0.06	0.05	[0.02, 0.10]	0.04***	[0.01, 0.08]	0.03***	[0.01, 0.08]	-0.01**	[-0.01, -0.00]
Starting SE3 without D | starting SE2 without D	0.02	0.03	[0.00, 0.07]	0.01	[-0.02, 0.05]	0.01	[-0.01, 0.05]	-0.01**	[-0.01, -0.00]
Starting SE4 without D | starting SE3 without D	0.10	0.07	[0.05, 0.17]	0.04**	[0.00, 0.10]	0.04**	[0.01, 0.10]	-0.01**	[-0.02, -0.00]
Starting SE5 without D | starting SE4 without D	0.06	0.10	[0.05, 0.17]	0.07***	[0.02, 0.14]	0.07***	[0.02, 0.14]	-0.01	[-0.02, 0.00]
Starting SE6 without D | starting SE5 without D	0.06	0.05	[0.01, 0.11]	0.00	[-0.04, 0.06]	0.02	[-0.02, 0.09]	0.00	[-0.02, 0.01]
Passing SE6 without D | starting SE6 without D	0.03	0.05	[0.02, 0.10]	0.03	[-0.01, 0.08]	0.04**	[0.00, 0.09]	-0.01**	[-0.01, -0.00]
Starting TE without D	**0.78**	**0.73**	**[0.62, 0.96]**	**0.29*****	**[0.17, 0.53]**	**0.36*****	**[0.21, 0.67]**	**-0.02**	**[-0.09, 0.06]**
Passing SE6 without D	0.60	0.62	[0.48, 0.81]	0.34***	[0.23, 0.54]	0.32***	[0.20, 0.59]	-0.04	[-0.11, 0.03]
Starting TE | passing SE6 without D	0.18	0.11	[0.06, 0.20]	-0.05	[-0.11, 0.04]	0.03	[-0.01, 0.12]	0.03*	[-0.00, 0.05]
PC 2y. after LS | LS before passing SE6	0.76	0.90	[0.54, 1.31]	0.88***	[0.49, 1.28]	1.04***	[0.54, 1.52]	0.13**	[0.02, 0.21]
PC 2y. after LS | LS with only SE6 d.	0.65	0.45	[0.25, 0.64]	0.43***	[0.22, 0.62]	0.57***	[0.23, 0.90]	0.10**	[0.02, 0.15]
PC 2y. after LS | LS with only SE6 d. or lower	**0.74**	**0.63**	**[0.39, 0.86]**	**0.61*****	**[0.35, 0.85]**	**0.70*****	**[0.34, 1.09]**	**0.11****	**[0.02, 0.18]**
PC 2y. after LS | LS with TE1 - TE4 d. and $\leqslant$ 1 year of D	**0.27**	**0.28**	**[0.08, 0.53]**	**0.27*****	**[0.06, 0.51]**	**0.42*****	**[0.09, 0.83]**	**0.05**	**[-0.01, 0.10]**
PC 2y. after LS | LS with TE3 or TE4 d. and $\leqslant$ 1 year of D	-0.09	0.03	[-0.16, 0.29]	0.02	[-0.17, 0.29]	0.15	[-0.13, 0.55]	0.05	[-0.01, 0.12]

The described gaps are log odds ratios of the observed and simulated fractions of the native respectively the immigrant youth. Some abbreviations are used: PE (primary education), SE1 (1st grade of secondary education), SE2 (2nd grade of secondary education), SE3 (3rd grade of secondary education), SE4 (4th grade of secondary education), SE5 (5th grade of secondary education), SE6 (6th grade of secondary education), TE (tertiary education), TE1 (1th grade of tertiary education), TE2 (2th grade of tertiary education), TE3 (3th grade of tertiary education), TE4 (4th grade of tertiary education), D (delay), LS (leaving school), y. (years), PC (permanent contract) and d. (degree). 95% confidence intervals, in brackets, were calculated using 999 random draws from the distributions of the underlying model's estimated parameters. In Column B ***(**)((*)) indicates significantly difference from the value in column A at the 1% (5%) ((10%)) level. In column C, column D and column E ***(**)((*)) indicates significantly difference from 0 at the 1% (5%) ((10%)) level.

2.8 Appendix: Steps in the Construction of the Simulated 95% Confidence Intervals of the Log Odds Ratios

The following steps are involved in the construction of the 95% confidence interval of the probability of any (log odds ratio of an) outcome of interest.

1. Randomly draw a vector of parameters from the asymptotic Normal distribution of all native respectively immigrant parameter estimates including the ones that refer to the endogenous variables (grade g and schooling delay V_t) and the unobserved heterogeneity distribution.

2. Consider the first estimated point of support of the unobservables and associate it to the R draws of the native and immigrant sample, so to $\mathbf{Z_N}$ respectively $\mathbf{Z_I}$.

3. Consider for all R vectors of observed and unobserved variables all possible paths that lead to the outcome of interest and calculate the chain of probabilities associated to each of these paths using the drawn parameter estimates as to predict the probabilities that these paths are realised.

4. Sum over all the possible paths to predict the R probabilities of realisation of the outcome of interest for the R draws of observed and unobserved explanatory variables.

5. Consider the draw of the next estimated point of support and repeat the steps 2 to 5 until the K estimated points of support are considered.

6. Calculate the weighted sum of the probabilities calculated in step 4 for each point of support where the weights correspond to the drawn estimated probability $p_k(\hat{\lambda})$.

7. Go to step 1. and repeat all subsequent steps until these steps are repeated $J = 999$ times.

8. The 95% confidence interval can be constructed by choosing the appropriate percentiles of the $J = 999$ simulated probabilities.

3

Do Employers Discriminate Less if Vacancies Are Difficult to Fill? Evidence From a Field Experiment

This chapter is joint work with Prof. Dr. Bart Cockx (Ghent University, Université catholique de Louvain, CESifo and IZA), Niels Gheyle (Ghent University) and Cora Vandamme (Argenta).

3.1 Introduction

It is well known that discrimination is not sustainable in a perfectly competitive product market (Becker, 1957). Similarly, discrimination is not possible in a perfectly competitive labour market (see, e.g., Cahuc and Zylberberg, 2004). Employers paying discriminated workers a lower wage than marginal productivity are driven out of the market by free entry, since employers without a preference for discrimination are willing to offer to these workers wages that do equal marginal productivity. However, recent contributions to the literature (see, e.g., Manning, 2003) have shown that employers, even if they operate in labour markets composed of many competing firms, can exercise a certain degree of monopsony power and can therefore discriminate against certain groups of workers without being driven out of the market. Monopsony power raises with search costs of employees and falls with search costs of employers. On the one hand, search costs incurred by employees limit their capacity to change employer and hence confer some power to employers to discriminate. On the other hand search costs at the employer side increase foregone output during the period that vacancies remain unfilled if a minority candidate is turned away. The primary objective of this study is to verify whether this second prediction holds: Do employers discriminate less if they have difficulties in filling their vacancies?

Contrary to the relationship between competition on the product market and discrimination,[1] the relationship between labor market tightness and discrimination has received little attention in the economic literature. Biddle and Hamermesh (2012) refer to Ashenfelter (1970) and Freeman (1973) arguing that "the perceived costs to employers of discriminating was higher in tight labor markets", but add that "neither found empirical evidence of cyclical movements in pure wage discrimination in the aggregate data." In addition, Booth et al. (2012) indicate that the heterogeneity by city in discrimination found in Australia is potentially partly driven by differences in labour market tightness. Apart from these authors, hardly any discussion of this relationship can be found in the literature. Biddle and Hamermesh (2012) are a rare exception in investigating this relationship, albeit *indirectly*, by studying how wage discrimination evolves over the business cycle. Building on the works of Black (1995) and Rosén (2003), they develop a theoretical equilibrium search model to get a better understanding of the underlying mechanisms of the cyclical variation in wage discrimination. At the same time this model forms the theoretical basis for our empirical analysis, since it confirms the aforementioned intuition that employers discriminate less if they face a tight labour market.[2]

In this study we are, to the best of our knowledge, the first to *directly* assess the relationship between labour market tightness and ethnic discrimination in the hiring process. To this end we conducted a correspondence test in Flanders, the Northern and economically most prosperous region of Belgium.[3] We sent out 752 fictitious job applications of school-leavers,

[1]See, e.g., Ashenfelter and Hannan (1986), Peoples and Saunders (1993), Black and Strahan (2001), Hellerstein et al. (2002), Black and Brainerd (2004) and, more recently, based on correspondence testing, Berson (2012).

[2]Biddle and Hamermesh (2012) state this result only in words, but it can be formally found by differentiating their Equation (9) with respect to φ: $\frac{\partial c^*}{\partial \varphi} = \frac{(1-\beta)\lambda}{r+s+(1-\beta)\varphi\lambda}[rU_A - c^* - rU_B] < 0$, where the negative sign follows from the fact that the term between braces on the right hand-side of (9) is a weighted average of $k + x$ and rU_A and from the fact that $k + x > rU_A$, so that $c^* > rU_A - rU_B$ or, equivalently, $rU_A - c^* - rU_B < 0$. Since φ is the rate at which workers arrive at employers, this rate decreases with labour market tightness and, hence, c^* increases with tightness and, since c^* is inversely related to discrimination, discrimination falls. Note that in this differentiation we hold U_A and U_B constant. This is because in the field experiment that we consider in our empirical analysis the labour market tightness for job seekers is given. They can apply for vacancies irrespectively of whether these are difficult to fill or not.

[3]Belgium is a federal state divided in three regions: Flanders, Wallonia and Brussels. In Flanders the

randomly assigned to individuals with a Flemish and a Turkish sounding name, to 376 vacancies for jobs requiring no work experience. Classifying these jobs on two measures of labour market tightness, autonomously constructed by the Public Employment Service of Flanders, we verify to what extent our measure of discrimination, the differential callback rate, differs between types of jobs. We perform sensitivity analysis to rule out that the found relationship just reflects correlation with other determinants of discrimination related to labour market tightness.

Our results confirm the negative relationship between labour market discrimination and labour market tightness. We find no significantly unequal treatment between the Flemish and Turkish job candidates in our experimental dataset when they apply for bottleneck occupations, that are occupations for which vacancies take long to fill. In contrast, Turkish job seekers applying for non-bottleneck occupations have to send out twice as many job applications in order to get the same number of job interviews as their Flemish counterparts.

Readers may take an interest in this study for a number of additional reasons. First, we focus on ethnic discrimination of school-leavers. Discrimination of this group is particularly relevant since discrimination at the first stage of the career may cause, through scarring (Arulampalam, 2001; Gregg, 2001; Gregg and Tominey, 2005), long-term adverse labour market outcomes even if discrimination does not play a role at later stages of the career.

Second, we provide evidence on hiring discrimination in the Flemish labour market. Flanders, and by extension Belgium, is an interesting case for a couple of reasons. In the 1990's the International Labour Office (ILO) conducted a series of ethnic discrimination studies in the three Belgian regions on the basis of audit and correspondence tests. Discrimination was found to be a significant and, compared with other OECD countries, more pronounced impediment to the employment of foreigners in Belgium (Arrijn et al., 1998). However, OECD (2008) argues that the results of the ILO studies probably had a stronger policy impact in Belgium than elsewhere. Affirmative action in combination with a stricter anti-discrimination legislation introduced in 2007 should have diminished labour market discrimination. Together with the very recent studies of Capéau et al. (2012b) and Capéau et al. (2012a)[4] our findings raise doubts on this conjecture.

This chapter is structured in the following way. In the next section we outline our experimental design. Subsequently we present a statistical analysis of the resulting dataset. A final

official language is Dutch, in Wallonia French and in Brussels both languages.

[4]Capéau et al. (2012b) and Capéau et al. (2012a) tested for the presence of discrimination in several dimensions in the three regions of Belgium: sex, age, ethnicity and nationality, pregnancy, and physical handicap. Their findings are, however, somewhat difficult to compare with the existing literature, since, in contrast to this literature, they compare callbacks between individuals who differ in more than one dimension at a time. We refer to their papers for further discussion.

section concludes and provides a brief discussion.

3.2 Experimental Design

3.2.1 Detecting Ethnic Discrimination by a Correspondence Test

Correspondence experiments to test for discrimination in the labour market have been extensively used (and refined) during the last decade. These experiments consist of sending carefully matched pairs of fictitious written job applications, randomly assigned to individuals revealing their minority status by their name or another individual characteristic, to real job openings and monitoring the subsequent callback. Concerning the identification of ethnic discrimination the extensive correspondence test conducted by Bertrand and Mullainathan (2004) is seminal. These authors show that, in the US labour market at the start of the former decade, applications with white-sounding names received 50% more positive callback on their job applications than those with African-American-sounding names. In Europe, pervasive levels of ethnic labour market discrimination are found in Greece, Ireland, Sweden and the UK (Drydakis and Vlassis, 2010; McGinnity et al., 2009; Bursell, 2007; Carlsson and Rooth, 2007; Wood et al., 2009). Besides, recent correspondence studies conclude that there is evidence of varying degrees of hiring discrimination based upon, for example, (i) gender in Austria, France and Spain, (ii) beauty in Sweden and (iii) sexual orientation in Austria, Greece and Sweden (Weichselbaumer, 2004; Petit, 2007; Albert et al., 2011; Rooth, 2007; Weichselbaumer, 2003; Drydakis, 2009; Ahmed et al., 2011). Furthermore, the correspondence methodology has also been applied to identify discrimination in other markets (an example is Carlsson and Eriksson (2012), in the Swedish housing market).

These field experiments have been widely viewed as providing the most convincing evidence on discrimination (Pager, 2007; Riach and Rich, 2002). Researchers using non-experimental data possess far less information than employers do. Native and foreign employees who according to these data appear similar to researchers may therefore be very different from the employers' perspective. By conducting a correspondence test, selection on individual unobservable characteristics is not an issue since all the employers' decision making information is controlled for by the researcher. Thereby strict equivalence between candidates is ensured. Moreover, this approach allows disentangling employer discrimination from alternative explanations of differential hiring rates between migrants and natives, such as differential employee preferences and network effects.

3.2.2 Construction of Applications and Matching with Vacancies

We generated template CVs and cover letters for eight profiles of school-leavers. These different profiles allow us to apply for vacancies with different requirements both in terms of schooling level and specialisation. First, three middle educated profiles with a secondary education diploma (ISCED[5] 3) in commerce, metallurgy and organisation help. Second, five high educated profiles holding a professional bachelor in business administration (ISCED 5) with a different specialisation (accounting and tax, finance and insurance, logistics, marketing and legal practice).[6]

All profiles were single males with the Belgian nationality graduated in June 2012. Depending on the region of the announced workplace in the vacancy, their residence was located in one of the suburbs of Antwerp or Ghent, the two largest cities of Flanders. Middle educated school-leavers were 18 years old and high educated school-leavers were 21 years old. So, none of the candidates experienced a grade retention in the past. In addition we added to each application the following features: Dutch mother tongue,[7] adequate French and English language skills, driving license, computer skills and student employment experience. Moreover, the cover letters signalled a motivated, structured and capable person. For the high educated school-leavers also sport club membership and student leadership were added. Last, we added a fictitious postal address (based on real streets in middle-class neighbourhoods) and date of birth to the applications. The CV and cover letters are available on request.

During five months, from November 2011 until March 2012, we randomly selected vacancies from the database of the Flemish Public Employment Service (PES or "VDAB" in Dutch), the major job search channel in Flanders, for which (at least) one of our eight profiles was adequately educated. We restricted ourselves to vacancies for which no work experience was required and which were posted less than a fortnight before the start of the experiment.[8]

The ethnicity of the candidate was only signalled by the name. Turkish names were used because the Turkish community forms the most significant ethnic minority in Ghent and the second most important in Antwerp. In addition, the unemployment rate for residents of non-EU-15 countries (among which Turkey) is very high. In 2011 23% of the active non-EU-15 residents were unemployed in Belgium, compared to 6% of the active Belgians.[9] Finally,

[5]ISCED stands for International Standard Classification of Education.

[6]This degree is among the highest that migrants obtain in Flanders (Duquet et al., 2006).

[7]Thereby, we isolate the effect of ethnicity from potential language effects.

[8]This choice was made in order to maximise the callback rate, since interviews with human resources managers revealed that filled vacancies are not always immediately removed from the PES database.

[9]Source: Eurostat. A study of the PES shows that the unemployment rate of individuals with a Turkish origin traditionally lies above the non-EU-15 average in Flanders (VDAB, 2009a).

typical Flemish and Turkish names can be easily distinguished.[10]

For each of the eight aforementioned profiles of school-leavers we created two types of CVs and cover letters: "Type A" and "Type B". This allowed us to send two applications, one of each type and of each ethnic group, to the same vacancy. To maximise comparability, both application types were identical in all job-relevant characteristics, such as number of months of work experience in student work,[11] language skills and quality of extra-curricular engagements. Type A and Type B candidates obtained education in the same type of school, with a comparable reputation. The applications just differed in inessential details, such as the name of the school, favourite sports and other particular engagements, and in fonts and lay-out.[12] In order to completely erase any dependence of call backs on the application type Flemish and a Turkish sounding name were alternately assigned to the Type A and Type B versions and, subsequently, sent in an alternating order to vacancies, each time with a one-day delay in between.

We matched to each assigned name an email address and a mobile phone number. These were registered with large commonly used internet and telecommunication providers. We logged for each application sent the number of announced (similar) job positions in the vacancy, the address of the workplace, the gender of the recruiter (if available), the date of the application, the application profile (one of the five high educated or one of the three middle educated profiles) and the application type (A or B).

3.2.3 Measurement of Callback

All applications were sent to the employer by email. Callbacks for interviews were received by telephone voice mail or by email. The content of the responses are available on request. Since we included postal addresses with a nonexistent street number in the applications, callback via regular mail could not be measured. However, several human resource managers confirmed that employers rarely, if ever, invite job candidates by regular mail for selection interviews. To minimise inconvenience to the employers, invitations were immediately declined. All callback later than 40 days after sending the application was neglected. This, however, turned out to

[10]Based on frequency data on first names and surnames we chose "Thomas Mertens" and "Jonas Vermeulen" as Flemish sounding names and "Emre Sahin" and "Okan Demir" as Turkish sounding names. We checked that these names were no stereotypes. Assigning different pairs of names to the middle and high educated individuals allowed to let both categories of individuals apply for vacancies of the same employer without risking detection.

[11]Note that restricting the analysis to school-leavers has an advantage from a methodological point of view. Controlling for human capital is easier for them, since we need not take labour market experience (beyond student work) into account.

[12]To be as realistic as possible, we adapted templates that the PES posts on its website as examples for job seekers.

be an artificial restriction since no response was received after 40 days.

In our analysis we distinguish between two definitions of positive callback. Positive callback sensu stricto means that the candidate is invited for an interview related to the job for which he applied. This definition is mostly used in the literature and therefore our benchmark definition. Positive callback sensu lato includes in addition to the sensu stricto definition also the receipt of an alternative job proposal and the request to provide more information or to contact the recruiter.

3.2.4 Variation in Labour Market Tightness

We matched each vacancy one-to-one with an occupation in the classification list of the PES.[13] For each occupation the PES provided us with two autonomously constructed measures of labour market tightness in 2011. First, the median duration to fill a vacancy in this occupation. This duration is right censored at vacancy withdrawal. Second, the so called "bottleneck" status of the occupation. Each year a list of bottleneck occupations is published by the PES. This list is obtained combining three statistical criteria and is then assessed by a number of labour market specialists. These three criteria are that (i) there must be at least 10 vacancies for the concerned occupation in the PES database, (ii) the vacancy filling rate must be lower than the median filling rate for all occupations together, and (iii) the median duration until a vacancy in this occupation is filled must be greater than the median for all occupations together. According to VDAB (2009b), the bottleneck status is driven by the relative size of the pool of adequately skilled workers, the wage level and the working conditions in these occupations. In the benchmark empirical analysis we rely on this second measure. The first measure is used in a sensitivity analysis as a robustness check.

Table 3.10 in the Appendix of this chapter lists the classifications of the occupations, some variables characterising these occupations and the number of fictitious applications that were sent to each of these occupations. First, both PES measures of labour market tightness for these occupations in 2011 are reported. The occupations with the minimum and maximum median vacancy duration in our experimental dataset are consultant in recruitment and selection (13 days) and demonstrator (109 days). "Bottleneck" occupations are industrial cleaner, classic cleaner, private cleaner, customs declaration officer, executive expedition operator, planning and logistics clerk, shipping agent at the quay, bookkeeper, accountant, seller, representative, call center employee and tele-seller. Second, the table contains two indicators of customer contact in the occupations, which will be used in the sensitivity analysis. Third,

[13] This occupation classification is a classification at 5-digit level. The PES classifies occupations in bottleneck and non-bottleneck occupations at this level.

it reports the number of observations (twice the number of vacancies) for each of the occupations by level of education. For three occupations (administrative clerk, commercial clerk and representative) applications were sent out for both middle and high educated profiles, depending on the particular requirements in the vacancy.

3.2.5 Research Limitations

In short we assess some research limitations inherent to our experimental design. For an in-depth discussion of the strengths and weaknesses of correspondents tests in general we refer to Riach and Rich (2002) and Pager (2007) and for an elaboration on the ethical aspects of this kind of tests to Riach and Rich (2004).

First, our experimental design can only demonstrate discrimination, if any, at the initial stage of the selection process. Since we simply measure callback rates for first interviews, we cannot make any statements about discrimination in the later stages of the selection process, let alone in wages. However, Bertrand and Mullainathan (2004) argue that reduced interview rates are expected to be reflected in reduced job offers and lower earnings. Moreover, since job interviews are costly, firms invite candidates for an interview only if they have a reasonably high chance of getting the job.

Second, we only investigate discrimination for a selection of occupations and for vacancies posted at the PES database. Possibly, discrimination is more or less pervasive in other sectors than those that are covered by the database and among employers who rely on other channels (for instance social networks) for filling their vacancies. It is unclear whether these limitations, taken together, may lead to an overestimation or rather an underestimation of discrimination in the Flemish youth labour market. However, it is important to keep in mind that we are especially interested in the relationship between discrimination and labour market tightness. If, therefore, the limitations mentioned cause a similar shift in the discrimination measures for the bottleneck and for the non-bottleneck occupations, our main research conclusions remain valid.

Last, as demonstrated by Heckman (1998), our design does not allow to distinguish between taste-based discrimination on the one hand and statistical discrimination on the other hand. Kaas and Manger (2012) and Carlsson and Rooth (2008) show how, to some extent, these forms of discrimination can be disentangled within the correspondence test framework. However, disentangling these forms of discrimination is outside the scope of this study.

3.3 Results

3.3.1 Descriptive Analysis

In this section we follow the international literature by reporting descriptive statistics on unequal treatment of Flemish and Turkish job candidates and on the relative callback probability of these groups.

Table 3.1 presents the aggregate experimental results adopting the sensu stricto definition of positive callback. Table 3.7 (in the Appendix of this chapter) displays the same statistics using the sensu lato definition. Since two applications were sent to each vacancy there are four possible outcomes: (i) positive callback for neither candidate, (ii) positive callback for both candidates, (iii) only positive callback for the Flemish candidate and (iv) only positive callback for the Turkish candidate. Overall, in 79 (139) of the 372 vacancies at least one candidate received positive callback sensu stricto (sensu lato). 29 (45) cases resulted in a positive callback for just the Flemish candidate and 7 (15) for the Turkish candidate only. The net discrimination rate is calculated as the ratio of the difference between the number of vacancies in which the Flemish and, respectively, Turkish candidate was treated favourably, and the total number of vacancies in which at least one candidate received a positive callback. Overall the net discrimination rate is 0.28 (0.22) adopting the sensu stricto (sensu lato) definition of positive callback. A standard χ^2 test of the hypothesis that the candidates of both ethnicities were equally often treated unfavourably is rejected at the 1% level. Based on this statistic we conclude that there is evidence of discrimination against Turkish school-leavers in the Flemish labour market.

Table 3.1 and Table 3.7 in the Appendix of this chapter also show the same descriptive statistics after splitting up the data in vacancies for bottleneck and non-bottleneck occupations. For the remainder of this section, we will focus, unless stated otherwise, on the results for this split-up and for the sensu stricto definition of positive callback. Note, however, that the results based on the alternative definition go in the same direction across all presented statistics.

Table 3.1 indicates that the net discrimination rate varies with labour market tightness in the expected direction. It is hardly different from zero for bottleneck occupations. In sharp contrast, this statistic is 0.50 for non-bottleneck occupations: while for 22 of the 195 vacancies only the Flemish candidate received a positive callback, just one vacancy resulted in a positive response for the Turkish candidate only. The more competition employers face in attracting workers, the lower the discrimination rate, since discrimination is then too costly.

Table 3.1: Unequal Treatment of Flemish and Turkish Job Candidates (Sensu Stricto)

Occupations	**Jobs** (No.)	**Neither callback** (No.)	**Both callback** (No.)	**Only Flemish callback** (No.)	**Only Turkish callback** (No.)	**ND**	χ^2
All	376	297	43	29	7	0.28***	13.44
Bottleneck	181	144	24	7	6	0.03	0.08
Non-bottleneck	195	153	19	22	1	0.50***	19.17

Note. ND: net discrimination rate. The null hypothesis is that both individuals are equally often treated unfavourably. ***(**)((*)) indicates significance at the 1% (5%) ((10%)) level.

Table 3.2: Positive Callback Rates (Sensu Stricto) for Flemish and Turkish Job Candidates

Occupations	**Callback rate Flemish**	**Callback rate Turkish**	**Callback ratio**	**t**
All	0.19	0.13	1.43***	3.73
Bottleneck	0.17	0.17	1.03	0.28
Non-bottleneck	0.21	0.10	2.05***	4.59

Note. The null hypothesis is that the callback rate is equal for both ethnicities. Standard errors are corrected for clustering of the observations at the vacancy level. ***(**)((*)) indicates significance at the 1% (5%) ((10%)) level.

Table 3.2 presents callback rates by ethnicity. These confirm the findings based on the net discrimination rate. The callback rate is defined as the number of positive callbacks relative to the total number of sent applications. The callback ratio is obtained by dividing the Flemish callback rate by the Turkish callback rate. The callback ratio is only significantly different from 1 for the individuals who apply for a non-bottleneck occupation. Candidates with Turkish sounding names need to send out more than twice as many job applications to be invited to as many job interviews as the Flemish candidates.

A counterintuitive result in Table 3.2 is that the callback rate for the Flemish candidates is lower when they apply for bottleneck occupations than when they apply for non-bottleneck occupations. This finding seems to be largely driven by the 168 observations (84 vacancies) with as an occupation industrial, classic and private cleaner. Callback rates for these cleaning occupations are both for Flemish and for Turkish candidates very low, namely 0.09. This may be a consequence of employers preferring female candidates for these jobs and of our candidates being to some extent overqualified for these jobs. If we drop these 170 observations from the dataset the callback rate sensu stricto (sensu lato) for bottleneck occupations increases for the Flemish from 0.17 (0.32) to 0.24 (0.39) and for the Turks from 0.17 (0.31) to 0.23 (0.41). As expected, the coefficient of the indicator of bottleneck occupations becomes, in this case, positive but is still not statistically significant.

Table 3.3: Main Empirical Analysis. The Probability of Positive Callback: Probit Estimates, Average Partial Effects

Variables	**Positive callback**			
	Sensu stricto		**Sensu lato**	
Turkish name * Bottleneck occupation	−0.01	(0.02)	−0.01	(0.03)
Turkish name * Non-bottleneck occupation	−0.11***	(0.02)	−0.15***	(0.03)
Log-likelihood	-328.93		-446.47	
Observations	752		752	

Note. Other control variable: indicator of bottleneck occupations. The reported average partial effects are averages over the Turkish population. Standard errors, corrected for clustering at the vacancy level and calculated using the delta method, are in parentheses. ***(**)((*)) indicates significance at the 1% (5%) ((10%)) level.

3.3.2 Empirical Analysis

In this section ethnic differences in positive callback rates are estimated on the basis of various probit models with the callback indicator (following both the sensu stricto and sensu lato definitions) as the dependent variable. Since characteristics of applicants are by construction orthogonal to ethnicity, adding these characteristics or not to the probit model does not affect the estimates of our main coefficients of interest, the interaction effects with ethnicity. We therefore choose to leave these characteristics out of the analysis.

The statistics in Table 3.3 (and Table 3.9 in the Appendix of this chapter) square with those reported in Table 3.2 and Table 3.8. In our experimental dataset, overall, a Turkish sounding name lowers the probability of receiving an invitation for a job interview by 11 percentage points after applying for a non-bottleneck occupation, while for bottleneck occupations the callback rate is not significantly different between the Turks and the Flemish. Equality of the related estimation coefficients for the probit model is rejected at the 1% significance level.

We conducted an extensive number of robustness checks on the aforementioned results. In a first robustness check, we estimate the probit model with the alternative variable capturing labour market tightness, namely the median vacancy duration time for the occupation for which the individual candidates. We normalise this variable by subtracting the sample mean and dividing by the sample standard deviation. Table 3.4 shows that an increase of the median vacancy duration by one standard deviation, that is by about 17 days, lowers discrimination by four percentage points. This result confirms that labour market discrimination is lower for occupations with high labour market tightness. In addition, we also looked into the effect of the standard deviation of the vacancy duration time for the occupation as higher standard deviations might be related to higher uncertainty about the arrival rate of new adequate (native) candidates after sending away a minority candidate. This exercise, however, led to

Table 3.4: Sensitivity Analysis 1. The Probability of Positive Callback: Probit Estimates, Average Partial Effects

Variables	**Positive callback**			
	Sensu stricto		**Sensu lato**	
Turkish name	−0.06***	(0.02)	−0.08***	(0.02)
Turkish name * Norm. median vacancy duration	0.04***	(0.01)	0.04***	(0.02)
Log-likelihood	-327.46		-444.23	
Observations	752		752	

Note. Other control variable: normalised median vacancy duration. The reported average partial effects are averages over the Turkish population. Standard errors, corrected for clustering at the vacancy level and calculated using the delta method, are in parentheses. ***(**)((*)) indicates significance at the 1% (5%) ((10%)) level. The median vacancy duration time for the occupation is normalised by subtracting the sample mean and dividing by the sample standard deviation.

insignificant results.

A concern is that the coefficients of both measures of labour market tightness, the median vacancy duration and the bottleneck status, may be affected by a simultaneity bias. We cannot exclude that vacancy durations may be longer as consequence of discrimination. However, if this were the case, the finding of less discrimination for bottleneck occupations would be strengthened, since we find do not find a positive but a *negative* relationship between vacancy duration and discrimination.

Another concern is that the bottleneck status of a job may correlate with other determinants of discrimination, so that the observed correlation is not causal. In a second robustness check we therefore include additional interactions between Turkish origin and four potential determinants of discrimination that may be correlated with the bottleneck status of an occupation. First, one could expect that labour market tightness is higher for jobs that require more education. Moreover, both theoretical[14] and empirical evidence[15] show that discrimination decreases with the level of education, so that our findings on labour market tightness could just reflect this relationship. Therefore, we include an indicator that identifies the high educated candidates, in casu those holding a professional bachelor in business administration. In our sample, individuals are high-educated in 34% of the applications for bottleneck occupations and in 61% of the applications for non-bottleneck occupations. Second, since customer induced discrimination (Becker, 1957) is expected to be higher in occupations with intensive customer contact, we include an indicator of intensive customer contact (see Section 3.2.4). In our sample, intensive customer contact is a characteristic of 35% (15%) of the bottleneck

[14]Taubman and Wales (1974) argue that higher education can act as a prejudices reducing screening device. In addition, if the level of education is reflected in the value of the production ("x") one can use the model of Biddle and Hamermesh (2012) to show that discrimination decreases with the level of education: It is clear from their equation (9) that c^* increases, and hence discrimination decreases, with x. The reason is that the opportunity cost of an unfilled vacancy increases with x.

[15]See Bursell (2007), Carlsson and Rooth (2007) and Wood et al. (2009).

Table 3.5: Sensitivity Analysis 2. The Probability of Positive Callback: Probit Estimates, Average Partial Effects

Variables	**Positive callback**			
	Sensu stricto		**Sensu lato**	
Turkish name * Bottleneck occupation	−0.09*	(0.05)	−0.08**	(0.04)
Turkish name * Non-bottleneck occupation	−0.22***	(0.06)	−0.25***	(0.06)
Turkish name * High educated	0.14***	(0.05)	0.17***	(0.06)
Turkish name * Customer contact	−0.03	(0.04)	−0.01	(0.05)
Turkish name * Norm. % foreign workers in sector	0.01	(0.02)	0.01	(0.02)
Turkish name * Log(average wage in occupation)	−0.16*	(0.09)	−0.05	(0.11)
Log-likelihood	-315.40		-423.31	
Observations	736		736	

Note. Other control variables: indicator for bottleneck occupations, indicator for high educated candidates, indicator for occupations with intensive customer contact, normalised fraction of foreign workers in the sector of the firm, natural logarithm of the average wage in the occupation. The reported average partial effects are averages over the Turkish population. Standard errors, corrected for clustering at the vacancy level and calculated using the delta method, are in parentheses. ***(**)((*)) indicates significance at the 1% (5%) ((10%)) level. The percentage of foreign workers in the sector of the employer is normalised by subtracting the sample mean and dividing by the sample standard deviation. 16 observations are dropped since neither the name of the firm nor its sector is given in 8 vacancies posted by labour market intermediaries.

(non-bottleneck) jobs. Third, according to the social distance theory (Akerlof, 1997) hiring discrimination should fall with the fraction of foreign workers in the firm (sector). Even if there is only weak empirical evidence for this theoretical prediction (Carlsson and Rooth, 2007; Bursell, 2007; Wood et al., 2009), we try to capture this relationship by including a variable measuring the fraction of workers with a non-Western nationality in the sector of the firm as a proxy of the fraction of foreign workers in the firm itself.[16] This variable was constructed by first identifying the sector of the employer that posted the vacancy[17] and then by merging this information to the fraction of workers with a non-Western nationality in the corresponding sector (2-digit level) in Flanders on December 31, 2009.[18] In our sample, the fraction of foreign workers is 5% (2%) in the bottleneck (non-bottleneck) jobs. Fourth, we include the average wage level in the occupation in 2010 as the bottleneck status of occupations is partly driven by the wage level in these occupations (see Section 3.2.4).[19] Therefore, the results presented in Table 3.3 potentially only reflect that discrimination is higher in well paid occupations. In our sample, the average wage level in the occupation is 2864 euro in the bottleneck jobs and 2946 in the non-bottleneck jobs.

[16]To our knowledge, these data are not available at the firm level in Belgium. Note that this proxy is also imperfect in the sense that all candidates in our empirical setting have the Belgian nationality.

[17]We did this by linking, on the basis of the online database of the Flemish business periodical "Trends", the name of the employer to the sector.

[18]Source: Datawarehouse of the Belgian federal public service of social security.

[19]This average is not measured for the profession classification of the PES but for the ISCO-08 classification at 3-digit-level which is, however, closely related to the former classification. ISCO stands for International Standard Classification of Occupations.

Table 3.5 reports the results for this second robustness check. The coefficients for the interactions between Turkish origin and the two last mentioned variables have the expected sign but are not significant. In contrast, the regression results provide, as expected, evidence of significantly less discrimination against the high educated subsample of Turkish candidates. In addition, we find weakly significant evidence for more discrimination in better paid occupations. However, the inclusion of these additional interaction variables does not affect our main conclusion. On the contrary, the differential discrimination against Turkish candidates between bottleneck and non-bottleneck occupations becomes even slightly more pronounced. Equality of the related estimation coefficients for the probit model is rejected at the 1% significance level.

We also examined[20] a number of alternative specifications in which Turkish origin is interacted with (i) the indicators both of moderate and of intensive customer contact; (ii) the fraction of Turkish (instead of non-Western) workers in the sector; (iii) the size of the firm in terms of its average number of workers in 2010 and (iv) other employer (or vacancy) characteristics (which we did not expect to be correlated with the bottleneck status of the occupation), such as the number of announced (similar) job positions by the vacancy, the province of the workplace or the gender of the recruiter.[21] None of these alternatives modifies our main conclusions in any way. The same holds true if we differentiate the interaction between Turkish origin and bottleneck status by level of education.

Heckman and Siegelman (1993) show that not controlling for group differences in the variance of unobservable job-relevant characteristics (and thereby of unobservable determinants of positive callback) can lead to spurious evidence of discrimination. To see this more clearly, assume that both the average observed and unobserved determinants of productivity are the same for Flemish and Turkish candidates for an unfilled vacancy, but that the variance of unobservable job-relevant characteristics is higher for Flemish than for Turkish youth. In addition, suppose that the employer considers the observed determinants of productivity, as inferred from the CV and the motivation letter, are relatively low compared to the job requirement. In that case it is rational for the employer to invite the Flemish and not the Turkish candidate, since, as the variance of unobservable job relevant characteristics is higher for the Flemish than for the Turkish candidates, it is more likely that the sum of observed and unobserved productivity is higher for the Flemish candidates. A correspondence test that detects discrimination against Turks could therefore overestimate the extent of discrimination. However, with other assumptions the bias may be in the opposite direction.

[20]These findings are available upon request.

[21]We were not able to include an interaction with a dummy indicating recruiters from an ethnic minority since hardly any recruiter had a foreign sounding name.

Neumark (2012) explicitly addresses this critique and provides a statistical procedure in order to recover unbiased estimates of discrimination. In what follows, we succinctly describe Neumark's approach. Subsequently, in a third robustness analysis, we apply this method to check to what extent our conclusions are sensitive to this critique. To the best of our knowledge, we are the first to follow Neumark in applying this methodology.

It is well known that in a standard probit model only the ratio of the coefficients to the standard deviation of the unobserved residual is identified. In estimations the standard deviation is usually arbitrary set to one. In our case this means that the variance of unobservable job-relevant characteristics is implicitly assumed to be equal (to one) for both ethnic groups, which, for reasons stated above, may therefore bias the intensity of discrimination. Neumark (2012) shows, however, that if the researcher observes job-relevant characteristics that affect the native and migrant populations' propensities of call back in the same way, one can identify the ratio of the standard deviation of the unobserved productivity components of these groups. The intuition is that if in a standard probit the estimated coefficients of these job-relevant characteristics differ by ethnicity, then this must be a consequence of a differential standard deviation, since by assumption the coefficient of these characteristics should be the same across ethnic groups (and since, as mentioned before, in a probit model only the ratio of the coefficients to the standard deviation are identified). To implement this idea, this just boils down to the estimation of a heteroskedastic probit model in which the variance of the error term is allowed to vary with ethnicity.

To identify the heteroskedastic probit model we assume that (i) the distance between the living place of the candidate and the announced working place and (ii) the particular application profiles, *beyond* their education level (high or middle educated), influence the callback rates in a similar way for Flemish and Turkish candidates.[22] The hypothesis of equality of the coefficients concerning these variables for both ethnic groups cannot rejected on the basis of a likelihood ratio test (p-value 0.88 or 0.87 following the sensu stricto or sensu lato definition of positive callback).

Table 3.6 reports the estimation results. In line with Neumark (2012), we get a (non-significantly) higher estimated variance of the error term for the foreign candidates. The overall marginal effects of the interaction variables at interest are closely comparable to the effects outlined in Table 3.3. They, however, can be decomposed in two parts. First, the partial effect of the variables at interest, holding the variance constant. Second, the effect of the

[22] Note that candidates apply for job vacancies that require a level of education that matches the attained level. Moreover, as mentioned, the extent of discrimination is expected to decline with the level of education, so that the level of education cannot be used to identify the differential variance in the heteroskedastic probit model.

Table 3.6: Sensitivity Analysis 3. The Probability of Positive Callback: Heteroskedastic Probit Estimates, Partial Effects

Variables	**Positive callback**			
	Sensu stricto		**Sensu lato**	
Overall average partial effect				
Turkish name * Bottleneck occupation	−0.01	(0.02)	−0.01	(0.03)
Turkish name * Non-bottleneck occupation	−0.11***	(0.03)	−0.14***	(0.03)
Average partial effect through level				
Turkish name * Bottleneck occupation	−0.06	(0.07)	−0.04	(0.06)
Turkish name * Non-bottleneck occupation	−0.16***	(0.06)	−0.16***	(0.05)
Average partial effect through variance				
Turkish name * Bottleneck occupation	0.05	(0.05)	0.03	(0.04)
Turkish name * Non-bottleneck occupation	0.05	(0.04)	0.03	(0.06)
$\ln(\sigma_T/\sigma_F)$	0.25	(0.30)	0.17	(0.34)
Log-likelihood	-304.73		-419.68	
Observations	752		752	

Note. Other controls: indicator of high educational attainment interacted with indicator of Turkish name, indicator of bottleneck occupation, indicator of high educational attainment, normalised variable capturing the distance (in minutes by car) between the announced work place and the living place of the candidate and six indicators for the eight application profiles except one reference profile for both high and middle level of education. Standard errors, corrected for clustering at the vacancy level and calculated using 500 bootstrap replications, are in parentheses. ***(**)((*)) indicates significance at the 1% (5%) ((10%)) level. $\ln(\sigma_T/\sigma_F)$ stands for the natural logarithm of the ratio between the standard deviation of unobservables for the Turkish and the Flemish subpopulation.

variables at interest via their impact on the variances of the unobservables. By disentangling these components we obtain that the effects on the level of the latent variable are larger in magnitude than the partial effects in Table 3.3.[23] The effect on the callback chance sensu stricto (sensu lato) of a Turkish sounding name applying to a non-bottleneck occupation increases in absolute value changes from minus 11 (15) to minus 16 (16) percentage points. The corresponding discrimination in case of application to a bottleneck occupation changes from minus 1 (1) to minus 6 (4) percentage points, but remains insignificant. Clearly, discrimination is more severe (although not significantly so) than in the analysis that ignores the role of ethnic group differences in the variance of the error term. However, the differential discrimination rate between bottleneck and non-bottleneck occupations is hardly affected.

As a fourth robustness check, available upon request, we extend the benchmark model by including an interaction between Turkish origin and a monthly proxy for the labour market tightness at a macro level: the number of vacancies divided with the number of unemployed in Flanders in the month the job application was sent out. The estimated coefficient for this interaction variable has the expected positive sign, implying that discrimination is lower in

[23]In contrast to Neumark (2012) who approximates the effect of a discrete change in the variables of interest by a partial derivative, we explicitly take the discrete nature of these variables into account and measure these effects on the basis of discrete changes in the callback probability.

times of more labour market tightness at the macro level. However, probably because of the limited variation in this macro variable, this effect is not significant.

3.4 Conclusions

To the best of our knowledge, this study is the first to test the theoretical relationship between labour market discrimination and labour market tightness directly. If employers have difficulties in filling a vacancy, turning a minority worker away is extra costly in terms of forgone output, since the vacancy then risks to remain vacant for a long time. In the correspondence test that we conducted, applicants with a Turkish sounding name were no longer discriminated against if they applied for occupations for which labour market tightness was high. In contrast, if they applied for occupations for which there are plenty of candidates, they had to send twice as many applications than candidates of native origin to be invited to a job interview. These results were found to be robust to a number of sensitivity analyses.

From a policy point of view, these findings suggest that labour market discrimination can be reduced by appropriate economic incentives; by increasing its cost. If thereby monopsony power is reduced, intuitively, such policies need not come at an efficiency cost, but whether this is the case clearly depends on the source of monopsony power and the precise nature of the policy. Further theoretical analysis is required before we can formulate any clear policy advice on this point. Our results also suggest to advise minorities to apply for jobs that are difficult to fill. However, such a policy advice may only work to the extent the competencies of minorities match the requirements for these jobs and that the tightness on the labour market is partly a consequence of minorities not being informed about for which occupations employers have difficulties in filling vacancies.

A well known limitation of correspondence tests is that they can only detect discrimination in the first stage of the hiring process. It is not because we detect no discrimination for bottleneck occupations at this first stage, that employers do not discriminate at a further stage. For instance, a possible reason that employers find too few candidates for particular occupations is that they do not pay enough relative to the job requirements. If this would be the main reason why bottleneck occupations exist, wage discrimination could remain an issue, even if employers do not discriminate in the hiring process, since, if as a consequence, disproportionately more minority workers are hired in these occupations, they will earn on average less than equivalent non-minority workers. Further research is therefore required to investigate the importance of this issue.

3.5 Appendix: Additional Tables

Table 3.7: Unequal Treatment of Flemish and Turkish Job Candidates (Sensu Lato)

Occupations	Jobs (No.)	Neither callback (No.)	Both callback (No.)	Only Flemish callback (No.)	Only Turkish callback (No.)	ND	χ^2
All	376	237	79	45	15	0.22***	15.00
Bottleneck	181	111	44	14	12	0.03	0.15
Non-bottleneck	195	126	35	31	3	0.41***	23.06

Note. ND: net discrimination rate. The null hypothesis is that both individuals are treated unfavourable equally often. ***(**)((*)) indicates significance at the 1% (5%) ((10%)) level.

Table 3.8: Positive Callback Rates (Sensu Lato) for Flemish and Turkish Job Candidates

Occupations	Callback rate Flemish	Callback rate Turkish	Callback ratio	t
All	0.33	0.25	1.32***	3.94
Bottleneck	0.32	0.31	1.04	0.39
Non-bottleneck	0.34	0.19	1.74***	5.09

Note. The null hypothesis is that the callback rate is equal for both ethnicities. Standard errors are corrected for clustering of the observations at the vacancy level. ***(**)((*)) indicates significance at the 1% (5%) ((10%)) level.

Table 3.9: The Probability of Positive Callback for an Interview: Probit Estimates, Average Partial Effects

Variables	**Positive callback**	
	Sensu stricto	**Sensu lato**
Turkish name	-0.06*** (0.02)	-0.08*** (0.02)
Log-likelihood	-331.02	-449.83
Observations	752	752

Note. The reported average partial effects are averages over the Turkish population. Standard errors, corrected for clustering at the vacancy level and calculated using the delta method, are in parentheses. ***(**)((*)) indicates significance at the 1% (5%) ((10%)) level.

Table 3.10: Occupations in Experimental Dataset: Descriptive Statistics

Occupation	**Median vacancy duration in days** (in 2011)	**Bottleneck occupation** (in 2011)	**Moderate customer contact**	**Intensive customer contact**	**Middle educated obs.**	**High educated obs.**
Consultant in recruitment and selection	13	No	Yes	No	0	2
Executive clerk	28	No	No	No	2	0
Administrative clerk	28	No	No	No	36	26
Tutor	29	No	Yes	No	0	2
Window cleaner	30	No	Yes	No	2	0
Industrial cleaner	32	Yes	No	No	2	0
Consultant in marketing and publicity	35	No	Yes	No	0	12
Accountancy clerk	35	No	No	No	0	44
Executive assistant human resources	38	No	No	No	0	4
Warehouseworker components and parts	40	No	No	No	2	0
Assistant bookkeeper	41	No	No	No	0	20
Notary clerk	41	No	No	No	0	10
Teller financial institutions	41	No	No	Yes	0	8
Customs declaration officer	41	Yes	No	No	0	2
Executive assistant general directorate	42	No	No	No	0	4
Classic cleaner	42	Yes	No	No	94	0
Seller	44	Yes	No	Yes	6	0
Adjuster of a packaging machine	46	No	No	No	10	0
Legal service clerk	47	No	No	No	0	26
Bank clerk	47	No	Yes	No	0	8
Production worker	47	No	No	No	62	0
Bookkeeper	50	Yes	Yes	No	0	56
Room attendant	52	No	Yes	No	2	0
Executive expedition operator	55	Yes	No	No	0	10
Car cleaner	55	No	Yes	No	12	0
Executive assistant sales, marketing and publicity	56	No	No	No	0	8
Commercial clerk	56	No	No	Yes	18	28
Planning and logistics clerk	56	Yes	No	No	0	26
Private cleaner	65	Yes	No	Yes	74	0
Accountant	68	Yes	Yes	No	0	18
Shipping agent at the quay	69	Yes	No	No	0	6
Investigator	70	No	Yes	No	2	0
Insurance clerk	73	No	Yes	No	0	30
Representative	80	Yes	No	Yes	30	6
Call center employee	84	Yes	Yes	No	22	0
Consultant in finance	106	No	Yes	No	0	6
Tele-seller	106	Yes	No	Yes	10	0
Demonstrator	109	No	No	Yes	4	0

4

On Grade Retention, Track Mobility and Secondary School Completion

This chapter is joint work with Prof. Dr. Bart Cockx (Ghent University, Université catholique de Louvain, CESifo and IZA) and Prof. Dr. Matteo Picchio (Marche Polytechnic University, Tilburg University, Ghent University and IZA).

4.1 Introduction

One of the most notable differences between school systems across OECD countries consists in grade retention policies.[1] Grade retention is used in many countries as a tool to improve poor academic performances. The hypothesis is that, by resitting the same grade, low-achieving students have extra time to catch up to the grade-level requirements, in terms both of knowledge and emotional maturity. By having more time to develop the skills needed in the subsequent grades, resitting students should be less at risk of failure in the future. Moreover, the threat of retention might be an incentive device to work more diligently and harder. However, retention might generate personal and academic costs with both short- and long-term effects, since it might: hurt pupils' self-esteem (Browman, 2005; Byrd et al., 1997); generate psychological

[1] See OECD (2004, p. 262) for a comparison of the features of school systems of OECD countries.

costs of separating students from their peers (Alexander et al., 1994); produce financial costs to the families and to society in terms of teaching resources (Eide and Goldhaber, 2005).

In the present study we examine the short-term and permanent effects of grade retention on later success rates in school. We use econometric modelling tools and identification analysis to examine the interrelated dynamics of secondary school grade retention, school track choices and achievements of a sample of Belgian pupils living in Flanders. We also shed light on the role played by family background and unobserved abilities, especially looking at how unobserved abilities interact with retention episodes in determining schooling pathways.

The empirical analysis is carried out using the SONAR dataset, a retrospective survey conducted in Flanders on the 1976, 1978 and 1980 cohorts. The SONAR dataset contains very rich information on education, but also on family and labour market experiences. Our sample is made up of 4,214 students belonging to the 1978 and 1980 cohorts. We exploit the ample information on secondary school performances and choices available for these two cohorts to estimate dynamic qualitative choice models.

The identification of the interrelated dynamics between grade retention, track mobility and schooling attainment is obtained by addressing some key challenges. First, educational achievements and choices are likely to be determined by a set of unobserved determinants, for instance behavioural and cognitive skills, with an unknown correlation structure. In order to disentangle the pure effects of past educational outcomes on future ones from the spurious effects determined by unobserved abilities, we take into account the presence of unobserved heterogeneity by semi-parametric maximum likelihood techniques (Heckman and Singer, 1984; Mroz, 1999). Second, at the start of secondary school pupils have already different years of delay due to retention episodes either in kindergarten or in primary school. If we assume that grade retention affects future outcome variables, we have an initial conditions problem. The years of delay at the beginning of secondary school cannot be easily assumed to be exogenous, since they are very likely correlated to the unobserved determinants. We solve for initial conditions by adding an equation for the years of delay at the beginning of secondary school which depends on unobserved heterogeneity and an exclusion restriction (Heckman, 1981a). Third, as pointed out by Fruehwirth et al. (2011), the effect of grade retention might be heterogeneous and vary by students' unobserved abilities. We allow therefore the effect of past retention episodes to vary across different levels of the unobserved determinants. Fourth, there might be sample selection attrition induced by students dropping out of secondary school. We model therefore also the probability of exiting school without the secondary education diploma at the end of each year from the end of compulsory education onwards, where the unobserved components determining the school drop-out are allowed to be correlated to the unobserved determinants of the other endogenous processes.

In contrast to most of the previous findings, we find that grade retention has a positive impact on the next evaluation and can permanently affect subsequent performances. The direction of the permanent effect depends on unobserved heterogeneity. While more able students are permanently penalised by retention, less able students benefit from it. We conclude that when looking for the optimal retention policy, the interaction effect between retention and students' abilities should be taken into account.

This study is organised as follows. In Section 4.2, we present the educational system of secondary school in Flanders (Belgium). Section 4.3 describes the data and summarises basic descriptive statistics of the variables used in the empirical analysis. Section 4.4 presents the econometric model. Section 4.5 reports the estimation results. Section 4.6 concludes.

4.2 The Flemish Secondary School Educational System

In this study we use data from Flanders, the Dutch speaking region of Belgium, situated in the northern part of the country. Belgium is a federal country with several competences devolved to its three Regions (Flanders, Brussels and Wallonia) and three Communities (Dutch, French and German speaking). While the federal authorities are competent for all matters of national importance, territorial and person-related issues are left to Regions and Communities. The Flemish Community is in charge of all aspects of education policy in Flanders.

Nationwide, the Belgian Constitution states that every child has the right to education, which is granted by a compulsory education law. Compulsory education starts on 1 September of the year in which the child turns 6 years old and ends on 30 June of the year in which (s)he reaches the age of 18.[2] Children start primary school in the year in which they turn 6 years old. However, they might start one year earlier or some years later if in kindergarten they are suggested to do so.[3] Grade retention and grade skipping are also allowed in primary school. Hence, pupils may start secondary school at different ages. In case of no retention or skipping in primary school and regular age at the beginning of primary school, pupils start secondary school in the year they turn 12 years old.[4]

In Flanders, when entering secondary school, students formally choose between hierarchical ordered tracks. Students are grouped or tracked according to their abilities and interests, a

[2] Starting from the age of 15 (conditional on passing the first two years of full-time secondary education) or the age of 16 (unconditionally), only part-time education is mandatory.

[3] In our sample, 1.4% of children started primary school in the year they turned 5 and 1.1% started it when 7 or 8.

[4] Out of 4,214 pupils in our sample, only 46 (1.1%) started secondary school in the year they turned 11 and 176 (4.2%) started secondary school with delay.

quite common practice in OECD countries to take into account the diversity of skills and preferences of pupils in education. In this study, as in Van de gaer et al. (2006) and Van Houtte et al. (2012), we refer to 'tracking' as the situation in which students are taught entirely different curricula depending on their curriculum choice which may be restricted after unsatisfactory performances. This is different from 'setting' or 'banding', where pupils in the same curriculum are taught at different difficulty levels given their ability (Gamoran et al., 1995). The Flemish secondary school system consists of several tracks which can be divided into four main education forms: i) general education (ASO) which emphasises general education and provides firm foundations for tertiary education; ii) technical education (TSO) which provides general foundations for practising a profession; iii) art education and iv) vocational education (BSO) which is oriented to the accumulation of skills for a specific profession. In this study, we do not consider the art education track, because of the small number of pupils in our sample choosing it. Our analysis is limited to track choices and track mobility between ASO, TSO and BSO. Students obtain the secondary school diploma if they successfully pass the 6 grades of ASO and TSO and the 7 grades of BSO. All the secondary school diplomas give access to tertiary education.

Track mobility in secondary school is allowed with the following constraints and features. First, track change is not permitted at the beginning of the last grade, hence at the beginning of grade 6 for the ASO and the TSO tracks and grade 7 for the BSO track. Second, tracks are hierarchical and moving upward is not allowed; it is not possible to go from BSO to TSO/ASO or from TSO to ASO. It is anyway possible at the beginning of each academic year to downgrade the track and move from ASO to TSO/BSO and from TSO to BSO. Finally, track mobility is also possible at a finer level within the ASO, TSO and BSO tracks. Within each major track, it is indeed possible to identify hierarchical subtracks with different curricula of different complexity for which the just mentioned track mobility constraints are satisfied. The data at hand allow us to identify two hierarchical subtracks for ASO, which we name ASO+ and ASO− and two hierarchical subtracks for TSO, labelled TSO+ and TSO−.[5]

At the end of each academic year, pupils receive an evaluation: A, B, or C. Pupils getting an A can access the next grade and, if they wish, can downgrade the track. Pupils obtaining a C must resit the grade and, if they wish, can downgrade the track. Pupils getting a B can decide whether to resit the grade or not. If they decide to resit, they can stay in the same track. If they decide not to resit the grade, they must downgrade the track.

Given the set-up of the Flemish educational system, there are different choices that pupils

[5]More concretely, ASO+ comprises the curricula including Latin and Ancient Greek and TSO+ comprises the curricula focussed on industrial sciences and on commerce.

(or/and their parents) have to make in each academic year. First, they have to decide the track. Second, if at the end of the year they get a B, they have to decide whether to resit the grade or not. Finally, they have to decide whether to downgrade the track. Once they turn 18 years old, they can also choose to drop-out the school without the diploma. We will model all these choices and students' performances (evaluation and secondary school completion) in a multiple-equations dynamic model for categorical outcome variables, where past choices and past performances are allowed to affect future schooling pathways.

4.3 Data and Sample

The dataset used in the empirical investigation comes from the SONAR survey. The SONAR survey retrospectively collected information on education, family background, family formation and labour market experiences for a sample of almost 9,000 of individuals living in Flanders and born in 1976, 1978, or 1980.[6] The 1976 cohort was interviewed thrice, at age 23, 26 and 29. The 1978 cohort was interviewed twice, at age 23 and 26. Finally, the 1980 cohort was interviewed only once at age 23. While we only know starting and ending years of primary school, for secondary school we have detailed information, year by year, on school track choices, evaluations, school drop-out and obtaining the diploma.

Since there is no detailed information on tracks for the 1976 cohort, we removed it from the sample and are left with 5,953 pupils. In order to have a sample of pupils with a homogeneous educational, social and family background, we removed from the sample pupils whose grandmother on mother's side have a foreign nationality, pupils who need special help, temporarily or permanently, and are therefore in special schools and pupils who start secondary school when older than 15. We also deleted those entering the art curriculum, those reporting a break of one or more years in secondary school attendance, those leaving school before the end of compulsory education and those with inconsistent or missing information on the progression of the grade, evaluation and grade mobility. After applying these selection criteria, we end up with a sample of 4,214 pupils. The exit from secondary school might take place with or without the diploma. In our sample there are students who are retained multiple times; the observed maximum number of years in secondary school is 11. If students move to part-time education, they are censored in the year they move to it. Hence, we use all the information

[6]A study of the representativeness of the sample was conducted by the SONAR group and reported in SONAR (2000b). The sample is representative with respect to gender. Comparing the sample with respect to other characteristics is more difficult because of a lack of comparable data. A cautious comparison with statistics of the Ministry of Education and the Labor Force Study reveals that the sample is representative with respect to family formation. The lower educated, the unemployed and respondents from lower social classes are instead somewhat under-represented.

until the transition to part-time education, but we disregard all the information from the moment of entering part-time education.[7]

Table 4.1 reports summary statistics of schooling attainment and choices which we model in the empirical analysis. First, we report some outcomes and decisions at the end of the schooling year averaged over the secondary education career. In our sample on average almost 90% get an A, the highest evaluation, while about 6% and 4% are assigned a B and a C, respectively. Around 5% of the pupils are retained on average at the end of the academic year.[8] We hardly see track transitions involving downgrades of more than 2 steps: only 48 track transitions involve a downgrade of three steps and nobody makes a 4-step downgrade, that is from ASO to BSO. Hence, given the starting track, information on track changes compressed in no downgrade, 1-step downgrade and 2-step downgrade is able to describe almost all the possible track transitions. In 90% of the cases, pupils stay in the same track, while 7.5% of the students start the new year with a 1-step track downgrade and 2.5% with a 2-step downgrade. Second, Table 4.1 shows the average cumulative delay at the beginning of grade 1 and grade 2 and at the end of secondary school (irrespective of whether one exits with or without a diploma). At the beginning of secondary education, the average number of years of schooling delay is 0.03. No student is retained at this first grade. By the end of secondary school pupils are on average retained for 0.32 years. Third, Table 4.1 reports the relative frequency of track choices at the beginning of grade 1 and grade 2. At the beginning of grade 1, we have only partial information about the school track choice. We only know whether the student is in the vocational track (BSO) or not (ASO/TSO). This partial observability generates a complication in modelling track choice at the start of grade 1 and subsequent downgrades. We explain how we deal with it in Subsection 4.4.4. Only starting from grade 2, we have more detailed information on the tracks and we can group track choices into five hierarchical categories: ASO+, ASO−, TSO+, TSO− and BSO. At the beginning of grade 1, 6.3% of pupils choose BSO. As a result of some downgrading decisions, this frequency increases almost up to 10% when moving to the second grade; 27% are instead in ASO+, 40% in ASO− and the remaining 23% is split almost evenly between TSO+ and TSO−. By the end of secondary education 19% of the pupils are in BSO, 13% are in ASO+, 36% in ASO−, 11% in TSO+ and 22% in TSO−. Finally, out of the 4,214 pupils who start secondary school, 4.4% enter part-time education and are therefore censored in our model, 86.5% are able to get the full-time secondary education diploma, while the remaining 9.2% drop out of secondary school without the diploma.

[7]Since only 184 students left full-time education for part-time education, we preferred not to model their transition to part-time education and their future schooling experiences.

[8]This figure is in line with the figures reported in OECD (2004, p. 262) for the whole Belgium.

Table 4.1: Summary Statistics of Outcome Variables: Schooling Attainment and Choices

	Mean	Std. Dev.
Outcomes and decisions at the end of the year		
Evaluation: A	0.897	0.304
Evaluation: B	0.059	0.235
Evaluation: C	0.044	0.206
Retention	0.054	0.226
No downgrade	0.900	0.300
1-step downgrade	0.075	0.263
2-step downgrade (or more)	0.025	0.157
Cumulative delay		
Cumulative delay at the beginning of grade 1	0.031	0.228
Cumulative delay at the beginning of grade 2	0.031	0.228
Cumulative delay at the end of secondary education	0.319	0.624
Track at the beginning of grade 1		
ASO/TSO	0.938	0.242
BSO	0.063	0.242
Track at the beginning of grade 2		
ASO+	0.272	0.445
ASO−	0.403	0.490
TSO+	0.095	0.293
TSO−	0.132	0.339
BSO	0.098	0.298
Track at the end of secondary school		
ASO+	0.133	0.340
ASO−	0.360	0.480
TSO+	0.105	0.306
TSO−	0.218	0.413
BSO	0.185	0.388
Exit from secondary school		
With diploma	0.865	0.342
Without diploma	0.092	0.289
Censored to part-time education	0.044	0.204
Number of pupils	4,214	
Number of pupils $\times$ number of years of schooling	26,313	

Note. The presented outcomes and decisions at the end of the year are yearly averages over the secondary education career.

Table 4.2: Summary Statistics of Covariates at the Beginning of Secondary School

	Mean	**Std. Dev.**
Female	0.502	0.500
Calendar day of birth	183.9	104.8
Father's education (years)	6.201	3.339
Mother's education (years)	5.809	3.032
Age at the beginning of primary school		
5 years old	0.014	0.115
6 years old	0.976	0.154
7 years old	0.011	0.103
8 years old	0.0002	0.014
Age at the beginning of secondary school		
11 years old	0.011	0.104
12 years old	0.947	0.223
13 years old	0.042	0.200
Cohort		
1978	0.497	0.500
1980	0.503	0.500
Presence of siblings		
0	0.138	0.345
1	0.465	0.499
2	0.257	0.437
3 or more	0.140	0.347
Number of pupils	4,214	

Note. Father's and mother's education measure the number of successful schooling years beyond primary school, which lasts 6 years.

Table 4.2 reports descriptive statistics of the covariates used in the econometric analysis. Most of the pupils start primary education in time (97.6%), that is in the year they turn 6. The fraction of those starting in time secondary school is smaller and equal to 94.7%, while the fraction of those who start late rises from 1.3% in primary school to 4.2% in secondary due to retention in primary school. Almost one half of the pupils have a sibling, 13.8% are only child and almost 40% have more than one sibling. Pupils' fathers are more educated than pupils' mothers, having on average 6.2 years of successful education beyond primary school against 5.8 years for mothers.

4.4 The Econometric Model

In this section, we write down the likelihood function and clarify the identifying assumptions. Finally, we deal with the problem of partial observability of tracks at the start of secondary school.

4.4.1 Model Specification and the Likelihood Function

If we aim at understanding the determinants of educational achievements in secondary school, we have to take into account that many determinants are potentially endogenous variables. For example, the total number of years of delay with which students start each grade and the different track choices they make might influence future schooling attainment and decisions, but are at the same time the results of past performances and choices. Performances and choices might be correlated across equations and over time due to the presence of unobserved heterogeneity. If we wish to disentangle the causal effects from the spurious ones, we have to control for it.

The six outcome variables that we model for each student i at each academic year t, with $i = 1, \cdots, N$ and $t = 1, \cdots, T$ are:

- Track choice at the beginning of secondary school (tr_i). Since tracks are hierarchically ordered, tr_i is an ordered response taking on the increasing values $\{\text{BSO}, \text{TSO}-, \text{TSO}+, \text{ASO}-, \text{ASO}+\}$.

- Evaluation at the end of each academic year (ev_{it}) or, if in the last grade, the success in getting the diploma (di_{it}). ev_{it} is an ordered response taking on the increasing values $\{\text{C}, \text{B}, \text{A}\}$. di_{it} is instead binary and equal to 1 if the student gets the diploma at the end of the academic year or equal to 0 if (s)he fails the last grade and has to resit.

- School drop out (out_{it}) if turning 18 (compulsory schooling age) in that calendar year or older than 18. out_{it} is a dummy indicator equal to 1 if the student drops out of school, 0 otherwise.

- Resitting decision (re_{it}) if the evaluation is B (ev_{it} = B). re_{it} is a dichotomous variable equal to 1 if the student chooses to resit when (s)he gets a B, 0 if (s)he chooses instead to downgrade.

- Track downgrade (dow_{it}) which is defined as an ordered response taking values on $\{0, 1, 2\}$, where 0 means 'no downgrade', 1 stands for '1-step downgrade' and 2 is '2-step downgrade'.

Furthermore, we have an initial conditions equation for the number of years of delay (in_i) at the beginning of secondary school. As mentioned before, pupils start secondary school at different ages due to different past retention histories either in primary school or in kindergarten. If we assume that past performances like past grade retention affect future outcome variables, we have an initial conditions problem. The years of delay at the beginning of secondary school cannot be easily assumed to be exogenous, since they are very likely correlated to the unobserved determinants. We solve for initial conditions by adding an equation for the years of delay at the beginning of secondary school which depends on unobserved heterogeneity and an exclusion restriction (Heckman, 1981a). in_i takes values on $\{-1, 0, 1\}$. It is equal to 0 when the student starts secondary school without delay, that is in the year in which (s)he turns 12, to -1 if one year in advance and to 1 if one year late.

Let $\mathbf{Y}_{it} \equiv (ev_{it}, di_{it}, out_{it}, re_{it}, dow_{it})$ be the vector collecting the five time-varying outcome variables and $\mathbf{z}_i$ be the vector of observed explanatory variables. Denote by $\mathbf{v}_i \in \mathbb{R}^7$ a random vector of equation-specific time-invariant covariates that are unobserved to the analyst. This vector of unobserved determinants has an unknown cumulative distribution function G.

We can always write the density of $(in_i, tr_i, \mathbf{Y}_i)$ conditional on $(\mathbf{z}_i, \mathbf{v}_i)$ as:

$$\begin{aligned} f(in_i, tr_i, \mathbf{Y}_i|\mathbf{z}_i, \mathbf{v}_i) &= f(in_i|\mathbf{z}_i, \mathbf{v}_i) \cdot f(tr_i|\mathbf{z}_i, \mathbf{v}_i, in_i) \\ &\cdot \prod_{t=1}^{T} f(\mathbf{Y}_{it}|\mathbf{z}_i, \mathbf{v}_i, \mathbf{Y}_{it-1}, \cdots, \mathbf{Y}_{i1}, tr_i, in_i) \\ &= f(in_i|\mathbf{z}_i, \mathbf{v}_i) \cdot f(tr_i|\mathbf{z}_i, \mathbf{v}_i, in_i) \cdot \prod_{t=1}^{T} f(\mathbf{Y}_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1}), \end{aligned} \quad (4.4.1)$$

where $\Im_{it-1}$ denotes the information set containing all the realisations of the endogenous variables from $t-1$ until the beginning of the processes, i.e. $\Im_{it-1} = (\mathbf{Y}_{it-1}, \cdots, \mathbf{Y}_{i1}, tr_i, in_i)$.

Assumption 1 *(Sequentiality):*
Within each academic year t and for $t = 1, 2, \cdots, T$, the five time-varying outcome variables in $\mathbf{Y}_t$ are realised sequentially with the following chronological order: performance at the end of the year, either evaluation or achieving the diploma, $(ev_t \vee di_t)$; school drop-out decision (out_t); resitting decision (re_t); track downgrade decision (dow_t).

Given Assumption 1 on the sequentiality of the realisations of the endogenous variables, it is meaningful to rewrite the conditional density in Equation (4.4.1) as:

$$\begin{aligned} f(in_i, tr_i, \mathbf{Y}_i|\mathbf{z}_i, \mathbf{v}_i) \; = \; & f(in_i|\mathbf{z}_i, \mathbf{v}_i) \cdot f(tr_i|\mathbf{z}_i, \mathbf{v}_i, in_i) \\ & \cdot \prod_{t=1}^{T} \Big[f(ev_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1})^{1-g_{it}} f(di_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1})^{g_{it}} \\ & \cdot \; f(out_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1}, ev_{it})^{s_{it}} \\ & \cdot \; f(re_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1}, ev_{it} = B)^{1-g_{it}} \\ & \cdot \; f(dow_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1}, ev_{it}, re_{it})^{c_{it}} \Big], \end{aligned} \tag{4.4.2}$$

where g_{it} is an indicator variable equal to 1 if the student is in the last grade of secondary school and 0 otherwise, s_{it} is a dummy equal to 1 if the student belongs to the set at risk of school drop-out (legally allowed to drop out) and c_{it} is equal to 1 if the student is in the ASO/TSO tracks and 0 if (s)he is in the BSO track (BSO students do not have the option to downgrade as already at the bottom of the track hierarchy). We cannot derive the likelihood function on the basis of the density in Equation (4.4.2), because we do not observe $\mathbf{v}_i$. Instead, we integrate $\mathbf{v}_i$ out after assuming that it is orthogonal to $\mathbf{z}_i$.

Assumption 2 *(Orthogonality):*
$\mathbf{v}_i \perp\!\!\!\perp \mathbf{z}_i$.

Under Assumption 2 on the orthogonality between the exogenous covariates and the unobservables we can integrate $\mathbf{v}_i$ out once we specify its cumulative distribution function $G(\mathbf{v}_i)$, yielding the following marginal density:

$$\begin{aligned} f(in_i, tr_i, \mathbf{Y}_i|\mathbf{z}_i) \; = \; & \int_{\mathbb{R}^7} f(in_i|\mathbf{z}_i, \mathbf{v}_i) \cdot f(tr_i|\mathbf{z}_i, \mathbf{v}_i, in_i) \\ & \cdot \prod_{t=1}^{T} \Big[f(ev_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1})^{1-g_{it}} f(di_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1})^{g_{it}} \\ & \cdot \; f(out_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1}, ev_{it})^{s_{it}} \\ & \cdot \; f(re_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1}, ev_{it} = B)^{1-g_{it}} \\ & \cdot \; f(dow_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1}, ev_{it}, re_{it})^{c_{it}} \Big] \mathrm{d}G(\mathbf{v}_i). \end{aligned} \tag{4.4.3}$$

Providing an empirical specification to the each of the probability density functions in Equation (4.4.3) leads to the sample log-likelihood function:

$$
\begin{aligned}
\ell(\boldsymbol{\theta}, \boldsymbol{\delta}) \;&=\; \sum_{i=1}^{N} \ln\Big[\int_{\mathbb{R}^7} \mathcal{L}_i(\boldsymbol{\theta}, \boldsymbol{\delta})\Big] \mathrm{d}G(\mathbf{v}_i; \boldsymbol{\delta}) \\
&=\; \sum_{i=1}^{N} \ln\Big\{\int_{\mathbb{R}^7} f(in_i|\mathbf{z}_i, \mathbf{v}_i; \boldsymbol{\theta}_{in}) \cdot f(tr_i|\mathbf{z}_i, \mathbf{v}_i, in_i; \boldsymbol{\theta}_{tr}) \\
&\cdot\; \prod_{t=1}^{T} \big[f(ev_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1}; \boldsymbol{\theta}_{ev})^{1-g_{it}} f(di_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1}; \boldsymbol{\theta}_{di})^{g_{it}} \\
&\cdot\; f(out_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1}, ev_{it}; \boldsymbol{\theta}_{out})^{s_{it}} \\
&\cdot\; f(re_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1}, ev_{it} = B; \boldsymbol{\theta}_{re})^{1-g_{it}} \\
&\cdot\; f(dow_{it}|\mathbf{z}_i, \mathbf{v}_i, \Im_{it-1}, ev_{it}, re_{it}; \boldsymbol{\theta}_{dow})^{c_{it}}\big]\Big\} \mathrm{d}G(\mathbf{v}_i; \boldsymbol{\delta}),
\end{aligned}
\tag{4.4.4}
$$

where $\mathcal{L}_i(\boldsymbol{\theta}, \boldsymbol{\delta})$ is the individual contribution to the likelihood and $\boldsymbol{\theta}$ and $\boldsymbol{\delta}$ are parameters fully characterising the probability density functions with respect to which the sample log-likelihood will be maximised.

Assumption 3 *(Logit and ordered logit probability density functions):*
The probability density functions of both dichotomous and ordered response outcome variables are assumed to have a logit form.

In Subsection 4.4.2, we clarify in more detail how the explanatory variables and past realisations enter the specification of the logit and ordered logit models of the probability density functions in the log-likelihood function (4.4.4). We also explain how we deal with the unobserved heterogeneity distribution and how we allow the unobserved determinants to interact with retention episodes.

4.4.2 The Empirical Specification

4.4.2.1 The Initial Conditions

Students start secondary school at different ages, meaning that they have different numbers of years of delay. This is due to a delayed beginning of primary school and/or retention in primary school. In our econometric model, years of delay at the beginning of each secondary school year can affect schooling choices and performances. This variable evolves over time according to the realisation of episodes of retention, which is also one of the outcome variables. As such, years of delay at the beginning of secondary school cannot be assumed to be a nonstochastic starting position for each student. It is very likely to be endogenous since correlated to the unobserved determinants of schooling choices and performances. This results in an initial conditions problem that we solve by specifying an ordered logit model for the years of delay, where unobserved characteristics are allowed to be correlated to those determining future outcomes and choices.

We specify the probability density function of the number of years of delay at the beginning of secondary school as an ordered logit model. This outcome variable takes on the values -1, 0 and 1. We define as $\alpha_{1,in} < \alpha_{2,in}$ the unknown cut points (threshold parameters) and as Λ the logit function. The unobserved heterogeneity component $v_{i,in}$ enters the specification as a shift in the threshold parameters. The probability density function of the initial conditions is:

$$\begin{aligned}
\Pr(in_i = -1|\mathbf{z}_i, v_{i,in}) &= \Lambda(\alpha_{1,in} + v_{i,in} - \mathbf{z}_i'\boldsymbol{\beta}_{in}), \\
\Pr(in_i = 0|\mathbf{z}_i, v_{i,in}) &= \Lambda(\alpha_{2,in} + v_{i,in} - \mathbf{z}_i'\boldsymbol{\beta}_{in}) - \Lambda(\alpha_{1,in} + v_{i,in} - \mathbf{z}_i'\boldsymbol{\beta}_{in}), \\
\Pr(in_i = 1|\mathbf{z}_i, v_{i,in}) &= 1 - \Lambda(\alpha_{2,in} + v_{i,in} - \mathbf{z}_i'\boldsymbol{\beta}_{in}). \qquad (4.4.5)
\end{aligned}$$

Students with a higher level of $v_{i,in}$ are less likely to end up into the top category, that is to start secondary school one year late.

4.4.2.2 The Track Choice at the Start of Secondary School

The track choice takes value on {BSO, TSO−, TSO+, ASO−, ASO+}. The probability density function of the choice of the hierarchically ordered tracks is:

$$\begin{aligned}
\Pr(tr_i = \text{BSO}|\mathbf{x}_i, in_i, v_{i,tr}) &= \Lambda(\alpha_{1,tr} + v_{i,tr} - \mathbf{x}_i'\boldsymbol{\beta}_{tr} - in_i\gamma_{tr}), \\
\Pr(tr_i = \text{TSO}-|\mathbf{x}_i, in_i, v_{i,tr}) &= \Lambda(\alpha_{2,tr} + v_{i,tr} - \mathbf{x}_i'\boldsymbol{\beta}_{tr} - in_i\gamma_{tr}) - \Lambda(\alpha_{1,tr} + v_{i,tr} - \mathbf{x}_i'\boldsymbol{\beta}_{tr} - in_i\gamma_{tr}), \\
\Pr(tr_i = \text{TSO}+|\mathbf{x}_i, in_i, v_{i,tr}) &= \Lambda(\alpha_{3,tr} + v_{i,tr} - \mathbf{x}_i'\boldsymbol{\beta}_{tr} - in_i\gamma_{tr}) - \Lambda(\alpha_{2,tr} + v_{i,tr} - \mathbf{x}_i'\boldsymbol{\beta}_{tr} - in_i\gamma_{tr}), \\
\Pr(tr_i = \text{ASO}-|\mathbf{x}_i, in_i, v_{i,tr}) &= \Lambda(\alpha_{4,tr} + v_{i,tr} - \mathbf{x}_i'\boldsymbol{\beta}_{tr} - in_i\gamma_{tr}) - \Lambda(\alpha_{3,tr} + v_{i,tr} - \mathbf{x}_i'\boldsymbol{\beta}_{tr} - in_i\gamma_{tr}), \\
\Pr(tr_i = \text{ASO}+|\mathbf{x}_i, in_i, v_{i,tr}) &= 1 - \Lambda(\alpha_{4,tr} + v_{i,tr} - \mathbf{x}_i'\boldsymbol{\beta}_{tr} - in_i\gamma_{tr}), \qquad (4.4.6)
\end{aligned}$$

where $\mathbf{x}_i \subset \mathbf{z}_i$ due to an exclusion restriction.

As exclusion restriction, we use the years of delay at the beginning of primary school. We assume therefore that, at the beginning of secondary school, choices and performances are not affected by the years of delay at the beginning of primary school but just by the years of delay at the start of secondary school, conditional on the other covariates and unobserved heterogeneity.

4.4.2.3 The Evaluation

At the end of each academic year, pupils receive an evaluation: A, B, or C. As mentioned before, an A allows students to move to the next grade. Students getting a C must resit the grade. Students with a B can decide to downgrade the track if they wish to avoid resitting the grade. The probability density function of the evaluation variable is specified as follows:

$$\begin{aligned}
\Pr(ev_{it}=C|\mathbf{x}_i, in_i, tr_i, \Im_{it-1}, v_{i,ev}) &= \Lambda\left[\alpha_{1,ev}+v_{i,ev}-\mathbf{x}_i'\boldsymbol{\beta}_{ev}-\phi_{ev}(in_i, tr_i, \Im_{it-1})\right], \\
\Pr(ev_{it}=B|\mathbf{x}_i, in_i, tr_i, \Im_{it-1}, v_{i,ev}) &= \Lambda\left[\alpha_{2,ev}+v_{i,ev}-\mathbf{x}_i'\boldsymbol{\beta}_{ev}-\phi_{ev}(in_i, tr_i, \Im_{it-1})\right] \\
&\quad -\Lambda\left[\alpha_{1,ev}+v_{i,ev}-\mathbf{x}_i'\boldsymbol{\beta}_{ev}-\phi_{ev}(in_i, tr_i, \Im_{it-1})\right], \\
\Pr(ev_{it}=A|\mathbf{x}_i, in_i, tr_i, \Im_{it-1}, v_{i,ev}) &= 1-\Lambda\left[\alpha_{2,ev}+v_{i,ev}-\mathbf{x}_i'\boldsymbol{\beta}_{ev}-\phi_{ev}(in_i, tr_i, \Im_{it-1})\right], \quad (4.4.7)
\end{aligned}$$

where $\phi_{ev}(in_i, tr_i, \Im_{it-1})$ is the impact of past outcome variables on future evaluations. We impose some parametric restrictions on the way in which the past is allowed to affect the future. We keep in mind that, from the policy perspective, it is of interest to understand whether and how students' performance is affected by past retention episodes and by past track downgrades. The impact of past outcome variables on future evaluations is modelled as follows:

$$\begin{aligned}
\phi_{ev}(in_i, tr_i, \Im_{it-1}) &= \eta_{ev} tr_{it} + \pi_{ev} dow_{it-1} \\
&+ \kappa_{ev} re_{it-1} + \tau_{ev} re_{it-1} \cdot ev_{it-1} + \psi_{ev} tre_{it-1}, \quad (4.4.8)
\end{aligned}$$

where tr_{it} is the track at the beginning of the t-th academic year, re_{it-1} is an indicator variable equal to 1 if the individual was retained at the end of the previous year (resitting therefore in the current year) and $tre_{it-1} = in_i + \sum_{s=1}^{t-1} re_{is}$ is the total years of delay at the beginning of the t-th academic year. The coefficients κ_{ev} and τ_{ev} capture the transitory effect of retention on the subsequent academic performance, while and ψ_{ev} is the permanent effect. η_{ev} captures track heterogeneity in the ability of the students to get good evaluations. Finally, π_{ev} is the effect of downgrading at the end of the last year on the current schooling achievement.

4.4.2.4 The School Drop-Out

In Belgium, compulsory education ends on 30 June of the year in which the youth reach the age of 18. From that date onwards, the student is at risk of school drop-out without diploma. Ignoring school drop-out might lead to sample selection attrition as it is not likely to be a random process. We model therefore also the probability of exiting school without the diploma at the end of each year, where the unobserved components determining the school drop-out are allowed to be correlated to the unobserved determinants of the other endogenous processes. In the sequentiality of the events, school drop-out takes place at the end of the academic year, after receiving the evaluation.[9] The school drop-out variable is binary and equal to 1 in case of drop-out. The logit model for pupils at risk of exit is:

$$\begin{aligned} \Pr(out_{it}=1|\mathbf{x}_i, in_i, tr_i, \Im_{it-1}, ev_{it}, v_{i,out}) &= \Lambda\big[\alpha_{out} + v_{i,out} + \mathbf{x}_i'\boldsymbol{\beta}_{out} \\ &+ \phi_{out}(in_i, tr_i, \Im_{it-1}, ev_{it})\big], \end{aligned} \tag{4.4.9}$$

Similar to Equation (4.4.8), the impact of past outcomes on the drop-out probability is:

$$\begin{aligned} \phi_{out}(in_i, tr_i, \Im_{it-1}, ev_{it}) &= \eta_{out} tr_{it} + \pi_{out} dow_{it-1} + \omega_{out} ev_{it} \\ &+ \kappa_{out} re_{it-1} + \tau_{out} re_{it-1} \cdot ev_{it-1} + \psi_{out} tre_{it-1}. \end{aligned} \tag{4.4.10}$$

Compared to Equation (4.4.8), ϕ_{out} has the extra argument, ev_{it}, the evaluation of the just ended academic year. Under the sequentiality assumption (Assumption 1), ev_{it} is predetermined with respect to the realisation of the drop-out variable. Thereby, it acts as a valid exclusion restriction in the drop-out equation.

4.4.2.5 The Resitting Choice for B Students

Students getting a B can choose either to resit or to downgrade the track. The choice is binary and, conditional on getting a B, the probability of resitting the grade is specified as a logit model:

$$\Pr(re_{it}=1|\mathbf{x}_i, in_i, tr_i, \Im_{it-1}, ev_{it}=B, v_{i,re}) = \Lambda\big[\alpha_{re} + v_{i,re} + \mathbf{x}_i'\boldsymbol{\beta}_{re} + \phi_{re}(in_i, tr_i, \Im_{it-1})\big]. \tag{4.4.11}$$

The function $\phi_{re}(in_i, tr_i, \Im_{it-1})$ is parametrised as Equation (4.4.8):

$$\begin{aligned} \phi_{re}(in_i, tr_i, \Im_{it-1}) &= \eta_{re} tr_{it} + \pi_{re} dow_{it-1} \\ &+ \kappa_{re} re_{it-1} + \tau_{re} re_{it-1} \cdot ev_{it-1} + \psi_{re} tre_{it-1}. \end{aligned} \tag{4.4.12}$$

4.4.2.6 The Track Downgrade

In Belgium, at the beginning of secondary school, students can choose among different tracks characterised by different curricula. This tracking system is aimed at grouping students with

[9] Very few students (71, 1.7% of the sample) drop out of school before the end of the academic year. In order to simplify the model and the timing of events, in these cases we advance the drop-out date at the end of the previous academic year, disregarding information on retention and track downgrade of the uncompleted academic year.

similar abilities and preferences. Choosing the right track is important as it will determine future work and education opportunities. However, the initial track choice is not always binding. Students are indeed allowed to switch track at the beginning of a new academic year, although under a set of constraints. The tracks are hierarchically ordered and students can only move from the more general (and more prestigious) tracks to the more specialised and vocationally oriented (and less prestigious) ones. The Belgian system of tracking is therefore often referred to as a 'cascade' system.

We model track transitions by defining a categorical ordered dependent variable for track downgrade. The ordered categories are no downgrade, one-step downgrade and two-step downgrade. They are coded as 0, 1 and 2, respectively.[10] Students in the BSO track are already at the bottom of the cascade and cannot downgrade further. Hence, we model track downgrade only for ASO/TSO students. For BSO students, track downgrade will not give any contribution to the likelihood function. The probability density function of track downgrade for ASO/TSO students is:

$$\begin{aligned}
\Pr(dow_{it} &= 0|\mathbf{x}_i, in_i, tr_i, \Im_{it-1}, ev_{it}, re_{it}, v_{i,dow}) = \\
&\qquad \Lambda\big[\alpha_{1,dow} + v_{i,dow} - \mathbf{x}_i'\boldsymbol{\beta}_{dow} - \phi_{dow}(in_i, tr_i, \Im_{it-1}, ev_{it}, re_{it})\big], \\
\Pr(dow_{it} &= 1|\mathbf{x}_i, in_i, tr_i, \Im_{it-1}, ev_{it}, re_{it}, v_{i,dow}) = \\
&\qquad \Lambda\big[\alpha_{2,dow} + v_{i,dow} - \mathbf{x}_i'\boldsymbol{\beta}_{dow} - \phi_{dow}(in_i, tr_i, \Im_{it-1}, ev_{it}, re_{it})\big] \\
&\quad - \Lambda\big[\alpha_{1,dow} + v_{i,dow} - \mathbf{x}_i'\boldsymbol{\beta}_{dow} - \phi_{dow}(in_i, tr_i, \Im_{it-1}, ev_{it}, re_{it})\big], \\
\Pr(dow_{it} &= 2|\mathbf{x}_i, in_i, tr_i, \Im_{it-1}, ev_{it}, re_{it}, v_{i,dow}) = \\
&\qquad 1 - \Lambda\big[\alpha_{2,dow} + v_{i,dow} - \mathbf{x}_i'\boldsymbol{\beta}_{dow} - \phi_{dow}(in_i, tr_i, \Im_{it-1}, ev_{it}, re_{it})\big]. \qquad (4.4.13)
\end{aligned}$$

The function $\phi_{dow}(in_i, tr_i, \Im_{it-1}, ev_{it}, re_{it})$ is linearly specified as follows:

$$\begin{aligned}
\phi_{dow}(in_i, tr_i, \Im_{it-1}) &= \eta_{dow} tr_{it} + \pi_{dow} dow_{it-1} + \omega_{dow} ev_{it} + \xi_{dow}(1 - re_{it}) \\
&+ \kappa_{dow} re_{it-1} + \tau_{dow} re_{it-1} \cdot ev_{it-1} + \psi_{dow} tre_{it-1}, \qquad (4.4.14)
\end{aligned}$$

where ξ_{dow} is the effect of not being retained at the end of the academic year on the probability of downgrading the track.

[10] In our dataset, we observe only 48 track transitions of three or more steps. Hence, given the knowledge of a starting point, information on track changes compressed in no downgrade, 1-step downgrade and 2-step downgrade is able to describe almost all the possible track transitions.

4.4.2.7 The Diploma Equation

In the last grade of the track (6th grade for ASO/TSO and 7th grade for BSO), students do not receive an evaluation with marks A, B, or C. If they succeed, they simply get the diploma. If they fail, they have to resit the last grade.[11] The performance variable of the last grade of secondary school is therefore binary. We specify the probability of success, which implies getting the secondary school diploma, as a logit model:

$$\Pr(di_{it} = 1|\mathbf{x}_i, in_i, tr_i, \Im_{it-1}, v_{i,di}) = \Lambda\big[\alpha_{di} + v_{i,di} + \mathbf{x}_i'\boldsymbol{\beta}_{di} + \phi_{di}(in_i, tr_i, \Im_{it-1})\big]. \quad (4.4.15)$$

The function $\phi_{di}(in_i, tr_i, \Im_{it-1})$ has a linear form as in Equation (4.4.8):

$$\begin{aligned} \phi_{di}(in_i, tr_i, \Im_{it-1}) &= \eta_{di} tr_{it} + \pi_{di} dow_{it-1} \\ &+ \kappa_{di} re_{it-1} + \tau_{di} re_{it-1} \cdot ev_{it-1} + \psi_{di} tre_{it-1}. \end{aligned} \quad (4.4.16)$$

4.4.2.8 The Unobserved Heterogeneity Distribution

In order to maximise the log-likelihood function in (4.4.4), we need to assign some parametric form to the joint distribution of the unobserved heterogeneity component $\mathbf{v}_i \equiv (v_{i,in}, v_{i,tr}, v_{i,ev}, v_{i,out}, v_{i,re}, v_{i,dow}, v_{i,di})$. In order to avoid too strict parametric assumptions, we follow Heckman and Singer (1984) and assume that $G(\mathbf{v}_i)$ is discrete with a finite and, a priori, unknown number M points of support. However, estimating our model with a seven-dimensional discrete distribution would be computational demanding. Our outcome variables belong to three types: i) the initial conditions; ii) schooling achievements (evaluation and diploma acquisition) and iii) educational choices (track choice, downgrade choice, resitting decision in case of B and drop-out decision). In order to reduce the estimation complexity of the model, we reduce the dimension of $\mathbf{v}_i$ to three by one-factor loading specifications: $v_{i,di} = \delta_{di} \cdot v_{i,ev}$, $v_{i,tr} = \delta_{tr} \cdot v_{i,dow}$, $v_{i,out} = \delta_{out} \cdot v_{i,dow}$ and $v_{i,re} = \delta_{re} \cdot v_{i,dow}$.

On the basis of Monte Carlo simulations for treatment effects in duration models, Gaure et al. (2007) find that the number of the points of support is best chosen by minimising the Akaike Information Criterion (AIC). We follow this recommendation. The probabilities associated to the points of support sum to one and are, $\forall\ m = 1, \ldots, M$, denoted by

$$p^m = \Pr(v_{i,in} = v_{i,in}^m, v_{i,ev} = v_{i,ev}^m, v_{i,dow} = v_{i,dow}^m) \equiv \Pr(\mathbf{v}_i = \mathbf{v}_i^m) \quad (4.4.17)$$

[11] Students failing the last grade of ASO/TSO can also choose to switch to grade 6 of the BSO track, which is taken into account in the model.

and specified as logistic transforms:

$$p^m = \frac{\exp(\lambda^m)}{\sum_{g=1}^{M} \exp(\lambda^g)} \quad \text{with} \quad m = 1, \ldots, M \quad \text{and} \quad \lambda_M = 0. \tag{4.4.18}$$

The sample log-likelihood function in Equation (4.4.4) can be rewritten as

$$\ell(\boldsymbol{\theta}, \boldsymbol{\delta}) = \sum_{i=1}^{N} \ln \Big[\sum_{m=1}^{M} p^m \mathcal{L}_{im}(\boldsymbol{\theta}, \boldsymbol{\delta}) \Big], \tag{4.4.19}$$

where $\mathcal{L}_{im}(\boldsymbol{\theta}, \boldsymbol{\delta})$ is the individual contribution to the likelihood function if the individual is of type m.

During the empirical analysis, we also use an alternative specification of the unobserved heterogeneity support points. In this alternative specification, we allow the points of support of v^m_{ev} and v^m_{dow} to interact with lagged retention and cumulated retention for each $m = 1, \cdots, M$:

$$v^m_{it,ev} = v^m_{i,ev}(1 + \psi_{ev} re_{it-1} + \zeta_{ev} tre_{it-1}) \tag{4.4.20}$$

$$v^m_{it,dow} = v^m_{i,dow}(1 + \psi_{dow} re_{it-1} + \zeta_{dow} tre_{it-1}). \tag{4.4.21}$$

By doing so, we allow the transitory and permanent effect of grade retention to be heterogeneous across unobserved determinants of preferences and choices. In other words, the points of support become time-varying, depending on the retention realisation. These time-varying components have to be plugged into models (4.4.7), (4.4.9), (4.4.11), (4.4.13) and (4.4.15).

4.4.3 Identification

The identification of the interrelated dynamics between grade retention, track mobility and schooling attainment is obtained by addressing some key challenges. In this subsection we summarise the aforementioned characteristics of our model that induce this identification.

First, educational achievements and choices are likely to be determined by a set of unobserved determinants, for instance behavioural and cognitive skills, with an unknown correlation structure. In order to disentangle the pure effects of past educational outcomes on future ones from the spurious effects determined by unobserved abilities, we take into account the presence of unobserved heterogeneity by semi-parametric maximum likelihood techniques (Heckman and Singer, 1984; Mroz, 1999). The identification of the unobserved heterogeneity distribution is based on multiple observations per student of the same processes.

Second, the imposed sequencing of schooling achievements and choices makes some of the outcome variables determinants of later outcomes within each academic year. This generates predetermined exclusion restrictions which are used to identify the interrelated dynamics of schooling achievements and choices.

Third, at the start of secondary school pupils have already different years of delay due to retention episodes either in kindergarten or in primary school. If we assume that grade retention affects future outcome variables, we have an initial conditions problem. We solve for initial conditions by adding an equation for the years of delay at the beginning of secondary school, which depends on unobserved heterogeneity and an exclusion restriction (Heckman, 1981a). As exclusion restriction, we use the number of years of delay at the start of primary school. We assume therefore that once we control for the years of delay at the start of secondary school, the years of delay at the start of primary school do not affect secondary school performances and choices.

Fourth, as pointed out by Fruehwirth et al. (2011), the effect of grade retention might be heterogeneous and vary by students' unobserved abilities. We allow therefore the effect of past retention episodes to vary across different levels of the unobserved determinants by imposing a specific functional forms on the interaction effect.

Finally, there might be sample selection attrition induced by students dropping out of secondary school. We model therefore also the probability of exiting school without the diploma at the end of each year, where the unobserved components determining the school drop-out are allowed to be correlated to the unobserved determinants of the other endogenous processes. The loading factor structure of the unobserved heterogeneity component and the fact that some students are at risk of drop-out for more than one year are of help in identifying the attrition equation.

4.4.4 Partial Observability of Tracks at the Start of Secondary School

As mentioned in Section 4.3, at the beginning of secondary school, we have only partial information about the school track choice. We only know whether students are in the vocational track (BSO) or not (ASO/TSO). Only starting from grade 2 we have detailed information on courses of study and we can group students into the five tracks. However, the cascade system of the institutional set-up jointly with the track position and track mobility of each student in subsequent grades convey some information about the possible starting track. For example, students who are in ASO+ in grade 2, surely were also in ASO+ in grade 1, as track upgrading is not allowed. For the same reason, students in ASO− in grade 2 were not in TSO and BSO tracks in grade 1.

We modify the likelihood function to take into account the partial observability of the track at the beginning of secondary school: we integrate over the possible tracks in grade 1, given future information about tracks and mobility. This is similar to the strategy used by Mroz and Picone (2011) to solve the partial observability of the time in which persons with diabetes progress to the next disease stage.

To show in what direction we modify the likelihood function and keep the notation simple, we rewrite the density in Equation (4.4.1) by ignoring the conditioning on the observed and unobserved covariates and the individual subscript i, yielding

$$f(in, tr, \mathbf{Y}) = f(in)f(tr|in)f(\mathbf{Y}|tr, in). \tag{4.4.22}$$

We assume that the probability of being in each track at the beginning of secondary school is related to the information we have in the future about tracks, mobility choices and performances. Denote by $f(tr|in, \mathbf{Y})$ this probability density function. If we integrate Equation (4.4.22) over the possible tracks, we get

$$f(in, \mathbf{Y}) = f(in) \int f(tr|in)f(\mathbf{Y}|tr, in)f(tr|in, \mathbf{Y})\mathrm{d}tr. \tag{4.4.23}$$

Once we parametrise $f(tr|in)$ and $f(\mathbf{Y}|tr, in)$, like we did in Subsection 4.4.2, we imply a particular parametrisation of $f(tr|in, \mathbf{Y})$:

$$\begin{aligned} f(tr|in, \mathbf{Y}) &= \frac{f(\mathbf{Y}|tr, in)f(tr|in)}{f(\mathbf{Y}|in)} \\ &= \frac{f(\mathbf{Y}|tr, in)f(tr|in)}{\int f(\mathbf{Y}, tr|in)\mathrm{d}tr} \\ &= \frac{f(\mathbf{Y}|tr, in)f(tr|in)}{\int f(\mathbf{Y}|tr, in)f(tr|in)\mathrm{d}tr}. \end{aligned} \tag{4.4.24}$$

Both the numerator and the denominator of Equation (4.4.24) depend indeed on the probability density functions that we have already parametrised in Subsection 4.4.2. Substituting Equation (4.4.24) into Equation (4.4.23) yields

$$f(in, \mathbf{Y}) = f(in) \int \frac{f(tr|in)^2 f(\mathbf{Y}|tr, in)^2}{\int f(\mathbf{Y}|s, in)f(s|in)\mathrm{d}s}\mathrm{d}tr. \tag{4.4.25}$$

Since tracks take value on five categories, the integrals in Equation (4.4.25) are just sums over the five possible realisations. The individual contribution to the likelihood function in Equation (4.4.3) and the sample log-likelihood function in Equation (4.4.4) are modified along the lines dictated by Equation (4.4.25).

Table 4.3: Estimated Probability Masses of the Discrete Unobserved Heterogeneity Distribution and Other Statistics

	Without unobserved heterogeneity (1)	**Time-invariant unobserved heterogeneity (2)**		**Time-variant unobserved heterogeneity (3)**	
Unobserved heterogeneity probability masses (λ_6 is normalised to 0)					
λ_1		-1.595***	(0.458)	-1.545***	(0.467)
λ_2		-2.929***	(0.338)	-2.583***	(0.279)
λ_3		-1.171***	(0.304)	-1.161***	(0.253)
λ_4		0.256	(0.267)	0.290	(0.281)
λ_5		0.099	(0.262)	0.163	(0.261)
Resulting probability masses					
p_1		0.051		0.052	
p_2		0.013		0.018	
p_3		0.078		0.076	
p_4		0.326		0.325	
p_5		0.279		0.286	
p_6		0.252		0.243	
Log-likelihood	-22,380.5	-22,222.2		-22,194.2	
AIC/N	10.679	10.615		10.604	
Number of parameters	120	144		148	
Number of pupils (N)	4,214	4,214		4,214	

Note. Standard errors in parentheses.

4.5 Estimation Results

The econometric model is made up of seven equations. The estimation results of the coefficients of each equation are reported and commented in the next subsections. We display estimation results of three different model specifications: without unobserved heterogeneity, with time-invariant unobserved heterogeneity and with time-invariant unobserved heterogeneity interacted with lagged retention and the cumulated years of delay.

Table 4.3 reports the estimation results of the probability masses of the discrete unobserved heterogeneity distribution. The number of points of support are chosen by minimising the AIC. For both specifications controlling for unobserved heterogeneity the resulting number is 6. The preferred model according to the AIC is the one that encompasses the interactions between the unobserved heterogeneity and lagged retention and cumulated retention.

Table 4.4: Estimation Results of the Initial Conditions Equation: Years of Delay at the Beginning of Secondary School

	Without unobserved heterogeneity (1)		**Time-invariant unobserved heterogeneity (2)**		**Time-variant unobserved heterogeneity (3)**	
Exogenous variables						
Years delay start primary school	5.317***	(0.253)	10.275***	(0.593)	10.505***	(0.621)
Female	−0.137	(0.164)	−0.221	(0.224)	−0.225	(0.224)
Cohort 1980	−0.217	(0.171)	−0.215	(0.228)	−0.198	(0.228)
Calendar day of birth/100	1.549***	(0.141)	1.960***	(0.242)	1.965***	(0.243)
Father's education/10	−0.634**	(0.302)	−0.774*	(0.421)	−0.751*	(0.422)
Mother's education/10	−1.484***	(0.312)	−1.636***	(0.465)	−1.696***	(0.460)
Number of siblings – Reference: No siblings						
1 sibling	0.232	(0.254)	0.340	(0.344)	0.356	(0.342)
2 siblings	0.072	(0.277)	0.177	(0.374)	0.204	(0.375)
3 or more	0.744***	(0.285)	1.029**	(0.401)	1.046***	(0.400)
Unobserved heterogeneity support points						
v_2			−0.871	(8.336)	−1.587	(7.551)
v_3			−10.582***	(1.875)	−11.184***	(1.732)
v_4			−1.313	(1.530)	−1.371	(1.442)
v_5			−10.419***	(1.952)	−10.981***	(1.772)
v_6			−1.997	(2.362)	−2.753	(2.103)

Note. Standard errors in parentheses.

4.5.1 Initial Conditions: Years of Delay at the Start of Secondary School

Table 4.4 reports the estimation results of the ordered logit model for the years of delay at the beginning of secondary school. The estimation results of the initial conditions equation are very stable across the three model specifications. We find that the years of delay at the start of primary school, the exclusion restriction, strongly and positively affects the probability of starting secondary school with delay. The relative age determined by birth date has a significantly negative effect on the years of delay at the beginning of secondary school: the later in the year the kid was born, the higher the probability that (s)he will cumulate years of delay. This evidence is consistent with those in Bedard and Dhuey (2006), Fredriksson and Öckert (2006), Hámori (2007), McEwan and Shapiro (2008), Strøm (2004) and Altwicker-Hámori and Köllő (2012), who find that school starting age has a positive effect on several measures of academic performance.

Parents' education has a significant impact on years of delay at the start of secondary school, especially mother's education: the higher the education of the mother of the student, the lower the probability that the pupil begins secondary education with delay. The number of siblings is also a significant determinant of the years of delay: in line with the effect on test scores in Hámori (2007), we find that pupils with more than two siblings have a higher probability of starting secondary school with delay. This effect might be explained by the fact that in larger families parents have less time to dedicate to each child. It might also be that the number of siblings capture particular social and cultural family background.

4.5.2 Track Choice at the Beginning of Secondary School

Table 4.5 displays the estimation results of the equation for the track choice at the beginning of secondary school. The tracks are hierarchically ordered from the bottom (BSO) to the top (ASO+). The years of delay at the beginning of secondary school significantly reduce the probability of choosing the ASO+ track and increase the probability of preferring the vocational track. When we control for unobserved heterogeneity the impact of past schooling performances captured by the years of delay at the start of secondary school gets smaller in absolute value. This means that part of the effect is spurious: unobserved characteristics, like ability and intelligence, jointly determine the probability of starting late secondary school and the track choice. Once we net out the spurious negative correlation between unobserved ability and the probability of starting late secondary school, the coefficient of the impact is reduced in size.

All the other regressors are highly significant in explaining the school track choice. Girls and pupils from highly educated parents are less likely to choose the vocational track and more likely to get into ASO+. The gender effect might be induced by gender heterogeneous preferences for vocational/technical tracks but also influenced by the socio-cultural environment, the performance expectations and their interaction.[12]

Both mother's and father's education strongly push up the probability of choosing the highest track (ASO+) and discourage the vocational track (BSO), meaning that parents take influence on the education of their children. This is a quite common association found in the educational research literature. See Haveman and Wolfe (1995) for a review of the literature on intergenerational mobility with respect to education and, among others, Bratti et al. (2012), Dustmann (2004) and Falter et al. (2011) for more recent findings on the effect of parental background on pupils' track choices. Also the family structure has an impact on track choice: the larger the number of siblings the higher the probability of choosing the vocational track (BSO). Finally, the younger the pupil, the higher the probability of choosing a lower track.

[12] Guiso et al. (2008) show that the more the culture is gender-equal, the better the girls score in math.

Table 4.5: Estimation Results of the Track Choice Equation

	Without unobserved heterogeneity (1)		**Time-invariant unobserved heterogeneity (2)**		**Time-variant unobserved heterogeneity (3)**	
Exogenous variables						
Female	0.318∗∗∗	(0.058)	0.448∗∗∗	(0.077)	0.428∗∗∗	(0.076)
Cohort 1980	−0.150∗∗∗	(0.058)	−0.238∗∗∗	(0.076)	−0.250∗∗∗	(0.075)
Calendar day of birth/100	−0.166∗∗∗	(0.029)	−0.273∗∗∗	(0.040)	−0.275∗∗∗	(0.039)
Father's education/10	1.651∗∗∗	(0.109)	2.262∗∗∗	(0.155)	2.233∗∗∗	(0.154)
Mother's education/10	1.741∗∗∗	(0.118)	2.310∗∗∗	(0.166)	2.283∗∗∗	(0.165)
Number of siblings – Reference: No siblings						
1 sibling	−0.111	(0.090)	−0.174	(0.118)	−0.178	(0.117)
2 siblings	−0.243∗∗	(0.099)	−0.351∗∗∗	(0.130)	−0.354∗∗∗	(0.128)
3 or more	−0.425∗∗∗	(0.110)	−0.584∗∗∗	(0.142)	−0.585∗∗∗	(0.140)
Endogenous variables						
Years delay start secondary school	−1.277∗∗∗	(0.136)	−0.804∗∗∗	(0.174)	−0.785∗∗∗	(0.176)
Unobserved heterogeneity loading factor						
Loading factor			−6.347∗∗∗	(2.063)	−6.403∗∗∗	(2.144)

Note. Standard errors in parentheses.

4.5.3 Evaluation at the End of the Academic Year

There are several studies in the educational research literature aimed at understanding whether grade retention has a positive or a negative impact on subsequent academic performances. See for instance the literature review in Xia and Kirby (2009) and the meta-analysis in Jimerson (2001). The conclusions are not uncontroversial. Most of the studies find a negative relationship between retention and subsequent academic achievement. However, if the analyst cannot control for all the determinants of grade retention and subsequent performances the estimate will be biased due to a selection bias. Innate ability, intelligence, cognitive skills and commitment to work are determinants of both grade retention and future educational achievements. If they are not properly taken into account, the impact of grade retention will be spurious and biased downwards.

In existing studies, the identification of the causal effect of grade retention mostly relies on controlling for confounding factors or on matching students on the basis of a set of observable characteristics. A few studies address the selection bias by instrumental variables (IV) relying on shifts and discontinuities determined by retention policies (Fruehwirth et al., 2011; Eide and Showalter, 2001; Greene and Winters, 2007; Jacob and Lefgren, 2004, 2009; Manacorda, 2012) or on the independence between the instrument and the selection variable, conditional on the outcome (D'Haultfœuille, 2010). For the French speaking region of Belgium, Belot and Vandenberghe (2013) exploit a reform which reintroduced the possibility of retention in the first grade of secondary education, finding no impact on academic performance.

In this study, we do not need to assume that we are controlling for all factors determining both the treatments (retention and track mobility) and the outcomes (some measures of subsequent performance), like in the matching literature. Moreover, we do not need a valid IV or an exclusion restriction. We rather exploit the longitudinal dimension of the dataset and the availability of multiple observations per student of the achievement and choice variables. This rich information allows us to flexibly identify the unobserved heterogeneity distribution and the correlation between the unobserved determinants of the performance outcomes (for instance evaluation), of the choices (for instance retention and track downgrade) and of the initial conditions (the years of delay at the beginning of secondary school).

The estimation results of the evaluation ordered logit equation are displayed in Table 4.6. In all three specifications, the transitory impact of retention (the coefficient of lagged retention) has a positive impact on the next evaluation. Hence, *ceteris paribus*, pupils who are resitting the grade are less likely to get a C and thereby to be retained again, than students who are not resitting. In contrast, based on model (2) controlling for time-invariant unobserved heterogeneity, for the permanent effect (the coefficient of total years of delay) we get a negative effect. In model (3), which allows the retention effects to be heterogeneous across different levels of the unobserved component v_{ev}, however, also the permanent effect is positive. This means that an episode of grade retention will also have a positive effect on the evaluation of all the next academic years. The evidence of a positive impact of grade retention on future schooling achievements contrasts with prior research. Two exceptions are D'Haultfœuille (2010) and Jacob and Lefgren (2004), who found that in the US and France, respectively, grade retention has a positive short-term effect on schooling performance. The former research based identification on a new method for models with endogenous selection and exploited the independence between an instrument and the selection variable, conditional on the outcome. The latter exploited a discontinuity generated by a school reform.

Two points are worthy of mention about the estimation results when moving to model (3). First, the permanent effect of grade retention (the coefficient of total years of delay) switches sign, from negative to positive. Hence, when we do not take into account that pupils might react differently to retention by abilities, like cognitive skills, intelligence and commitment, the permanent effect of retention is underestimated. Second, the interaction between the unobserved heterogeneity component and total years of delay is significantly negative: if the unobserved component is small enough ($v_{ev} << 0$), for instance if the student is very smart,

Table 4.6: Estimation Results of the Evaluation Equation

	Without unobserved heterogeneity (1)		Time-invariant unobserved heterogeneity (2)		Time-variant unobserved heterogeneity (3)	
Exogenous variables						
Female	0.423***	(0.045)	0.504***	(0.055)	0.467***	(0.054)
Cohort 1980	−0.055	(0.045)	−0.083	(0.053)	−0.103**	(0.052)
Calendar day of birth/100	0.026	(0.022)	−0.013	(0.026)	−0.012	(0.026)
Father's education/10	0.215***	(0.083)	0.473***	(0.103)	0.432***	(0.102)
Mother's education/10	0.200***	(0.088)	0.481***	(0.110)	0.423***	(0.108)
Number of siblings – Reference: No siblings						
1 sibling	−0.015	(0.069)	−0.036	(0.081)	−0.035	(0.081)
2 siblings	−0.143*	(0.074)	−0.212**	(0.088)	−0.208**	(0.089)
3 or more	−0.161*	(0.085)	−0.233**	(0.102)	−0.231**	(0.101)
Grade – Reference: Grade 1						
Grade 2	−1.747***	(0.127)	−1.532***	(0.138)	−1.529***	(0.137)
Grade 3	−1.597***	(0.135)	−1.414***	(0.144)	−1.413***	(0.144)
Grade 4	−1.654***	(0.129)	−1.537***	(0.137)	−1.523***	(0.138)
Grade 5	−1.494***	(0.134)	−1.438***	(0.144)	−1.414***	(0.144)
Grade 6	−1.225***	(0.222)	−1.255***	(0.240)	−1.186***	(0.238)
Endogenous variables						
Track – Reference: BSO						
ASO+	0.964***	(0.126)	0.167	(0.186)	0.298	(0.187)
ASO−	−0.541***	(0.080)	−1.150***	(0.126)	−1.058***	(0.122)
TSO+	−0.681***	(0.093)	−1.063***	(0.122)	−1.028***	(0.120)
TSO−	−0.907***	(0.078)	−1.187***	(0.101)	−1.174***	(0.100)
Total years of delay	−0.522***	(0.055)	−0.476***	(0.076)	0.370**	(0.146)
Lagged retention	1.369***	(0.138)	1.419***	(0.155)	1.033***	(0.367)
Lag B if retained last year	0.866***	(0.303)	0.782**	(0.311)	0.767**	(0.330)
Lag A if not retained last year	0.958***	(0.090)	0.675***	(0.105)	0.662***	(0.106)
Downgrade at the end of previous year – Reference: No downgrade						
1-step downgrade	0.597***	(0.098)	0.443***	(0.106)	0.474***	(0.106)
2-step downgrade	0.916***	(0.170)	0.648***	(0.179)	0.713***	(0.182)
Unobserved heterogeneity						
Unobserved heterogeneity support points (v_1 is normalised to 0)						
v_2			1.876***	(0.273)	1.632***	(0.221)
v_3			0.401***	(0.173)	0.245*	(0.140)
v_4			−0.111	(0.105)	−0.088	(0.070)
v_5			−0.876***	(0.192)	−1.113***	(0.211)
v_6			−1.330***	(0.239)	−1.295***	(0.249)
Interactions of unobserved heterogeneity support points with retention variables						
Interaction with total years of delay					−0.279***	(0.040)
Interaction with lagged retention					0.114	(0.114)

Note. Standard errors in parentheses.

grade retention generates a net negative permanent effect. Less able pupils ($v_{ev} > 0$) will instead be permanently, as well as momentarily, favoured by grade retention. The fact that more able students might be permanently penalised by an episode of grade retention suggests that psychological costs dominate possible benefits. For less able pupils the psychological costs might instead be dominated by the benefit of having more time to develop the knowledge and emotional maturity required at each educational grade. Also Fruehwirth et al. (2011) find that the retention effect varies by the abilities of pupils retained in kindergarten in the US. However, in contrast to our results, they find that lower able pupils are more negatively affected by grade retention. We conclude that when assessing the effectiveness of grade retention, one should carefully consider the heterogeneity in responses to grade retention by unobservable behavioural and cognitive abilities.

In the top track (ASO+) and in the vocational track (BSO), it is easier to get top evaluations. Pupils downgrading track are more likely to get good evaluations in the next academic year. Although there might be negative effects induced by changing peers and sometimes school, students who come from a higher track are likely to have an excess of knowledge relatively to the new track-level requirements, so they succeed more easily.

About the impact of exogenous regressors, the results are as expected. Girls perform better than boys, as it is generally found in the educational literature.[13] Parents' education is positively associated to the probability of getting an A. Pupils in larger families are more likely to perform worse.

4.5.4 Resitting Decision for B Students

Students getting a B can choose either to resit or to downgrade the track. The estimation results of the resitting equation for B students are reported in Table 4.7. Few regressors are significant. The higher the education of the father, the higher the probability that the pupil will prefer to resit instead of downgrading the track. The social status of the father of the pupil is therefore not only a determinant of schooling success, but also of resitting/downgrading choices. The lagged retention indicator has a significantly negative impact on the probability of choosing retention. This means that retained students who get a B are more likely to downgrade than to resit again compared to non-retained students. The cost of losing an academic year seems therefore to be increasing with the number of times students resit the same grade.

[13] See for instance the results in Van Houtte (2004) for Flanders.

Table 4.7: Estimation Results of the Resitting Decision for B Students

	Without unobserved heterogeneity (1)		**Time-invariant unobserved heterogeneity (2)**		**Time-variant unobserved heterogeneity (3)**	
Exogenous variables						
Female	−0.135	(0.164)	−0.147	(0.168)	−0.143	(0.167)
Cohort 1980	0.024	(0.167)	0.030	(0.169)	0.028	(0.169)
Calendar day of birth/100	0.123	(0.077)	0.128	(0.080)	0.127	(0.080)
Father's education/10	0.716**	(0.298)	0.701**	(0.321)	0.707**	(0.322)
Mother's education/10	0.160	(0.319)	0.145	(0.331)	0.153	(0.332)
1 or more siblings	0.055	(0.259)	0.053	(0.260)	0.054	(0.262)
Grade – Reference: Grade 2						
Grade 3	0.573**	(0.227)	0.544**	(0.229)	0.542**	(0.232)
Grade 4	0.722***	(0.196)	0.712***	(0.208)	0.712***	(0.210)
Grade 5	0.969	(1.016)	0.776	(1.030)	0.775	(1.077)
Endogenous variables						
Track – Reference: BSO						
ASO+	−0.556	(0.514)	−0.628	(0.649)	−0.656	(0.656)
ASO−	0.432	(0.322)	0.269	(0.359)	0.269	(0.359)
TSO+	0.992***	(0.356)	0.805**	(0.368)	0.810**	(0.369)
TSO−	1.144***	(0.326)	0.955***	(0.336)	0.963***	(0.336)
Total years of delay	−0.384	(0.270)	−0.409	(0.281)	−0.278	(0.322)
Lagged retention	−0.957	(0.668)	−0.984	(0.687)	−1.638**	(0.753)
Downgrade in the previous year	0.076	(0.323)	0.080	(0.327)	0.070	(0.332)
Unobserved heterogeneity loading factor						
Loading factor			0.281	(0.761)	−0.193	(0.806)

Note. Standard errors in parentheses. Due to the small number of students getting a B, there is not enough variation to be finer in distinguishing between number of siblings and the number of downgrade steps at the end of the previous academic year. For the same reason, we could not identify the effect of the interactions between lagged retention and lagged evaluation. Nobody resits grade 1 and therefore the reference category is grade 2.

4.5.5 Track Downgrade

The estimation results of the ordered logit model for track downgrade are reported in Table 4.8. A positive coefficient means that the corresponding regressor has a positive impact on the probability of making a two-step downgrade and a negative impact on the probability of remaining in the same track.

Parents' education has a negative effect on track downgrade. This evidence, jointly with the finding that the higher the education of the father the higher the probability that students getting a B will prefer to resit instead of downgrading, are in line with the predictions of sociological theories claiming that educational choices are influenced by social status maintenance and structural risk aversion. In a society where education is an investment good for social status upgrade, more advantaged families (higher educated parents) might have a greater incentive to invest in their children's education in order to preserve their advantage (Thurow, 1972). Moreover, higher education might become a social norm which children are persuaded to follow under the pressure of their family and peers (Boudon, 1974). Although also families in less advantaged class positions might invest in their children's education to give them a chance to raise their social and economic position, the failure of getting an education degree for a student from a lower educated family is likely to have more serious consequences than those for a student from families with larger resources (Breen and Goldthorpe, 1997; Goldthorpe, 1996). Since easier and/or shorter educational tracks minimise the risk of failing and entering the labour market without any (vocational) degree (Hartlaub and Schneider, 2012), it is not surprising to find students with lower educated parents to be more likely to choose the vocational track and to prefer downgrading to resitting.

The likelihood of downgrading is the highest in ASO+ and the lowest in TSO−. Furthermore, the evaluation obtained at the end of the current academic year is a strong determinant of downgrading. As expected, students getting an A are less likely to downgrade than students getting a B and, above all, a C. Students who have experienced a track change are less likely to downgrade in the following year, meaning that downgrading stabilises the track pathways of students. Finally, the total years of delay and ending the year without the need to resit the next one positively affect the probability of track downgrade.

4.5.6 School Drop-Out Without Diploma

Table 4.9 reports the estimation results of the drop-out equation. First, girls and younger students are less likely to drop out of secondary school without a diploma.[14] Second, the

[14]Eide and Showalter (2001) find the same gender difference in drop-out rates in the US.

Table 4.8: Estimation Results of the Track Downgrade for ASO/TSO Students

	Without unobserved heterogeneity (1)		Time-invariant unobserved heterogeneity (2)		Time-variant unobserved heterogeneity (3)	
Exogenous variables						
Female	0.035	(0.056)	−0.004	(0.058)	0.002	(0.058)
Cohort 1980	−0.041	(0.056)	−0.026	(0.056)	−0.024	(0.056)
Calendar day of birth/100	−0.003	(0.027)	0.019	(0.028)	0.019	(0.027)
Father's education/10	−0.600∗∗∗	(0.105)	−0.757∗∗∗	(0.116)	−0.748∗∗∗	(0.116)
Mother's education/10	−0.416∗∗∗	(0.114)	−0.578∗∗∗	(0.125)	−0.568∗∗∗	(0.124)
Number of siblings – Reference: No siblings						
1 sibling	−0.099	(0.082)	−0.091	(0.084)	−0.090	(0.083)
2 siblings	−0.161∗	(0.091)	−0.141	(0.092)	−0.140	(0.092)
3 or more	−0.09	(0.106)	−0.063	(0.109)	−0.066	(0.109)
Grade – Reference: Grade 1						
Grade 2	2.544∗∗∗	(0.131)	2.454∗∗∗	(0.133)	2.464∗∗∗	(0.133)
Grade 3	1.318∗∗∗	(0.136)	1.237∗∗∗	(0.137)	1.243∗∗∗	(0.138)
Grade 4	1.851∗∗∗	(0.131)	1.814∗∗∗	(0.132)	1.818∗∗∗	(0.133)
Endogenous variables						
Track – Reference: TSO−						
ASO+	1.193∗∗∗	(0.104)	1.747∗∗∗	(0.176)	1.712∗∗∗	(0.175)
ASO−	0.369∗∗∗	(0.093)	0.641∗∗∗	(0.108)	0.608∗∗∗	(0.107)
TSO+	0.674∗∗∗	(0.105)	0.799∗∗∗	(0.109)	0.778∗∗∗	(0.109)
Current evaluation – Reference: C						
A	−4.297∗∗∗	(0.197)	−4.292∗∗∗	(0.200)	−4.299∗∗∗	(0.201)
B	−0.721∗∗∗	(0.183)	−0.727∗∗∗	(0.187)	−0.732∗∗∗	(0.187)
Total years of delay	0.233∗∗∗	(0.090)	0.209∗∗	(0.092)	0.310∗∗	(0.158)
Lagged retention	0.414∗∗	(0.198)	0.445∗∗	(0.203)	−0.121	(0.346)
No current retention	2.216∗∗∗	(0.182)	2.232∗∗∗	(0.186)	2.232∗∗∗	(0.186)
Lag B if retained last year	−0.541∗	(0.295)	−0.545∗	(0.299)	−0.534∗	(0.304)
Lag A if not retained last year	0.105	(0.125)	0.119	(0.127)	0.127	(0.127)
Downgrade at the end of previous year – Reference: No downgrade						
1-step downgrade	−0.428∗∗	(0.167)	−0.290∗	(0.170)	−0.292∗	(0.171)
2-step downgrade	−1.117∗∗∗	(0.367)	−0.940∗∗	(0.385)	−0.950∗∗	(0.388)
Unobserved heterogeneity						
Unobserved heterogeneity support points (v_1 is normalised to 0)						
v_2			−1.146∗∗∗	(0.381)	−1.073∗∗∗	(0.368)
v_3			−1.019∗∗∗	(0.343)	−0.990∗∗∗	(0.344)
v_4			−0.917∗∗∗	(0.314)	−0.889∗∗∗	(0.316)
v_5			−0.894∗∗∗	(0.308)	−0.880∗∗∗	(0.313)
v_6			−0.478∗∗∗	(0.207)	−0.464∗∗	(0.210)
Interactions of unobserved heterogeneity support points with retention variables						
Interaction with total years of delay					0.041	(0.049)
Interaction with lagged retention					−0.237∗∗	(0.102)

Note. Standard errors in parentheses.

Table 4.9: Estimation Results of the Drop-Out Equation

	Without unobserved heterogeneity (1)		**Time-invariant unobserved heterogeneity (2)**		**Time-variant unobserved heterogeneity (3)**	
Exogenous variables						
Female	−0.183	(0.144)	−0.489∗∗	(0.200)	−0.366∗∗	(0.183)
Cohort 1980	0.088	(0.138)	0.103	(0.178)	0.156	(0.169)
Calendar day of birth/100	−0.218∗∗∗	(0.067)	−0.237∗∗∗	(0.085)	−0.218∗∗∗	(0.082)
Father's education/10	−0.454∗	(0.270)	−1.008∗∗∗	(0.365)	−0.917∗∗∗	(0.348)
Mother's education/10	−0.158	(0.281)	−0.830∗∗	(0.375)	−0.725∗∗	(0.348)
1 or more siblings	0.000	(0.194)	0.075	(0.260)	0.010	(0.241)
Final grade	−3.267∗∗∗	(0.288)	−2.353∗∗∗	(0.343)	−2.748∗∗∗	(0.353)
Endogenous variables						
BSO	2.138∗∗∗	(0.248)	2.100∗∗∗	(0.349)	1.992∗∗∗	(0.300)
Current evaluation – Reference: C						
A	−1.450∗∗∗	(0.228)	−0.738∗∗	(0.300)	−1.079∗∗∗	(0.281)
B	0.215	(0.543)	0.575	(0.637)	0.284	(0.645)
Total years of delay	−0.036	(0.106)	−0.214	(0.140)	−0.211	(0.137)
Lagged retention	0.231	(0.236)	−0.123	(0.283)	−0.000	(0.280)
Downgrade in the previous year	−0.560∗	(0.320)	−0.555	(0.407)	−0.541	(0.394)
Unobserved heterogeneity loading factor						
Loading factor			−18.430∗∗	(7.604)	−16.696∗∗	(6.931)

Note. Standard errors in parentheses. As students can drop-out of school without the diploma only when they turn 18 years old, the sample at risk of exit is small and there is not enough variation to distinguish between different tracks, grades and the number of downgrade steps at the end of the previous academic year. For the same reason, we could not identify the effect of the interactions between lagged retention and lagged evaluation.

higher parents' education, the lower the propensity to drop-out. Third, students reaching the final grade, therefore closer to the target, or getting an A are less likely to drop out without the diploma. Last, BSO students have a significantly higher probability of not completing secondary school. This finding might be explained by the fact that BSO students' opportunity cost of not getting the diploma might be lower than the one of ASO/TSO students for at least two reasons. First, students in vocational tracks might access the labour market without the diploma in specialised/blue collar jobs more easily than similar ASO/TSO students because of the specific human capital they acquired in the BSO track. Second, BSO students might be less interested in enrolling in tertiary education.

4.5.7 Secondary School Graduation

Table 4.10 reports the estimation results of the diploma equation, i.e. the impact of covariates on the probability of getting the diploma once students make it to the last grade of their

Table 4.10: Estimation Results of the Diploma Equation

	Without unobserved heterogeneity (1)		**Time-invariant unobserved heterogeneity (2)**		**Time-variant unobserved heterogeneity (3)**	
Exogenous variables						
Female	0.724***	(0.182)	2.023***	(0.642)	2.794***	(0.858)
Cohort 1980	0.034	(0.173)	0.027	(0.345)	−0.652	(0.456)
Calendar day of birth/100	0.166**	(0.084)	0.115	(0.169)	0.018	(0.210)
Father's education/10	0.079	(0.340)	1.087	(0.764)	1.528	(0.996)
Mother's education/10	−0.144	(0.353)	1.134*	(0.684)	2.007**	(0.943)
Number of siblings – Reference: No siblings						
1 sibling	0.162	(0.253)	−0.140	(0.590)	−0.639	(0.810)
2 siblings	−0.003	(0.284)	−0.723	(0.659)	−1.509	(0.978)
3 or more	−0.263	(0.311)	−1.323*	(0.737)	−2.348**	(1.159)
Endogenous variables						
Track – Reference: BSO						
ASO+	1.420***	(0.385)	0.787	(0.818)	−0.171	(1.301)
ASO−	0.873***	(0.283)	0.334	(0.659)	−0.151	(1.126)
TSO+	−0.139	(0.313)	0.712	(0.649)	−2.285*	(1.318)
TSO−	0.231	(0.279)	0.611	(0.687)	−0.433	(1.076)
Total years of delay	−0.441***	(0.138)	0.150	(0.314)	0.647	(0.542)
Lagged retention	0.042	(0.431)	3.241***	(1.103)	7.170	(6.588)
Unobserved heterogeneity loading factor						
Loading factor			−9.422**	(3.999)	−23.152*	14.000

Note. Standard errors in parentheses. As students are not allowed to change track at the beginning of the last grade, there is no control for lagged downgrade decision in the diploma equation.

track.[15] Once again we find that girls and pupils in smaller families perform better and that they are significantly more likely to get the diploma. Parents's education is positively correlated to the probability of getting the diploma, although only the impact of mother's education is significantly different from zero. Finally, once students are in the last grade, the transitory and the permanent retention effects are positive but not significantly different from zero.

[15] As mentioned before, the last grade of ASO/TSO tracks is the 6th grade, while the last grade of the BSO track is the 7th grade.

4.6 Conclusions

We empirically analysed the short-term and permanent effect of grade retention on later success rates in school. We exploited econometric modelling tools and identification analysis to examine the interrelated dynamics of secondary school grade retention, school track choices and achievements of a sample of Belgian pupils. We also shed light on the role played by family background and unobserved abilities, especially looking at how unobserved abilities interact with retention episodes in determining schooling pathways.

The empirical analysis was based on the rich schooling information contained in the SONAR dataset, a retrospective survey conducted in Flanders on the 1976, 1978 and 1980 cohorts. Our sample was made up of 4,214 students belonging to the 1978 and 1980 cohorts. We exploited the ample information on secondary school performances and choices to estimate dynamic qualitative choice models.

In contrast to most of the previous findings, we found that grade retention has a positive impact on the next evaluation and can permanently affect subsequent performances. The direction of the permanent effect depends on unobserved heterogeneity. While more able students are permanently penalised by retention, less able students benefit from it. The fact that more able students might be permanently penalised by an episode of grade retention suggests that psychological costs dominate possible benefits. For less able pupils the psychological costs might instead be dominated by the benefit of having more time to develop the knowledge and emotional maturity required at each educational grade. We conclude that when assessing the effectiveness of grade retention, one should carefully consider the heterogeneity in responses to grade retention by unobservable behavioural and cognitive abilities.

5

Overeducation at the Start of the Career: Stepping Stone or Trap?

This chapter is joint work with Prof. Dr. Bart Cockx (Ghent University, Université catholique de Louvain, CESifo and IZA) and Prof. Dr. Dieter Verhaest (HUBrussel and Ghent University).

5.1 Introduction

Numerous studies have shown that many young workers are overeducated at the start of their career (see, e.g., Battu et al., 1999; Dolton and Vignoles, 2000). A worker is considered to be overeducated if her/his education level is higher than the level that is typically required to perform adequately (McGuinness, 2006). This phenomenon suggests a less-than-optimal allocation of graduates over jobs and is potentially costly for society (Groot and Maassen van den Brink, 2000; McGuinness, 2006). For overeducated workers, this translates in lower earnings (Hartog, 2012; Korpi and Tåhlin, 2009) and lower job satisfaction (Tsang, 1987; Allen and van der Velden, 2001). Therefore, one might wonder why young job seekers actually accept jobs with requirements below their educational attainment. One potential answer is that by accepting such positions these job seekers avoid scarring effects of staying unemployed (see, e.g., Arulampalam, 2001). Another (additional) explanation is that this might be the shortest pathway to a job that matches the attained educational level. This stepping stone hypothesis has been formulated most clearly by Sicherman and Galor (1990). According to their career mobility theory, overeducation is an investment in work experience which enhances promotion opportunities to higher level positions inside or outside the firm.[1] However, overeducation might just as well retard the transition to an adequate job. Job search intensity on-the-job may decrease (Holzer, 1987) and job specific human capital investments may lock workers into bad positions (Pissarides, 1994). Further, the sources underlying unemployment scarring may equally apply to overeducation. According to McCormick (1990), overeducation even acts as a stronger negative signal to employers than unemployment and de Grip et al. (2008) show that overeducation also results in cognitive decline.

Several empirical studies have already provided interesting insights into this debate by investigating the mobility behaviour of overeducated workers. Sicherman (1991), Robst (1995) and Rubb (2006), for instance, find for the US that overeducated workers are more likely to move to occupations with higher human capital requirements than adequately educated workers with similar educational backgrounds. This is consistent with the career mobility thesis. In addition, Rubb (2003) reports a yearly transition rate from overeducation to adequate employment of about 20%, suggesting that overeducation is a temporary problem for most US workers. Finally, relying on data for a large Dutch firm, Groeneveld and Hartog (2004) find some evidence that overeducated workers experience more internal promotions,[2] suggesting that overeducation may foster career mobility. However, a number of studies for other countries challenge this conclusion. Battu et al. (1999) find that the match between the educational degree and the job requirements remains fairly stable around 60% 1, 6 and 11 years after graduation for two cohorts who graduated from higher education in the UK. Dolton and Vignoles (2000) arrive at similar conclusions. Bauer (2002) finds, using the German GSOEP data from 1984 to 1998, that relatively few employees change their mismatch status. This is confirmed by Büchel and Mertens (2004) who report that overeducation results in less upward occupational mobility and less wage growth in the German labour market. This is especially the case for young workers with low-quality education (Pollmann-Schult and Büchel, 2004). More recently, Verhaest and van der Velden (2013) studied the persistence of overeducation in 14 countries. They find substantial heterogeneity in this persistence both

[1] Within such a context, the term of overeducation may sound confusing as there is no overinvestment in education if the whole career is taken into account. However, we follow the literature and conceptualise overeducation rather as a situation of underutilisation of education given one's job at a particular point in time (see McGuinness (2006), and the aforementioned definition).

[2] This evidence was found for the internally oriented work units of the firm, but not for the more externally oriented ones.

across countries and within countries according to the quality and orientation of their human capital. Finally, Mavromaras and McGuinness (2012) estimate a dynamic random effects probit model allowing for correlated unobserved heterogeneity on Australian data. They find substantial state dependence in overskilling[3] for workers with a high educational degree, but none for workers with vocational education. Based on a similar model and consistent with the aforementioned findings, Mavromaras et al. (2013) report that neither overeducation nor overskilling has any significant effect on job mobility of female university graduates in Australia. By contrast, overeducation, especially in combination with overskilling, positively affects (voluntary) quits, but not (involuntary) layoffs of male graduates. Whether this leads to more upward occupational mobility remains an open question, however.

From this overview of the literature we conclude that evidence for the career mobility theory is mixed. Moreover, most researchers just study the persistence of overeducation. However, even if overeducation is persistent, this does not necessarily mean that one slows down the transition to an adequate job. This transition may still be accelerated. After all, an individual who does not accept a job for which he is overeducated may remain unemployed and therefore without adequate job even longer than when he accepts such a job. In other words, overeducation is not necessarily a "trap". It may still be a "stepping stone" to an adequate job. This chapter studies the stepping stone hypothesis.

An answer to this research question is not only interesting from a theoretical perspective, but also from a policy point of view. By analysing the stepping stone hypothesis for unemployed youth, this paper provides more insight in the strategy that policy makers should follow in fighting youth unemployment, currently one of main priorities of the European Union (European Foundation for the Improvement of Living and Working Conditions, 2012). For, if overeducation is a stepping stone for young unemployed graduates, then the policy maker has an interest to encourage or enforce acceptance of job offers to this group as early as possible, irrespectively of whether the educational attainment required in the job matches or is below that of the job candidate. By contrast, if overeducation is a trap, then the policy maker faces a trade-off. In that case the benefits of the shorter unemployment spell induced by accepting a job for which one is overeducated, should be weighed against the losses of the delayed entry in an adequate job.

Our analysis also innovates in the overeducation literature from a methodological point of view. As pointed out by Leuven and Oosterbeek (2011), many studies on overeducation and its consequences fail to account for possible non-random selection into overeducation.[4]

[3]Overskilling is a situation whereby an individual has more skills and knowledge than those utilised in the job (see, e.g., Allen and van der Velden, 2001; Green and McIntosh, 2007)

[4]The aforementioned articles of Mavromaras and McGuinness (2012) and Mavromaras et al. (2013) are

For instance, overeducated workers may be less able than adequately educated workers. If so, the comparison of the outcomes between adequately qualified and overeducated workers may partly reflect this ability difference and is therefore not causal. To solve this selection problem, we apply the Timing of Events approach as developed by Abbring and van den Berg (2003). To identify the selection from the treatment effect, this method exploits that unobserved time-constant individual determinants of the transition to an adequate job affect this transition throughout the period that one is searching for an adequate job whereas the treatment (transition into overeducation) may only influence this transition as from the moment at which the treatment occurs. The selection effect can therefore be identified from the pre-treatment data if the treatment is not anticipated and the timing of the treatment is random, even without any exclusion restrictions.

A second methodological critique on the overeducation literature by Leuven and Oosterbeek (2011) concerns the error in the measurement of overeducation. Even if we do not directly address this critique, we indirectly deal with it by assessing the sensitivity of our results to two alternative measures of overeducation. One is based on a job analysis approach and another on a modified self-assessment method. Moreover, we argue in the text that the latter measure is close to a measure of "genuine" overeducation, as defined by Chevalier (2003). This is important, since Chevalier argues that genuinely overeducated workers are more likely to move to a higher level job than those who are apparently overeducated. Since our findings are not sensitive to the choice of these measures, we are quite confident that they are not driven by incorrect measurement of the overeducation.

The analysis is based on a retrospective survey of a representative sample of two birth cohorts, born in 1978 and 1980 and living in Flanders, the Dutch-speaking region in the North of Belgium. From this sample we retain male unemployed youth who started searching for a job after graduating from formal education. An advantage of analysing data right after graduation is that there is a closer connection between the concept of overeducation and overskilling, since individuals have not yet acquired any skills through experience on-the-job. Moreover, the unusual richness of the database sustains the credibility of our findings. On the one hand, it contains detailed information on the timing of labour market transitions: starting dates of job search, transitions from unemployment to employment and even job-to-job and position changes within a same firm. This is crucial for the application of the Timing of Events method and also to capture career mobility even if it occurs within the firm (Groeneveld and Hartog, 2004), ensuring that our study cannot be criticised on the grounds of underestimating career mobility.

rare exceptions with respect to the literature on the job mobility of overeducated workers.

The chapter is organised as follows. In the next section we summarise the institutional setting: the educational system and the youth labor market in Flanders. Section 5.3 describes the dataset in further detail and provides some selected descriptive statistics. Section 5.4 discusses the econometric framework. Section 5.5 contains and discusses the estimation results. Section 5.6 concludes.

5.2 The Transition from Education to Work in Flanders: Institutional Context

As many other countries, Belgium has experienced substantial youth labour market problems since the nineties (Blanchflower and Freeman, 2000). In 2005, for instance, youth unemployment rate was three times that of adults, compared to an OECD average ratio of 2.3 (OECD, 2007).[5] This high unemployment is particularly observed among new labour market entrants (Gangl, 2003) while relatively few youngsters experience repeated unemployment spells once they have found a job (Couppié and Mansuy, 2003). This poor performance at labour market entry is partly related to the strict employment protection legislation (EPL) especially for experienced white collar workers (OECD, 2007). Moreover, different from other countries, school graduates in Belgium can claim unemployment benefits (UB) after a so-called "waiting period" of nine months even if they did not acquire any work experience, and during the first six months they may refuse job offers that do not match their acquired skills in school without losing their entitlement to UB. However, in 2012 (beyond the observation period of the empirical analysis), the Belgian government has tightened UB eligibility requirements. The aforementioned waiting period has been prolonged to twelve months and the period during which inadequate job offers could be refused has been shortened to three months. This provides incentives to search more intensively for jobs and to be less selective in job acceptance behaviour, so to accept lower paid jobs and jobs for which one is overeducated. This may therefore deteriorate the quality of the job match in the short run, but also in the long run if these jobs are no stepping stone to adequate jobs, but a trap, a point that will be clarified in our empirical analysis.

Apart from labour market regulations, also educational institutions matter for the transition from education to work. As most other regions and countries, Flanders has experienced a substantial increase in the average level of education of the population over the past decades.[6]

[5] While there are substantial regional disparities in youth unemployment rates, with Flanders approaching the OECD average, all regions face a similarly high youth versus adult unemployment ratio.

[6] In Belgium, educational policy is a regional competence. The Flemish educational landscape is described in Chapter 4 of this thesis and in De Ro (2008).

This is at least partly attributed to the increase in the age of compulsory education age to 18 years since the beginning of the eighties. At that age, full-time students without grade retention have completed their higher secondary education (ISCED[7] 3 or 4). In Flanders, full-time secondary education is organised along four tracks (general, technical, arts and vocational) and lasts six years. Along with full-time education, also an apprenticeship track can be followed from the age of 15 onwards. Those with a full-time higher secondary education degree can afterwards choose to start a seventh specialisation grade or to enrol immediately, without any entry exam, into higher education.[8] Our data concern education registrations prior to adoption of the Bologna process. Three kinds of higher degrees could be obtained: (i) non-university of the “short type” (lasting 3 years), (ii) non-university of the “long type” and (iii) university education of the “long type” (4 years or more). Since the Bologna reform, programs of type (i) deliver Bachelor degrees (“Lower tertiary education”), while programs of type (ii) and (iii) deliver Master degrees (“Higher tertiary education”).

5.3 Data and Descriptive Statistics

5.3.1 The Sample of Analysis

Our analysis is based on data from a representative sample of two cohorts (birth years 1978 and 1980) of the SONAR survey conducted when respondents were 23 years old. These data are supplemented with data from two follow-up surveys, completed at age 26 for the 1978 cohort (response rate of 69%) and at age 29 for the 1980 cohort (response rate of 64%). Detailed information regarding the sampling procedures and general summary statistics can be found in SONAR (2000a) and SONAR (2005).

The SONAR data contain detailed information regarding school and labour market careers, which makes them very suitable for our analysis. The level of acquired educational attainment is measured at the moment that the youngster reports to have left formal education for the first time. The labour market history is registered on a monthly basis. Each month is assigned either to a working or to a non-working status,[9] depending on the status in which one spends most of the time. Further, if employed, both job-to-job transitions and position changes within a job are recorded. Part-time jobs held during vocational education, student and vacation jobs are defined to be part of the educational career.

The analysis targets workers who are unemployed right after graduation. We therefore

[7] ISCED stands for International Standard Classification of Education.

[8] Students from the vocational track are obliged to follow the seventh specialisation year to start tertiary education. Candidates for studying medicine must first pass an entry exam.

[9] Not working is simply defined as a residual category, meaning neither in work nor in education.

select individuals from the first moment they report to have started searching for a job since graduation.[10] Furthermore, in the benchmark model we only consider men, since career mobility of women could be influenced by fertility and by their traditionally higher responsibility in household activities, including child rearing. However, in a sensitivity analysis, we include women in the analysis and conclude that our findings are actually hardly affected. Finally, we exclude individuals who did not attain a degree of lower secondary education, because below this level of education no one is overeducated by definition. After eliminating 114 observations for which explanatory variables are missing, the final dataset contains 1,434 individuals. In the sensitivity analysis that includes women the sample size increases to 2,956.

5.3.2 Measures of Overeducation

In the main analysis we define overeducation according to a job analysis approach. Each position in the SONAR data has been coded following the Standard Occupation Classification of Statistics Netherlands (CBS, 2001). This classification groups jobs according to a set of tasks to be executed and assigns to each occupation the educational level that is the most appropriate. The following five functional levels are considered: less than lower secondary (ISCED 0 or 1), lower secondary (ISCED 2), higher secondary (ISCED 3 or 4), lower tertiary (ISCED 5–Bachelor) and higher tertiary (ISCED 5–Master) education. Hence, an individual is considered to be overeducated if the functional level of his job exceeds his attained educational level. Those with a functional level above their educational level, the so-called undereducated, are considered to be adequately educated in this study. Considering this –small –group as a separate category would further complicate the analysis. Moreover, undereducated individuals generally earn at least as much and are at least as satisfied with their jobs as adequately educated workers (Hartog, 2012; Verhaest and Omey, 2009). Hence, this justifies pooling them with the adequately educated.

Apart from job analysis, several other measurement approaches, for instance based on self-assessments, have been applied in the literature (see Leuven and Oosterbeek, 2011, for an overview). As any method, job analysis has some disadvantages. One often formulated criticism is that, within a particular occupation, there may be substantial heterogeneity in the tasks to be executed. However, since the SBC-classification is rather built upon tasks to be executed than on occupational titles, this problem should be less severe for our measure. Another criticism is that occupational or task classifications may be relatively inflexible to upgrades in educational requirements. Subjective measures or, better, measures of “genuine”

[10] We do not retain individuals who started searching for jobs during their studies, since for these individuals we cannot identify the moment at which they found a job (if this occurs before graduation). Without information on this moment the Timing of Events approach cannot be implemented.

overeducation à la Chevalier (2003) may perform better in this respect in that they discard "apparent" overeducation induced thereby in the job analysis methodology. On the other hand, Hartog (2012) argues that, as a consequence of a social desirability bias, job requirements may be overestimated in a self-assessment approach and the measure of overeducation therefore biased downwards. From this discussion follows that there is no agreement among researchers which measure of overeducation should be used. This justifies a sensitivity analysis.

Ideally, we would like to build a measure of "genuine" overeducation à la Chevalier (2003). However, this is not possible, since this requires a response on a question that gauges the extent to which the employee is utilising his skills and knowledge (Green and Zhu, 2010). The survey does contain a statement[11] to which the respondent should formulate the extent of agreement that could be used as a proxy for this question. However, this statement was only asked with respect to the first job occupied after graduation. To study the stepping stone hypothesis we also need this information for all occupied jobs until the first that is adequate. As a way out, we develop a "modified self-assessment method", the method of which is explained below, that is highly associated with the aforementioned measure of genuine overeducation. According to the modified self-assessment measure 34% of the individuals in our sample are overeducated in their first job after graduation compared to 33% according to the measure of genuine overeducation. The measures classify 75% of the jobs in the same way.

The modified self-assessment is defined in the following way. The SONAR survey included the following survey question regarding the first job: "What is (was), according to your own opinion, the most appropriate educational level to execute your job?" As this question was not included for subsequent jobs, we adopted the following construction procedure for our alternative measure. First, relying on this information on first jobs, we computed the median subjectively assessed required level within each occupation.[12] Second, we assessed both for the first job and for later jobs whether someone was overeducated or not by comparing his educational level with this computed median of the subjectively required level.

According to the job analysis method within the retained sample 59% of the first jobs were filled by overeducated workers. This is a substantial fraction. The fact that we choose to focus on a more disadvantaged group of unemployed graduates, as to address our policy question regarding youth unemployment, explains only part of this high fraction. The fraction of overeducated workers drops to 49% if we retain only graduates who directly transit from school to work. The high fraction seems therefore more related to the fact that we focus

[11]"I can show in my job what I am capable of doing".

[12]For this derivation, we rely on the full SONAR dataset. Nevertheless, the number of self-assessments is relatively low for some occupations at more detailed levels. Therefore, we base the computation on the most detailed occupational level for which we have at least 20 observations available in our data. For a similar procedure and discussion in the case of realised matches measures, see Verhaest and Omey (2010).

on young graduates. van Smoorenburg and van der Velden (2000) and Giret and Hatot (2001) report on the basis of job analysis approaches that 65% and 57% of the first jobs after graduation were filled by overeducated workers, respectively in the Netherlands and France.

The degree of overeducation according to the modified self-assessment method is 34% and is therefore much lower than according to the job analysis approach. This is again consistent with the literature, in particular if we focus on the degree of overeducation in the first job after graduation. Subjective measures of overeducation range for this group between 23% in the Netherlands (van Smoorenburg and van der Velden, 2000), 40% in the UK (Dolton and Vignoles, 2000) and 48% in France (Giret and Hatot, 2001). In a study analysing 14 countries, Verhaest and van der Velden (2013) find that the median incidence of self-assessed overeducation is 24% in the job that is occupied six months after graduating from university. For Flanders, the region that we study here, 27% of these university graduates were overeducated according to this method.

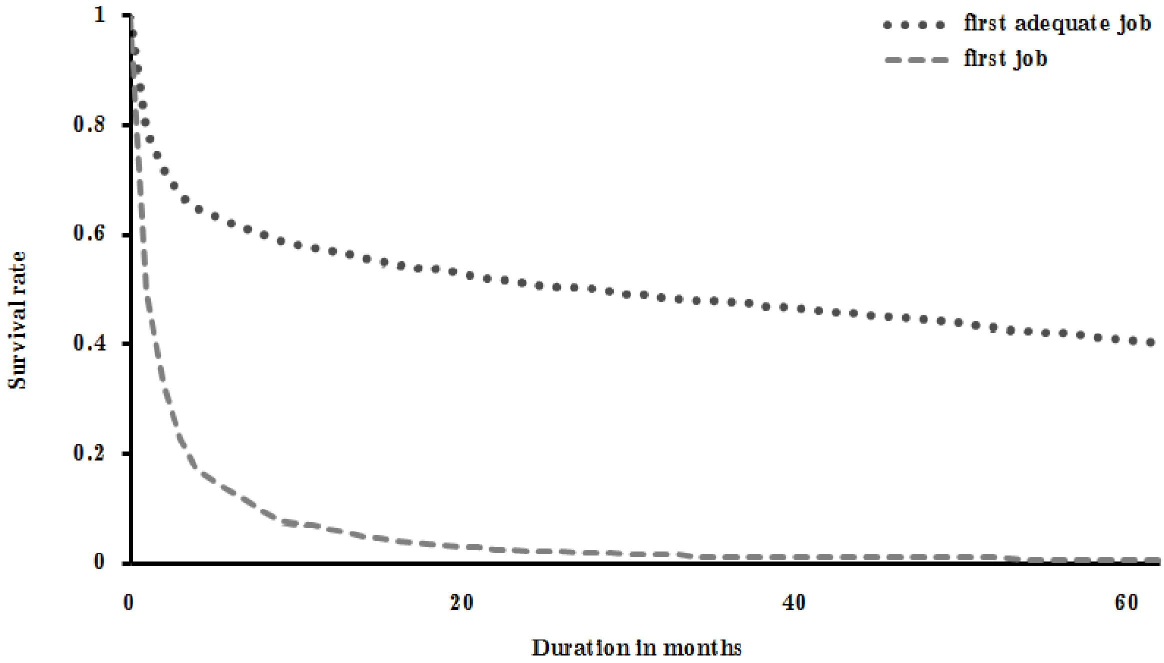
5.3.3 Descriptive Analysis

Based on the aforementioned information, we determine for each sampled individual the timing at which he entered an adequate job or a job for which he was overeducated since he started searching for a job. 788 young men (55% of the sample) find an adequate job (as measured by job analysis) before the end of the observation period. 546 (38%) of these men directly enter an adequate job (subsample 'E'), while 242 (17%) are temporarily overeducated before entering the adequate job (subsample 'OE'). 549 young men (38%) enter a job for which they are overeducated and do not subsequently transit to an adequate job before they are right censored (subsample 'OC'). 97 (7%) individuals are right censored before making any transition (subsample 'C'). The treatment group consists of those individuals who enter a job for which they are overeducated; subsamples 'OE' and 'OC'. The control group consists of the same individuals until the moment they enter overeducation, plus those individuals who are never overeducated; subsample 'E' (until entry in an adequate job) and 'C'.

The right-censoring for 646 ($= 549 + 97$) of the observations occurs for one of the following reasons: (i) end of the observation period and sample attrition (69% of the 646 right censored observations); (ii) transition to a job for which the functional level is not registered (15%); (iii) return to full time education (7%); (iv) transition to self-employment (7%) and (v) transition to disability (2%).

Figure 5.1 reports non-parametric Kaplan-Meier estimates of the monthly transition into a first job (irrespectively of whether one is overeducated for it or not) and into a first adequate job (directly or indirectly, so after a temporary spell of overeducation). Overeducation is

Figure 5.1: Kaplan-Meier Estimates: Duration Between Start Job Search and Entry in the First (Adequate) Job (Job Analysis Method)

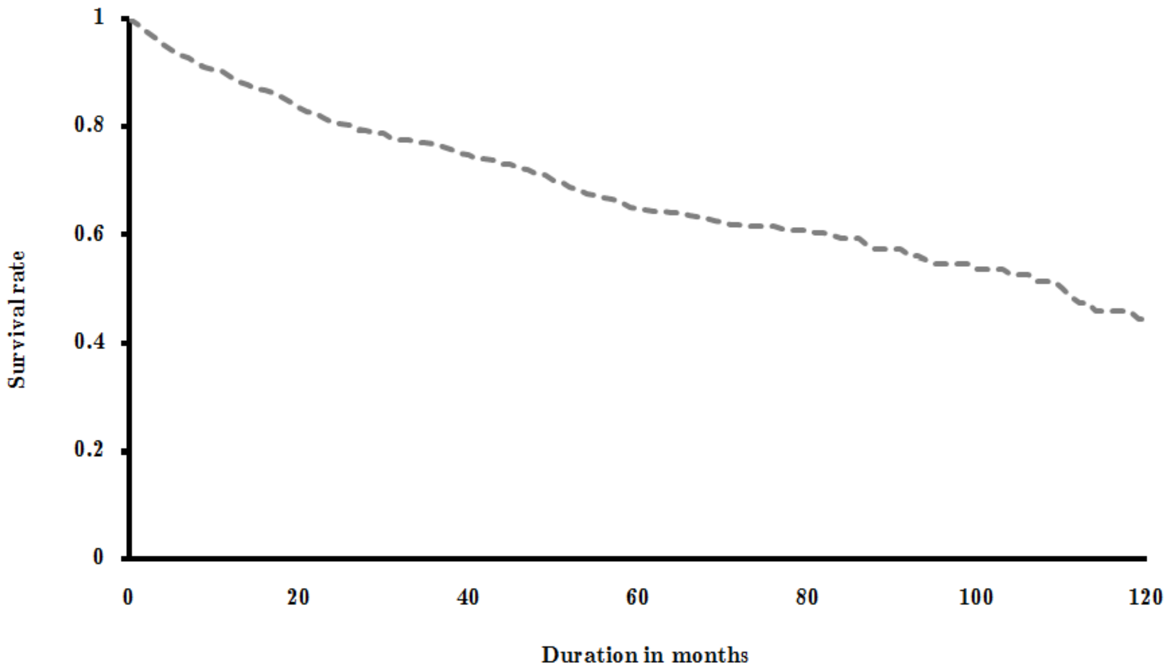


measured according to the job analysis method.

The corresponding figure if overeducation is measured by the self-assessment method is reported in the Appendix of this chapter. The median duration until the transition to a first job is 1 month and 29 months until a first transition to an adequate job. This illustrates clearly that most young graduates very rapidly find a job at the start of their career, but also that these graduates are overeducated for most of these jobs, since they enter an adequate job at a much slower rate.

Figure 5.2 reports the non-parametric Kaplan-Meier estimates of the number of months that elapse after accepting a job for which one is overeducated until entry in an adequate job. The median duration is as high as 110 months. As the median duration since the start of job search until (direct or indirect) entry into an adequate job is only 29 months, this means that most direct transitions into an adequate job occur much more rapidly than the indirect transitions. However, since this comparison does not take selection on (un)observable characteristics into account, we cannot conclude from this descriptive evidence, that accepting a job for which one is overeducated is a trap rather than a stepping stone to an adequate job. Overeducated individuals might have very low chances to enter adequate jobs anyway, so that for these individuals it might have taken even longer before they would have found an

Figure 5.2: Kaplan-Meier Estimates: Duration Between Start Overeducated Employment and Entry in the First Adequate Job (Job Analysis Method)



adequate job if they would have rejected all jobs for which they were overeducated. The Timing of Events method that we apply in this research takes the selection on (un)observable characteristics into account and leads therefore to a better founded answer to our research question.

Our analysis controls for a rich set of observed characteristics. A vector of time-constant variables measured before the start of the job search spell captures the respondents' (i) level of educational attainment (highest attained level of education, number of uncertified years of schooling beyond the highest level and an indicator of whether one obtained an additional degree at the same level of education), (ii) school achievement (years of schooling delay at the age of 16 and the grade obtained in tertiary education),[13] (iii) school orientation in secondary school (general, technical, vocational or arts), (iv) social background (mother's and father's level of education and migrant status as captured by the nationality of the grandmother at mother's side), (v) birth cohort (1978 or 1980), (vi) work experience during school (internship or student job) and (vii) timing of the start of job search (quarter in the year and number of months since leaving school). In addition, the monthly Belgian youth unemployment rate (ILO definition) is included as a time varying variable as to capture seasonal and business

[13]This information is not available at the level of secondary education.

cycle variation. Details concerning some of these variables are outlined in the notes of Table 5.1.

In Table 5.1 we report descriptive statistics for these explanatory variables used in the econometric analysis below. We separately report statistics on the four subsamples identified in the beginning of this subsection: 'E', 'OE', 'OC' and 'C'. Subsample C is relatively small and contains the most educated graduates, since, as this group studies longer, the observation period is systematically shorter for this group. On the other hand, it comprises more foreign youth, which squares with the well documented negative correlation between foreign ethnicity and labour market success. There are also fewer individuals in this subsample with any internship or student work experience during their educational careers.

There is no clear pattern in the differences observed between the other three subsamples except that in subsample E, which is restricted to youth with a direct transition to adequate employment, the parent's level of education and the number of years of schooling delay at the age of 16 are on average higher and the youth unemployment rate is lower than in the two other subsamples, containing youth who are (first) overeducated. It is a priori unclear in which direction this could have biased the aforementioned descriptive evidence on our research question.

5.4 Econometric Model

5.4.1 The Selection Problem

We aim at identifying whether an unemployed graduate can accelerate the transition to an adequate job by temporarily accepting a job for which he is overeducated ("the treatment") rather than only accepting adequate jobs, or whether instead he might get trapped in overeducation by following such a strategy. To answer this question, we face a double selection problem. First, young men who are more likely to accept a job for which they are overeducated may have a systematically lower (or higher) likelihood of finding an adequate job than those who are less likely to be overeducated. If we ignore this 'classic selection problem', then a simple comparison of the speed of transition to an adequate job between those who directly enter an adequate and those who do so only after an intermediate period of employment as overeducated worker, will underestimate (overestimate) the treatment effect on this speed. Second, even if there is no systematic relationship between the unobserved determinants of treatment and entry in an adequate job, then we are still confronted with a "dynamic selection problem". Since the treatment does not occur at the start of the unemployment spell, treatment can only occur for youth who did not find an adequate job beforehand. Consequently, the treatment

Table 5.1: Summary Statistics

Sample:	**Total**	**C**	**E**	**OC**	**OE**
Number of individuals:	(1434)	(97)	(546)	(549)	(242)
Highest level of educational attainment					
Highest attained level of education					
Lower secondary education	0.24 (0.43)	0.14 (0.35)	0.27 (0.44)	0.24 (0.43)	0.21 (0.41)
Secondary education	0.43 (0.50)	0.30 (0.46)	0.44 (0.50)	0.41 (0.49)	0.52 (0.50)
Lower tertiary education	0.17 (0.38)	0.22 (0.41)	0.19 (0.39)	0.14 (0.35)	0.17 (0.38)
Higher tertiary education	0.16 (0.37)	0.34 (0.48)	0.10 (0.31)	0.21 (0.41)	0.10 (0.30)
Additional successful years of education after highest attained level of education	0.40 (0.73)	0.32 (0.65)	0.37 (0.67)	0.48 (0.81)	0.36 (0.69)
Additional degree at highest attained level of education	0.06 (0.23)	0.11 (0.32)	0.04 (0.19)	0.07 (0.25)	0.05 (0.23)
Schooling achievement					
Tertiary education: grade					
No grade	0.85 (0.36)	0.77 (0.42)	0.85 (0.36)	0.86 (0.35)	0.87 (0.34)
With honours (cum laude)	0.13 (0.34)	0.21 (0.41)	0.13 (0.34)	0.12 (0.33)	0.12 (0.32)
With great/highest honours (cum magna/maxima laude)	0.02 (0.14)	0.02 (0.14)	0.02 (0.14)	0.02 (0.14)	0.02 (0.14)
Years of schooling delay (at age of 16)	0.43 (0.60)	0.40 (0.64)	0.47 (0.65)	0.42 (0.57)	0.38 (0.54)
School orientation at age of 16 in secondary school					
General secondary education	0.35 (0.48)	0.57 (0.50)	0.33 (0.47)	0.35 (0.48)	0.30 (0.46)
Technical secondary education	0.37 (0.48)	0.27 (0.45)	0.40 (0.49)	0.34 (0.47)	0.43 (0.50)
Arts secondary education	0.01 (0.11)	0.02 (0.14)	0.01 (0.10)	0.01 (0.10)	0.02 (0.13)
Vocational secondary education	0.27 (0.44)	0.14 (0.35)	0.26 (0.44)	0.30 (0.46)	0.25 (0.43)
Social background					
Mother's educational level	5.33 (3.25)	5.86 (3.40)	5.47 (3.16)	5.13 (3.32)	5.24 (3.20)
Father's educational level	5.72 (3.41)	6.47 (3.60)	5.86 (3.25)	5.57 (3.50)	5.47 (3.42)
Grandmother (mother's side) foreign	0.06 (0.24)	0.13 (0.34)	0.07 (0.25)	0.04 (0.20)	0.06 (0.24)
Birth cohort (birth year)					
1978	0.48 (0.50)	0.42 (0.50)	0.49 (0.50)	0.50 (0.50)	0.41 (0.49)
1980	0.52 (0.50)	0.58 (0.50)	0.51 (0.50)	0.50 (0.50)	0.59 (0.49)
Work experience during education					
Any internship during education	0.59 (0.49)	0.51 (0.50)	0.55 (0.50)	0.63 (0.48)	0.61 (0.49)
Any student job during education	0.81 (0.40)	0.72 (0.45)	0.81 (0.39)	0.79 (0.41)	0.86 (0.34)
Timing of the start of job search					
Start of job search: quarter in the year					
First quarter	0.09 (0.28)	0.13 (0.34)	0.10 (0.30)	0.07 (0.25)	0.10 (0.29)
Second quarter	0.17 (0.37)	0.10 (0.31)	0.18 (0.39)	0.16 (0.36)	0.17 (0.37)
Third quarter	0.66 (0.47)	0.70 (0.46)	0.63 (0.48)	0.67 (0.47)	0.68 (0.47)
Fourth quarter	0.09 (0.28)	0.06 (0.24)	0.09 (0.28)	0.10 (0.31)	0.06 (0.24)
Months between leaving school and starting job search	2.04 (4.03)	2.61 (4.68)	2.12 (4.76)	2.03 (3.57)	1.67 (2.68)
Seasonal and business variation					
Youth unemployment rate at start date of job search	18.79 (2.69)	19.21 (3.12)	18.36 (2.75)	19.05 (2.54)	18.98 (2.59)

Reported figures are means and standard deviations in parentheses. Subsample 'C' contains men who are right censored before making any transition during the observation period; subsample 'E' contains men who directly transit to an adequate job; subsample 'OC' contains men who enter a job for which they are overeducated and who are subsequently right censored before making any transition to an adequate job; subsample 'OE' contains men who are temporarily overeducated before entering an adequate job. The indicator variable 'grandmother (mother's side) foreign' is 1 if the respondent's grandmother on mother's side has not the Belgian nor a North American, British, Scandinavian, Western European or Australian nationality. Unless otherwise stated, variables reflecting the educational career of the respondent are measured at the moment the youngster stops schooling (for the first time). The parental education level equals the number of successful schooling years beyond primary school. For instance it is equal to 6 if the parent has successfully completed general or technical secondary education, but did not successfully complete any year of higher education. The 'youth unemployment rate' is the seasonally unadjusted monthly Belgian unemployment rate for males aged less than 25 years following the ILO definition (source: Eurostat).

effect is measured for a population that has less chances of finding an adequate job than if it was measured at the start of the unemployment spell. This biases the treatment effect towards zero because the treatment effect is confounded with the unobserved lower chances of finding adequate employment in this population (Lancaster, 1990). The Timing of Events method proposed by Abbring and van den Berg (2003) takes this double selection problem into account and identifies therefore the true causal impact of a transition to overeducation on the speed of transition to an adequate job. We first write down the econometric model and then discuss why we believe that the main identifying assumptions of the Timing of Events method are satisfied.

5.4.2 The Econometric Model

In the following, the index o indicates overeducation and the index e refers to adequate employment. The transitions of interest into overeducation and adequate employment are represented by two random latent durations: T_o and T_e, with t_o and t_e denoting their realisations. We assume that all individual differences in the joint distribution of both durations can be characterised by explanatory variables X and V. X denotes the observed variables as described in Section 5.3.3 with realisation x.[14] V, on the other hand, is unobservable to the researcher and transition-specific. More concretely, V is a vector (V_e, V_o) with realisation (v_e, v_o). X and V are assumed to be independently distributed (see Section 5.4.3 for further discussion).

Abbring and van den Berg (2003) assume that T_e and T_o are independent conditionally on X and V, so that the joint distribution of $(T_o, T_e)|(X, V)$ can be written as the product of the distributions of $T_e|(X, V_e)$ and $T_o|(X, V_o)$ which are in turn completely determined by their hazard rates $\theta_e(t|t_o, x, V_o)$ and $\theta_o(t|x, V_o)$, where t is the elapsed job search duration. These hazard rates are then specified according to the following Mixed Proportional Hazard (MPH) form:

$$\begin{cases} \ln \theta_o(t|x, V_o) & = & \ln \lambda_o(t) + x\prime\beta_o + V_o, \\ \ln \theta_e(t|t_o, x, V_e) & = & \ln \lambda_e(t) + x\prime\beta_e + \delta(t|t_o, x) I(t > t_o) + V_e, \end{cases} \tag{5.4.1}$$

where $I(.)$ is an indicator function, which is 1 if the argument is true and 0 otherwise, and $\delta(t|t_o, x)$ is the treatment effect of overeducation on the speed of transition to an adequate job.[15] Observe that it can be any function of t, t_o and x, but cannot depend on any unobserved factor.

In the benchmark model we allow the treatment effect to depend on both the duration

[14]To avoid cumbersome notation, we ignore that youth unemployment rate is a time-varying covariate.

[15]Note that in this specification we do not model transitions from a job for which one is overeducated back to unemployment. If this happens, they remain at risk for a transition to adequate employment as members of the "treated" group.

since entry in overeducation $(t - t_o)$ and on the elapsed unemployment duration until entry in overeducated employment t_o. The first factor $(t - t_o)$ aims at capturing a gradually decreasing locking-in effect and/or steadily growing investment effect. Locking-in may reflect investment in specific human capital (Pissarides, 1994), cognitive decline (de Grip et al., 2008), habituation (Verhaest and Omey, 2009) or reduced job-search effort on-the-job (Holzer, 1987). The investment effect reflects the gradually increasing promotion opportunities with work experience, as described in the career mobility theory (Sicherman and Galor, 1990). The dependence on the elapsed unemployment duration, on the other hand, aims at testing whether long-term unemployed benefit more from a stepping stone effect (if any) than short-term unemployed, since accepting any job might reduce the scarring effects of long-term unemployment. We include quadratic terms to allow for nonlinearity of these effects over time.

$$\delta(t|t_o) = \delta_0 + \delta_1(t - t_o) + \delta_2(t - t_o)^2 + \delta_3 t_o + \delta_4(t_o)^2. \tag{5.4.2}$$

In Section 5.5.2, in which we report a number of sensitivity analyses, we discuss some extensions in which the treatment effect depends on some other explanatory variables.

$\lambda_o(t)$ and $\lambda_e(t)$ represent the baseline hazard functions for transitions into overeducation and adequate employment. The hazard rate is said to be duration dependent if these functions are time-variant. Positive (negative) duration dependence in the transition into overeducation, respectively adequate employment, means that $\lambda_o(t)$, respectively $\lambda_e(t)$, are increasing (decreasing) in t. We follow the literature by specifying these baseline hazards as piecewise constant:

$$\begin{cases} \ln \lambda_o(t) & = \alpha_m^o, \\ \ln \lambda_e(t) & = \alpha_m^e, \end{cases} \quad \text{for} \quad t \in [t_{m-1}, t_m), \tag{5.4.3}$$

where m is an indicator of the time interval and where in the application $m \leq 8$ and $t_0 = 0$, $t_1 = 1$, $t_2 = 2$, $t_3 = 3$, $t_4 = 4$, $t_5 = 6$, $t_6 = 9$, $t_7 = 18$ and $t_8 = +\infty$.

We estimate the benchmark model by Maximum Likelihood. We distinguish between four types of likelihood contributions, conditional on the unobserved heterogeneity distribution, depending on the labour market history of the youth described in Section 5.3.3; $l_c(V)$, $l_e(V)$, $l_{oc}(V)$ and $l_{oe}(V)$. We refer to the working paper version of this chapter for the derivation of these conditional contributions taking the time-grouped nature of the data into account (Baert et al., 2012).

To obtain the unconditional likelihood contributions, we integrate the four conditional contributions over the unobserved heterogeneity distribution. We follow Heckman and Singer (1984) and assume that (v_e, v_o) is randomly drawn from a discrete distribution with a finite and a priori unknown number K of points of support. Since we include a constant term in X, v_{e1} and v_{o1} are normalised to 0. The probabilities associated to these points of support are specified as logistic transforms:

$$p_k = \frac{\exp(\gamma_k)}{\sum_{j=1}^{K} \exp(\gamma_j)}, \quad \text{with} \quad k = 1, 2, \ldots, K; \gamma_1 = 0. \tag{5.4.4}$$

Hence, the likelihood contribution for individual i in subsample $n \in \{c, e, oc, oe\}$ unconditional on unobserved heterogeneity is:

$$l_{ni} = \sum_{k=1}^{K} p_k \cdot l_n(v_{ek}, v_{ok}), \quad \text{with} \quad n \in \{c, e, oc, oe\}. \tag{5.4.5}$$

We can then write the unconditional log-likelihood as the sum of the unconditional individual log-likelihood contributions:

$$L = \sum_{i=1}^{N} [J_{ci} \ln(l_{ci}) + J_{ei} \ln(l_{ei}) + J_{oci} \ln(l_{oci}) + J_{oei} \ln(l_{oei})], \tag{5.4.6}$$

where J_{ni} equals 1 if l_{ni} is the contribution of individual i to the likelihood and J_{ni} equals 0 otherwise. We maximise this log-likelihood according to the procedure described in Gaure et al. (2007). In particular, we increase the number of points of support until the likelihood function does not show any improvement and subsequently select the model that minimises the Akaike Information Criterion (AIC) to reduce the risk of bias induced by an over-parameterised model.

5.4.3 Identification

Unlike some other methods that aim at resolving the selection problem, the Timing of Events method does not require any exclusion restrictions. All observed determinants may affect both the transition to overeducation (the treatment) and the transition to an adequate job (the outcome of interest). However, the method requires another set of identifying assumptions (Abbring and van den Berg, 2003) of which we discuss the credibility of the four most important ones.

Firstly, it is essential that the moment at which employment is entered may not be anticipated. Since the timing of job offers cannot be anticipated and neither the employer nor the job searcher has in general an interest to postpone hiring once the hiring decision is taken, we believe that the time lag between the moment at which the job was offered and the moment at which the job is entered, is relatively short, so that the no-anticipation assumption is approximately satisfied. The fact that we observe many transitions to a job in the first few months since individuals started job search (see Figure 5.1) demonstrates that in many cases individuals enter employment shortly after they have been offered a job, suggesting that bias induced by anticipation is not so important. Note that by ignoring anticipation we tend to overestimate the treatment effect, or stated otherwise in case of a negative treatment effect: the entrapment in overeducation might be even larger than what our estimates suggest. This occurs because those who anticipate job entry have during this period a lower propensity of entering a job and have incorrectly been assigned to the control group, so those who are still searching for a job. Assigning these individuals during that period to the treatment group will therefore decrease the transition rate to an adequate job of the treated and increase the transition rate of the control group. However, we cannot correct for this bias, since we do not observe the timing of job offers.

Secondly, observed and unobserved determinants affect the transition rates to overeducation and adequate employment of the untreated individuals proportionally. This is the so called Mixed Proportional Hazard (MPH) assumption. The Monte Carlo analysis of Gaure et al. (2007), specially designed to evaluate the reliability of the Timing of Events Method of Abbring and van den Berg (2003), has shown that this assumption is crucial, at least if only time constant explanatory variables are available.[16] We are concerned that the MPH assumption may not be satisfied across levels of educational attainment and that this therefore may be a source of bias. On the one hand, one may argue that the highest educated individuals have more opportunities to be overeducated, simply because accepting any job with requirements below the highest level of educational attainment leads to overeducation. By contrast, individuals with a lower secondary degree are only overeducated if they accept a job that does not require any educational attainment. On the other hand, there is the evidence on job polarisation indicating that technological change has resulted in a substantial drop in the number of medium-skilled jobs over the past decades (Goos et al., 2009). Consequently, the medium educated individuals are more likely overeducated than the lower or higher educated. We therefore perform a sensitivity analysis in which we estimate our model separately for each level of educational attainment, relaxing thereby the proportionality assumption in this dimension. Since our findings are robust to this sensitivity analysis, we are convinced that the MPH assumption is reasonable in this application.

Thirdly, X and V should be independently distributed. This is a strong assumption, but it can be relaxed if one is willing to assume that the unobserved heterogeneity conditional on x can be written as $v_k \exp(x\mu_k)$ (for $k = e, o$), where V_k is then independently distributed

[16] Note that despite the youth unemployment rate is a time-varying explanatory variable, it is of no use for identification, since its variation is the same across all observations in the sample.

from X and where μ_k is some unknown parameter vector. In this case it is not difficult to see that the treatment effect is still consistently estimated, but that the parameters associated to the covariates of x can no longer be given a structural interpretation. This is very similar to widely used extension of the random effect probit model as established by Chamberlain (1980). See Cockx et al. (2013) for further discussion.

Finally, based on Monte Carlo analysis Baker and Melino (2000) have shown that Maximum Likelihood estimates of a flexible specification of both the baseline hazard and the unobserved heterogeneity distribution in single spell duration models tend to be biased towards finding an excessively dispersed distribution of unobserved heterogeneity, especially if the sample size is small, as it is in our case. However, these researchers show that these biases can essentially be eliminated by selecting the model on the basis of a criterion that penalises the model with too many points of support. Gaure et al. (2007) arrive at very similar conclusions and propose to use the Akaike Information Criterion (AIC) to select the appropriate model. We therefore follow the latter procedure to select the appropriate model. Moreover, to ascertain that the small sample size does not bias our findings, we show that our findings are robust to sensitivity analyses in which we (i) increase the sample size (by integrating women in the analysis); (ii) restrict the specification of the baseline hazard, since this is shown to reduce the bias (Baker and Melino, 2000); and (iii) reduce the number of explanatory variables.

5.5 Results

5.5.1 Main Results

In this subsection we first discuss the main results. In a subsequent subsection we report a number of sensitivity analyses. In the main text we focus on the treatment effects. The complete estimation results, including those that do not correct for selection on unobservables, can be found in the working paper version of this chapter (Baert et al., 2012).

On the basis of the AIC we retain for the benchmark specification of the treatment effect the model with three heterogeneity types. However, the probability assigned to one of the points of support is small (4%), so that the estimates are not very different from a model with two points of support.

The main estimation results of the benchmark model are summarised in Table 5.2. First, the point estimate for δ_0, indicating the treatment effect in the first month of overeducation, is very negative and highly statistically significant. In the full sample the monthly transition rate into adequate employment drops by about 98%[17] for this month. Second, the point

[17] $0.98 = 1 - \exp(-4.080)$.

Table 5.2: Results of the Benchmark Model

Treatment effect		
Constant: δ_0	$-4.080{*}{*}{*}$	(0.354)
Interaction with $(t - t_o)$: δ_1	-0.014	(0.012)
Interaction with $(t - t_o)^2$: $\delta_2 \times 100$	0.011	(0.011)
Interaction with t_o: δ_3	$0.232{*}{*}{*}$	(0.088)
Interaction with $(t_o)^2$: δ_4	-0.004	(0.004)
Log-likelihood	-4594.661	
AIC	9331.321	
Parameters	71	
Observations	1434	

***(**)((*)) indicates significance at the 1% (5%) ((10%)) level. Standard errors in parentheses.

estimates for δ_1 and δ_2, capturing the effect heterogeneity in the elapsed duration since inflow into a job for which one is overeducated, are not statistically significant. Third, by contrast, δ_3 is highly significant. This means that young men delay the transition to an adequate job by accepting jobs for which they are overeducated more in case they do this at the start of the unemployment spell rather than later on. The magnitude of this adverse effect declines with unemployment duration and, if we ignore the insignificant second order term (δ_4), it becomes even positive beyond an unemployment duration of 17.6 months.[18] This would be an important finding, since it would mean that only short-term unemployed graduates get trapped in overeducation. Long-term unemployed would accelerate the transition to an adequate job by temporarily accepting a job for which they are overeducated. However, this conclusion crucially hinges on ignoring the insignificant second order term δ_4. If we take it into account, the treatment effect attains a maximum at 29 months[19] and decreases again thereafter. At this maximum the transition rate to an adequate job after entry in overeducation is still 51%[20] below what it would have been if one would have only accepted adequate jobs. In the sensitivity analysis below, we will argue that this is the correct interpretation of our findings. The transition to overeducation is thus never a stepping stone, but always a trap.

In order to get some sense of the meaning of the size of the treatment effect, we calculate for the treated group, so for all men who are overeducated in the first job that they enter, the first quartile and median duration until transition in an adequate job, both in the case of treatment and in the counterfactual of no treatment. In this counterfactual we impose that these individuals do not enter any job for which they are overeducated (we set the

[18] $17.6 = 4.08/0.232$.
[19] $29 = 0.232/(2 \cdot 0.004)$.
[20] $-4.08 + 0.232 \cdot 29 - 0.004 \cdot 29^2 = -0.716$ and $1 - \exp(-0.716) = 0.51$.

transition rate to zero, so $\theta_o(t|x, V_o) = 0$, but instead only transit directly to adequate jobs at the estimated transition intensity (without treatment, so $\delta(t|t_o, x) = 0$). The median (first quartile) duration until a transition in an adequate job in case of treatment is 115.8 (39.4) months compared to 3.0 (1.1) months in the counterfactual. These figures reconfirm that if unemployed graduates aim at entering a job for which they have the appropriate level of education, they should not accept a job for which they are overeducated.

We briefly discuss some secondary results. Recall, however, that a structural interpretation of the coefficients for the observed covariates is hazardous given the potential dependence between X and V.[21] We find that the employment gap between foreign and native youth is significantly larger for the transition into jobs for which one is overeducated than for the transition into adequate jobs. In addition we get a stable positive effect of an honours degree in tertiary education on the transition to adequate employment. Conducting any student work during education affects the transition to overeducation and adequate employment with a similar magnitude.

The fact that the high educated ceteris paribus less rapidly find a job than the lower educated is most likely a consequence of only retaining in the sample young men who started job search after leaving education: this group is a negatively selected subsample of the higher educated. One might argue that this might cause non-proportionality of the unobserved determinants of the transitions from unemployment, since for lower levels of education this negative selection is less an issue because they are less likely to find a job immediately after leaving education. Conditional on the level of educational attainment, one might therefore expect higher V's for the lower educated than for the higher educated. However, we will investigate this point in the sensitivity analysis reported in the next subsection and demonstrate that this concern is not an issue for our data.

5.5.2 Sensitivity Analysis

In Table 5.3 we report some robustness checks of our main result that young men get trapped in jobs for which they are overeducated. First, we test the robustness of our results to the alternative measure of overeducation as defined in Section 5.3.2. This alternative measure results in a rearrangement of the sample of graduates over the four subsamples: subsamples 'C', 'E', 'OC' and 'OE' comprise 97, 886, 267 and 184 individuals respectively. As explained in Section 5.3.2, this alternative measure is closely related to a measure of genuine overeducation. If so, Chevalier (2003) argues that the likelihood of promotion should increase. The estimated negative treatment effects are indeed somewhat less important. However, they do not reverse

[21] See the discussion of the identifying assumptions in Section 5.4.3.

the main conclusions of the benchmark model. For the treated group median (first quartile) duration until a transition in an adequate job in case of treatment is 85.7 (17.3) months compared to 1.4 (1.0) months in the counterfactual of no treatment. For the benchmark model these numbers were 115.8 (39.4) months if treated and 3.0 (1.1) months in the counterfactual.

The alternative measure leads to two main differences compared to the benchmark model. A first one is that the quadratic interaction term with unemployment duration (δ_4) is now very significantly different from zero. This provides a first confirmation that we should take this interaction term into account when interpreting the results. A second main difference is the significance of δ_1 and δ_2, the interactions with the time since entry in a job for which one is overeducated $(t - t_o)$: the entrapment effect is (slightly) more pronounced during the first 70 months.

In a second sensitivity analysis we re-estimate our model separately on the four subsamples defined according to their highest attained level of education.[22] We do so because we are concerned that the MPH assumption might fail across levels of educational attainment (see our discussion of identifying assumption 2 in Section 5.4.3 as well as our discussion at the end of the previous subsection). However, panels B1 until B4 indicate that the estimates of the treatment effects are not different across these subgroups. For all educational levels, accepting a job for which one is overeducated prolongs the transition to an adequate job. Moreover, δ_1 and δ_2 are insignificant for all subsamples. As in the pooled analysis, δ_3 is large and positive, but only significantly for the two lowest levels of education. Finally, δ_4 is systematically negative and even significantly (at the 10% level) for graduates with a secondary education. If this quadratic term is taken into account, the treatment effect remains negative for all possible unemployment durations. These findings are therefore reassuring and consistent with the main findings reported in Section 5.5.1.

In a third sensitivity test we introduce more heterogeneity in the treatment effect. On the one hand it could be argued that in a booming economy it would be easier to promote from overeducation to an adequate job. On the other hand we want to capture the difference in treatment effects according to differences in skills, since one may argue that, within each educational level, the higher skilled transit faster to an adequate job. This would be evidence of promotion induced by genuine overeducation, since the more skilled are more likely to be

[22]For the models reported in panels B1 and B3 the lowest AIC was obtained with two points of support, while for the model reported in panel B2 three points of support were required. For the model reported in panel B4 one point of support was optimal. This is probably due to the relatively small number of observations in this sample. Since the findings of this model are quite different from the models accounting for heterogeneity and since among the models that account for heterogeneity the one with 3 points of support yields the lowest AIC, we choose to report the parameter estimates of the latter model rather than those of the model that disregards unobserved heterogeneity.

Table 5.3: Results of the Sensitivity Analyses

	A. Sensitivity analysis 1 Modified subjective overeducation measure	B1. Sensitivity analysis 2 Subpopulation "Lower secondary education"	B2. Sensitivity analysis 2 Subpopulation "Higher secondary education"	B3. Sensitivity analysis 2 Subpopulation "Lower tertiary education"	B4. Sensitivity analysis 2 Subpopulation "Higher tertiary education"	C. Sensitivity analysis 3 Additional treatment interactions	D. Sensitivity analysis 4 Population extended with women
Treatment effect							
Constant: δ_0	-4.103***	-4.322***	-4.893***	-4.149***	-4.044	-4.216***	-3.445***
	(-0.267)	(-0.618)	(-0.689)	(-0.606)	(-2.906)	(-0.605)	(-0.38)
Interaction with $(t - t_o)$: δ_1	-0.045***	-0.006	-0.006	0.015	0.081	-0.015	0.004
	(-0.016)	(-0.022)	(-0.017)	(-0.042)	(-0.088)	(-0.012)	(-0.009)
Interaction with $(t - t_o)^2$: $\delta_2 \times 100$	0.032**	0.005	0.011	-0.02	-0.143	0.012	-0.006
	(-0.015)	(-0.018)	(-0.015)	(-0.053)	(-0.125)	(-0.01)	(-0.009)
Interaction with t_o: δ_3	0.320***	0.357***	0.421***	0.219	0.2	0.229**	0.181***
	(-0.072)	(-0.121)	(-0.152)	(-0.214	(-0.35)	(-0.09)	(-0.056)
Interaction with $(t_o)^2$: δ_4	-0.009**	-0.008	-0.011*	-0.009	-0.003	-0.004	-0.004**
	(-0.003)	(-0.005)	(-0.006)	(-0.013)	(-0.012)	(-0.004)	(-0.002)
Interaction with unemployment rate: δ_5						0.012	
						(-0.028)	
Interaction with years of schooling delay (age of 16): δ_6						-0.222	
						(-0.153)	
Interaction with female: δ_7							-0.027
							(-0.121)
Log-likelihood	-4058.329	-1024.419	-1978.052	-765.183	-670.047	-4593.34	-9588.461
AIC	8258.657	2156.839	4078.103	1648.365	1466.093	9332.68	19318.92
Parameters	71	54	61	59	63	73	71
Observations	1434	338	620	244	230	1434	2956

***(**)((*)) indicates significance at the 1% (5%) ((10%)) level. Standard errors in parentheses. Concerning Panel B1 one individual is dropped since he was the only one with arts secondary education at the age of 16. Concerning Panel B3 one individual is dropped since he was the only one with vocational secondary education at the age of 16.

genuinely overeducated and the latter are more likely to promote (see the discussion in Section 5.1). To this end, we further include interactions with the unemployment rate and the years of schooling delay at the age of 16. Panel C indicates that neither of the two additional interaction effects is statistically significant.

In a fourth set of sensitivity analyses, we aim at addressing the critique that we might fail to identify the selection on unobservables as a consequence of a too small sample size and the relatively large number of estimated parameters. First, we enlarge our sample by including women. Recall that we did not include women in the benchmark analysis, since we were concerned that, as a consequence of fertility considerations and for their traditionally higher responsibility in household activities, including child rearing, women would more likely get trapped in jobs for which they are overeducated. This sensitivity analysis reveals, however, that this concern was void. The findings of this analysis are reported in Panel D of Table 5.3. The interaction of the treatment effect with female gender is very small (-0.027) and insignificantly different from zero. The constant term is slightly lower, implying that by accepting a job for which one is overeducated in the first month of unemployment, the transition rate decreases by 97%[23] instead of by 98% in the benchmark model. Different from the benchmark model, the quadratic term of the interaction effect of the treatment with the unemployment duration at the start of the spell of overeducation (t_o^2) is now significantly negative. This is a further confirmation that in the benchmark model the coefficient of this quadratic term is insignificant as a consequence of lack of precision and not because it is truly zero.

We further investigate the claim that selection on unobservables might not be well identified as a consequence of overfitting, by (i) reducing in the benchmark model the number of explanatory variables[24] and by (ii) reducing the number of duration intervals in the baseline hazard.[25] However, this influences the findings only negligibly.[26]

5.5.3 Discussion

The finding that overeducation strongly retards the transition to adequate employment clearly challenges Sicherman and Galors' (1990) career mobility thesis. This adds to the more indirect evidence provided by other researchers who studied upward mobility of overeducated individuals, and who concluded that many individuals remain overeducated for very long periods

[23] $0.97 = 1 - \exp(-3.445)$.

[24] We exclude the following variables: "additional successful years at school after highest attained level of education", "additional degree at highest attained level of education", "tertiary education: grade", "years of schooling delay (at age of 16)", "year of birth", "any internship during education" and "months between leaving school and starting job search".

[25] We reduce the number of duration intervals from eight to four: $t_0 = 0$, $t_1 = 6$, $t_2 = 18$ and $t_3 = +\infty$.

[26] These results are not reported, but are available upon request.

(see Section 5.1). It also confirms the conclusions of studies that test this theory in an indirect way by investigating the relationship between overeducation and training participation or skill acquisition. These studies in general find that overeducated workers participate less often in training and acquire less additional skills than adequately educated workers with a similar educational background (Robst, 1995; van Smoorenburg and van der Velden, 2000; Verhaest and Omey, 2013). Finally, it is also consistent with the finding that overeducated workers experience no more wage growth than adequately educated workers (Büchel and Mertens, 2004; Korpi and Tåhlin, 2009).

An often-formulated criticism on the literature of overeducation is that many workers are only "apparently overeducated" (Chevalier, 2003; Green and McIntosh, 2007; Green and Zhu, 2010), that is because of occupational upgrading and/or lower quality of human capital, they are formally overeducated but not overskilled. Hence, for these individuals, making a transition to a job for which they are formally adequately educated may simply be not an option. While this might be true for some of the individuals, we have no indications that this drives our results. Firstly, our results were largely similar if based on an adapted subjective overeducation measure which, as argued in Section 5.3.2, is a good proxy of genuine overeducation. Secondly, by focusing our analysis on young graduates we do not face the problem that work experience confounds the measure of overeducation. This reduces the likelihood of mismeasuring overeducation. Thirdly, we accounted for selection on unobservables, implying that our basic estimates are likely reflecting the true causal effect of overeducation and not just unobserved ability (Leuven and Oosterbeek, 2011). Fourthly, additional estimates did not deliver indications of heterogeneous effects depending on skill level as measured by the years of schooling delay. Finally, since a job change in our data is defined to be a change either in employer or in the tasks to be executed with the same employer, we account for possible changes in skill requirements due to task upgrading and internal promotions.

Accounting for selection on unobservables in the analysis of the mobility behaviour of overeducated workers is an important contribution of this study. Remarkably, we found that the entrapment effect is underestimated if such selection is assumed to be absent. Those with favourable unobserved characteristics for a transition to an adequate job are thus also more likely to make the transition to a job for which they are overeducated. This suggests that selection on unobservables is mainly driven by other factors than differences in unobserved ability. A reasonable explanation is that highly motivated job seekers have a higher search intensity, resulting in higher transition rates for both types of jobs, and that more able youth is not more selective than less able youth in their job search and acceptance behaviour. Motivated youth seems to be willing to accept any job, which according to our findings may have a long-lasting negative impact on the quality of the job match. We return to this point at the end of this section where we discuss policy implications.

If overeducation is not a stepping stone to an adequate job, one may wonder why. What explains the strong entrapment effect immediately after entry into overeducation? And why does this entrapment effect not decrease with job tenure? An explanation for the strong initial entrapment effect may be that, by accepting a position that does not match with one's level of educational attainment, the worker may transmit a negative signal to prospective employers (McCormick, 1990). Furthermore, it may be difficult to maintain the same job search intensity on-the-job than when one is unemployed (Holzer, 1987). If so, this may also reduce the incentive for employers to create vacancies for adequately matched jobs and thereby reinforce the low transition rate to an adequate job (Dolado et al., 2009).[27] Furthermore, the entrapment effect might not decrease with job tenure as a consequence of investments in specific human capital (Pissarides, 1994), cognitive decline (de Grip et al., 2008), or habituation (Verhaest and Omey, 2009).

Another interpretation of the strong entrapment effect may be that it reflects that vacancies for adequate positions are cyclically or structurally lacking. However, we have no indications that this is the dominant explanation. First, in our sample entry in the labour market is spread out over a relatively long period (from 1996 to 2006) covering both years of economic upturn and downturn. Second, the median search duration until an adequate job, under the counterfactual that no one enters overeducation, is simulated to be only three months on the basis of our benchmark model (see Section 5.5.1). This does not fit with a structural mismatch between the qualifications of the graduates and those needed by the labour market. Third, we found that the entrapment effect is not significantly related to the unemployment rate. Neither did we find that the entrapment effect is significantly higher for the medium educated, who may be affected by job polarisation.

These findings may follow from the fact that the relationship between labour demand and the entrapment effect is theoretically less clear than it may seem. While a lack of vacant adequate positions is likely to result in longer spells of overeducation, it will also decrease the likelihood of finding an adequate position for the unemployed. Of course, it is possible that some of the previously mentioned mechanisms underlying the entrapment effect (such as via job search intensity or signalling) depend on the availability of jobs, but it is a priori unclear in which direction this will affect the entrapment effect. For instance, the average quality of overeducated workers will be higher in a slack labour market so that negative signalling effects

[27] The low likelihood of finding an adequate position may also be explained by an efficiency wage type of argument (Skott, 2006).

resulting from overeducation may then be lower.

Given the strong entrapment effect of overeducation, we need to understand why individuals are prepared to accept such jobs. One reason can be that the financial and psychological costs of overeducation are, although significant, still relatively modest in comparison to those of being unemployed (Verhaest and Omey, 2009). Albrecht and Vroman (2002) developed a matching model in which they show that it may be optimal for skilled workers to accept unskilled jobs as long as the productivity differences between skilled and unskilled jobs are not too large. In that case, the expected earning gains resulting from future adequate employment no longer outweighs the income loss resulting from unemployment. This argument applies especially for individuals who are already long-term unemployed, since, if these individuals reject job offers for which they are overeducated, they risk remaining unemployed even longer as a consequence of the negative duration dependence in the exit rate from unemployment. For instance, based on the benchmark model reported in Table 5.2, we predict that the median remaining unemployment duration of a young graduate who is already one year unemployed and who does not accept jobs for which he is overeducated is 11.9 months. Moreover, this cost in terms of expected remaining unemployment duration continues to increase with elapsed unemployment duration.

The preceding argument does not hold, however, for short-term unemployed graduates. For instance, for a young graduate who is just one month unemployed and who follows the aforementioned job search strategy the median remaining unemployment duration is only 1.8 months. So, in this case there should be another explanation why young graduates accept jobs for which they are overeducated. One explanation is that individuals are credit constrained. Since school-leavers are only entitled to unemployment benefits after nine months of registered unemployment, this is a natural explanation, especially if they are no longer financially dependent on their parents' income. 15% of the individuals in the sample retained for the analysis left the parental home at the moment at which they started job search. Furthermore, after expiration of this waiting period, the benefit level is low. For singles and cohabiting individuals it is not very different from the means-tested social assistance benefit level.[28]

Further reasons why short-or long-term unemployed graduates may accept jobs for which they are overeducated are that (i) they may be insufficiently informed about the long-term cost of doing so or (ii) that they are too impatient, that is they display "hyperbolic" time preferences (see, e.g., Frederick et al., 2002; DellaVigna and Paserman, 2005; Fang and Silverman, 2009).

[28]Depending on factors such as cohabitation status and number of children, the unemployment benefit is only 4 –24% higher than the social assistance benefit level (source: National Employment Office and Federal Public Service for Social Integration of the Belgian government). In addition, in contrast to beneficiaries of social assistance, the unemployed are usually not entitled to reduced rates of telecommunication, electricity, heating and public transport.

Impatient individuals with so called hyperbolic preferences tend to choose activities with immediate rewards and delayed costs (accepting a job for which one is overeducated) to those with immediate costs and delayed rewards (continuing search for an adequate job). Which of all these explanations apply, is matter for future research.

These findings have policy implications. Since overeducation is not a stepping stone to an adequate job, it may pay for young graduates to search for jobs selectively and therefore reject jobs for which they are overeducated.[29] This suggests that policy makers should take care in not providing incentives to young unemployed graduates to accept any job too early in the unemployment spell, since this may induce persistent mismatch between the individuals' qualifications and those required on the job. This is not only costly for the individual concerned, but also for society as a whole. Consequently, the short-term benefits in terms of job transitions that can be induced by a tightening of entitlement conditions to UB, such as the Belgian government implemented in 2012 (see Section 5.2), must be carefully traded-off against the long-term costs in terms of qualification mismatch in the job, especially if this tightening targets the short-term unemployed for whom the costs of waiting for an adequate job are not yet counterbalanced by the costs of the risk of not finding any job at all. Besides, the form of the policy intervention that aims at preventing that unemployed graduates accept to early jobs for which they are overeducated depends on the reason for which they accept these jobs. If the reason is related to a credit constraint, then the policy should be aimed at lifting this constraint. By contrast, if it is lack of information or impatience that is the main explanation, guidance in the job acceptance behaviour may be more appropriate.

5.6 Conclusions

In this research project we investigated whether overeducation at the start of the career speeds up the transition to adequate employment. Contrary to many other contributions in this research area, we handled selection on both observables and unobservables. For this, we applied the Timing of Events approach. Our findings indicate that, even for long-term unemployed young people, accepting a job for which one is overeducated substantially retards the transition to an adequate job. By accepting a job for which one is overeducated rather than only accepting adequate job matches, monthly transition rates into adequate employment fall by 51–98%, depending on the elapsed unemployment duration. This result was found to be robust

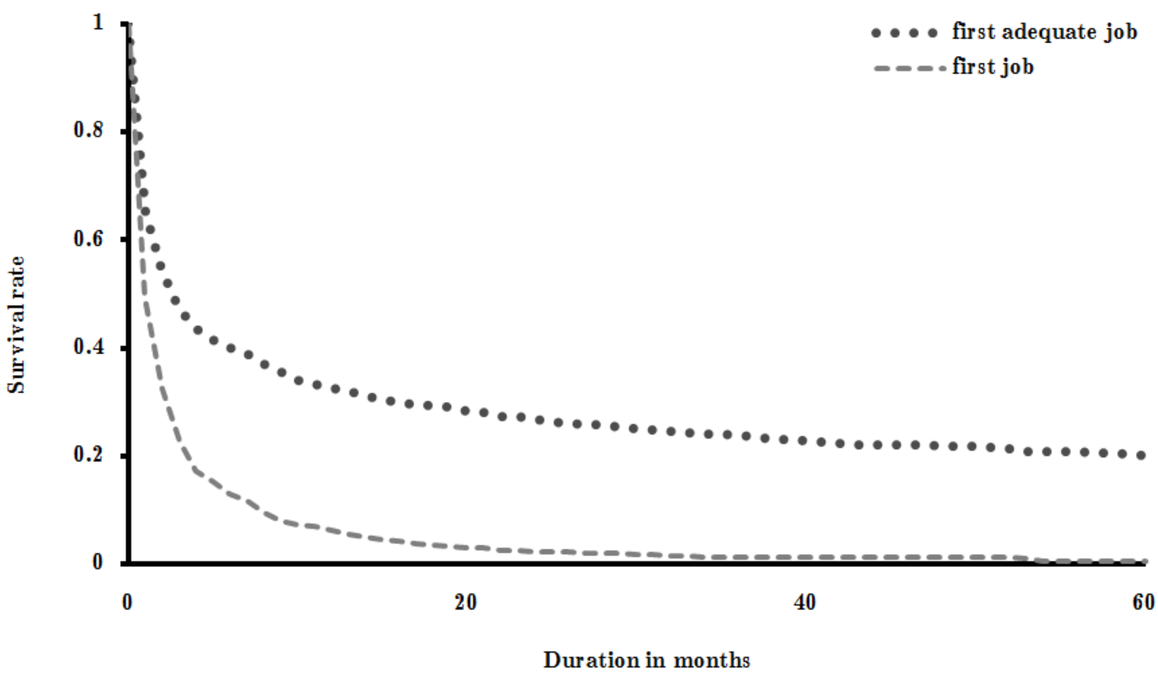
[29] We implicitly assume that workers earn more in adequate jobs than in jobs for which they are overeducated. Although Leuven and Oosterbeek (2011) challenge this, since many studies giving evidence for lower wages for the overeducated do not properly account for selection bias and measurement error issues, there are a couple of more recent studies that attempt to address this criticism and still arrive at the same conclusion (see, e.g., Dolton and Silles, 2008; Korpi and Tåhlin, 2009; Verhaest and Omey, 2012).

against various sensitivity checks including the comparison of our results with respect to two measures of overeducation. We argued that this entrapment effect is likely to be explained by a combination of factors, such as a negative signal of overeducation, reduced job search intensity, job-specific human capital investments, cognitive decline and habituation. Furthermore, for long-term unemployed graduates the substantial costs entailed by the expected duration of the remaining unemployment spell and for short-term unemployed credit constraints, lack of information or impatience may explain why graduates nevertheless choose for such jobs.

From a policy point of view, these results imply that policy makers should, especially early in the unemployment spell, carefully trade-off the benefits and costs of forcing young graduates to accept any job, so ignoring whether the educational requirement of this job matches the qualifications of these graduates, since this may lead to a persistent qualification mismatch in the job. This is not only costly for the individual concerned, but also for society as a whole.

5.7 Appendix: Additional Figures

Figure 5.3: Kaplan-Meier Estimates: Duration Between Start Job Search and Entry in the First (Adequate) job (Modified Self-Assessment Method)



6
General Conclusions

At the end of this PhD thesis we wrap up with three key results and their policy implications ("to take away for policy makers") and three key directions for future research ("to take away for researchers"). A first key result of this thesis is that ethnicity, even after controlling for socio-economic background characteristics, affects success in school and first labour market transitions in Flanders. Chapter 2 showed that a small pure ethnic gap exists in educational attainment if school delays are neglected. This pure ethnic gap becomes substantial if grade retention is taken into account. This is an important finding, since it means that ethnic schooling gaps, in particular gaps in schooling delay, cannot be eliminated by focussing policy to disadvantaged groups irrespectively of their ethnic background. Moreover, Chapter 2 revealed that the pure ethnic gap in the school-to-work transition is very important. In addition, in Chapter 3 it was shown that this pure ethnic gap in the labour market is to a large extent induced by ethnic discrimination in the Flemish youth labour market. This discrimination is not only unacceptable from an ethical perspective, but has also important economic consequences. Given the significant challenges due to an ageing population that face the Belgian labour market, it is important to call on all groups of the population so that there is no room for the (partial) exclusion of minority groups. The legal framework to punish discrimination is available in Belgium, so that the main benefit seems to lie in a more vigorous detection of discrimination. One could investigate whether this could not happen based on a systematic application of the experimental method we have reported in Chapter 3. On the other hand, the latter chapter indicates that employers discriminate less if they have to attract workers for occupations for which labour market tightness is high. Our results therefore suggest advising minorities to apply for jobs that are difficult to fill. However, such a policy advice may only work to the extent the competencies of minorities match the requirements for these jobs and that the tightness in the labour market is partly a consequence of minorities not being informed about the occupations for which employers have difficulties in filling vacancies. Moreover, the benefit of applying for bottleneck professions in terms of hiring chances should be weighed against the potential losses induced by the particular working conditions in these occupations.[1]

A second key result is that grade retention in school has important short- and long-term consequences. In contrast to most of the previous empirical contributions, we find that grade retention has a positive impact on the next evaluation and can permanently affect subsequent educational achievements. The direction of the permanent effect is "essentially" heterogeneous: while from a long-term perspective more able students are penalised by retention, less able students permanently benefit from it in terms of later success rates. Given these consequences of grade retention and given the aforementioned pure ethnic gaps in grade retention, we believe that decision makers in school should use this instrument with care. Furthermore, we conclude that in the design of the optimal retention policy, the interaction between retention and students' abilities should be taken into account.

A third key result is that, for unemployed school-leavers, accepting a job for which one is overeducated substantially delays the transition to a first adequate job. Since overeducation is not a stepping stone to an adequate job, it may pay for young graduates to search for jobs selectively and therefore reject jobs for which they are overeducated. In Chapter 5 we therefore argued that policy makers should take care in not providing incentives to young unemployed graduates to accept any job too early in the unemployment spell, since this may induce persistent mismatch between the acquired qualifications and those required in the job. This is not only costly for the young graduates, but also for society as a whole. Consequently, the short-term benefits in terms of job transitions that can be induced by a tightening of entitlement conditions to unemployment benefits, such as the Belgian government implemented in 2012, must be carefully traded-off against the long-term costs in terms of qualification mismatch in the job. This is especially the case if this tightening targets the short-term unemployed for whom the costs of waiting for an adequate job are not yet counterbalanced by the costs of the risk of not finding any job at all.

[1] According to VDAB (2009b), besides the relatively size of the pool of adequately skilled workers and the wage level, the bottleneck status of an occupation is driven by its working conditions.

We now briefly review some directions for future research. First, in order to develop adequate policy action to tackle pure ethnic gaps in educational attainment one should identify the exact mechanisms behind these gaps. In Chapter 2 we discussed the most important candidates in this respect among which discrimination, ethnic differences in preferences and expectations and class and school segregation of migrants. However, our model is not capable to identify their relative importance. Therefore, more research is needed, potentially in other fields, to uncover the relative importance of these mechanisms in explaining ethnic gaps in educational attainment in Flanders and abroad. We hope that our proposed methodology, by which we can identify the instants at which these gaps emerge, can be a useful tool to target this research on the key schooling years. Second, Chapter 2 innovated by explaining observed ethnic gaps not only in educational attainment but also in the time it takes to realise this attainment. As mentioned before, by doing this it is found that important pure ethnic gaps exist in grade retention. We suggest that similar pure ethnic gaps might arise with respect to other measures of educational achievement within a particular level of educational attainment, such as scores on standardised tests of achievement. Identifying these gaps would be of great relevance in view of policy action aimed at realising equal opportunities in education. Third, we believe the important empirical evidence for a negative relationship between labour market discrimination and labour market tightness in Flanders we presented in Chapter 3 should be complemented with evidence from other countries. To this end, one could re-inspect the data gathered in similar large scale field experiments on ethnic discrimination with respect to this relationship.

References

Abbring, J. and van den Berg, G. (2003). The non-parametric identification of treatment effects in duration models. *Econometrica*, 71:1491–1517.

Agirdag, O., Houtte, M. V., and Avermaet, P. V. (2011). Ethnic school context and the national and sub-national identifications of pupils. *Ethnic and Racial Studies*, 34:357–378.

Ahmed, A., Andersson, L., and Hammerstedt, M. (2011). Are gays and lesbians discriminated against in the hiring situation? Institute for Labour Market Policy Evaluation Working Paper Series 21.

Akerlof, G. (1997). Social distance and social decisions. *Econometrica*, 65:1005–1027.

Albert, R., Escot, L., and Fernandez-Cornejo, J. (2011). A field experiment to study sex and age discrimination in the Madrid labour market. *International Journal of Human Resource Management*, 22:351–375.

Albrecht, J. and Vroman, S. (2002). A matching model with endogenous skill requirements. *International Economic Review*, 43:283–305.

Aldashev, A., Gernandt, J., and Thomsen, S. (2009). Language usage, participation, employment and earnings evidence for foreigners in West Germany with multiple sources of selection. *Labour Economics*, 16:330–341.

Alexander, K., Entwisle, D., and Dauber, S. (1994). *On the success of failure*. Cambridge University Press, New York.

Allen, J. and van der Velden, R. (2001). Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search. *Oxford Economic Papers*, 5:434–452.

Altwicker-Hámori, S. and Köllő, J. (2012). Whose children gain from starting school later? evidence from Hungary. *Educational Research and Evaluation*, 18:459–488.

Angrist, J. and Krueger, A. (1991). Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics*, 106:979–1074.

Arrijn, P., Feld, S., and Nayer, A. (1998). *Discrimination in access to employment on grounds of foreign origin: the case of Belgium*. International Labour Office, Geneva.

Arulampalam, W. (2001). Is unemployment really scarring? effects of unemployment experiences on wages. *Economic Journal*, 111:585–606.

Ashenfelter, O. (1970). Changes in labor market discrimination over time. *Journal of Human Resources*, 5:403–430.

Ashenfelter, O. and Hannan, T. (1986). Sex discrimination and product market competition: The case of the banking industry. *Quarterly Journal of Economics*, 101:149–173.

Baert, S., Cockx, B., and Verhaest, D. (2012). Overeducation at the start of the career: stepping stone or trap? IZA Discussion Paper Series 6562.

Baker, M. and Melino, A. (2000). Duration dependence and nonparametric heterogeneity: A monte carlo study. *Journal of Econometrics*, 96:357–393.

Battu, H., Belfield, C., and Sloane, P. (1999). Overeducation among graduates: A cohort view. *Education Economics*, 7:21–38.

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

Becker, G. (1957). *The economics of discrimination*. University of Chicago Press, Chicago.

Bedard, K. and Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *Quarterly Journal of Economics*, 121:1437–1472.

Belot, M. and Vandenberghe, V. (2013). Evaluating the “threat” effects of grade repetition. *Education Economics*. Forthcoming.

Belzil, C. and Poinas, F. (2010). Education and early career outcomes of second-generation immigrants in France. *Labour Economics*, 17:101–110.

Berson, C. (2012). Does competition induce hiring equity? Documents de travail du Centre d'Economie de la Sorbonne 12019.

Bertrand, M. and Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? a field experiment on labor market discrimination. *American Economic Review*, 94:991–1013.

Biddle, J. and Hamermesh, D. (2012). Wage discrimination over the business cycle. IZA Discussion Paper Series 6445.

Black, D. (1995). Discrimination in an equilibrium search model. *Journal of Labor Economics*, 13:309–334.

Black, S. and Brainerd, E. (2004). Importing equality? the impact of globalization on gender discrimination. *Industrial and Labor Relations Review*, 57:540–559.

Black, S. and Strahan, P. (2001). The division of spoils: Rent-sharing and discrimination in a regulated industry. *American Economic Review*, 91:814–831.

Blanchflower, D. and Freeman, F. (2000). The declining economic status of young workers in OECD countries. In Blanchflower, D. and Freeman, F., editors, *Youth Employment and Joblessness in Advanced Countries*. University of Chicago Press, Chicago.

Booth, A., Leigh, A., and Varganova, E. (2012). Does racial and ethnic discrimination vary across minority groups? evidence from a field experiment. *Oxford Bulletin of Economics and Statistics*, 74:547–573.

Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in Western society*. Wiley-Interscience, New York.

Bratti, M., Cappellari, L., Groh-Samberg, O., and Lohmann, H. (2012). School tracking and intergenerational transmission. In Ermisch, J., Jäntti, M., and Smeeding, T., editors, *From parents to children. The intergenerational transmission of advantage*. Russell Sage Foundation, New York.

Breen, R. and Goldthorpe, J. (1997). Explaining educational differentials: Towards a formal rational action theory. *Rationality and Society*, 9:275–305.

Browman, L. (2005). Grade retention: Is it a help or hindrance to student academic success? *Preventing School Failure*, 49:42–46.

Büchel, F. and Mertens, A. (2004). Overeducation, undereducation, and the theory of career mobility. *Applied Economics*, 36:803–816.

Bursell, M. (2007). What's in a name? a field experiment test for the existence of ethnic discrimination in the hiring process. Stockholm University Linnaeus Center for Integration Studies Working Paper Series 7.

Byrd, R., Weitzman, M., and Auinger, P. (1997). Increased behavior problems associated with delayed school entry and delayed school progress. *Pediatrics*, 100:654–661.

Cahuc, P. and Zylberberg, A. (2004). *Labor Economics*. MIT Press, Massachusetts.

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

Cameron, S. and Heckman, J. (2001). The dynamics of educational attainment for black, hispanic and white males. *Journal of Political Economy*, 109:455–499.

Capéau, B., Eeman, L., Groenez, S., and Lamberts, S. (2012a). Standardised scores as a way to measure and compare discrimination across dimensions. Ecore Discussion Papers 59.

Capéau, B., Eeman, L., Groenez, S., and Lamberts, S. (2012b). Two concepts of discrimination: Inequality of opportunity versus unequal treatment of equals. Ecore Discussion Papers 58.

Card, D. (2005). Is the new immigration really so bad? *Economic Journal*, 115:300–323.

Carlsson, M. and Eriksson, S. (2012). Do reported attitudes towards immigrants predict ethnic discrimination. Uppsala University Department of Economics Working Paper Series 6.

Carlsson, M. and Rooth, D.-O. (2007). Evidence of ethnic discrimination in the Swedish labor market using experimental data. *Labour Economics*, 14:716–729.

Carlsson, M. and Rooth, D.-O. (2008). Is it your foreign name or foreign qualifications? an experimental study of ethnic discrimination in hiring. IZA Discussion Paper Series 3810.

Carneiro, P., Hansen, K., and Heckman, J. (2003). Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review*, 44:361–422.

CBS (2001). *Standaard Beroepenclassificatie 1992—Editie 2001.* Central Bureau of Statistics, Heerlen/Voorburg.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies*, 47:225–238.

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

Chiswick, B. (2008). The economics of language: An introduction and overview. IZA Discussion Paper Series 3568.

Chiswick, B. and DebBurman, N. (2004). Educational attainment: analysis by immigrant generation. *Economics of Education Review*, 23:361–379.

Cockx, B., Göbel, C., and Robin, S. (2013). Can income support for part-time workers serve as a stepping stone to regular jobs? an application to young long-term unemployed women. *Empirical Economics*, 44:189–229.

Colding, B. (2006). A dynamic analysis of educational progression of children of immigrants. *Labour Economics*, 13:479–492.

Colding, B., Husted, L., and Hummelgaard, H. (2009). Educational progression of second-generation immigrants and immigrant children. *Economics of Education Review*, 28:434–443.

Constant, A., Krause, A., Rinne, U., and Zimmermann, K. (2010). Economic preferences and attitudes of the unemployed: Are natives and second generation migrants alike? *International Journal of Manpower*, 32:825–851.

Couppié, T. and Mansuy, M. (2003). Young people and new entrants in european labour markets: the timing of gradual integration. In Müller, W. and Gangl., M., editors, *Transitions from Education to Work in Europe*. Oxford University Press, Oxford.

de Grip, A., Bosma, H., Willems, D., and van Boxtel, M. (2008). Job-worker mismatch and cognitive decline. *Oxford Economic Papers*, 60:237–253.

De Ro, J. (2008). *Education in Flanders. A broad view on the Flemish education landscape.* Agency for Educational Communication Publications, Brussels.

DellaVigna, S. and Paserman, M. (2005). Job search and impatience. *Journal of Labor Economics*, 23:527–588.

D'Haultfœuille, X. (2010). A new instrumental method for dealing with endogenous selection. *Journal of Econometrics*, 154:1–15.

Dolado, J., Jansen, M., and Jimeno, J. (2009). On-the-job search in a matching model with heterogeneous jobs and workers. *Economic Journal*, 119:200–228.

Dolton, P. and Silles, M. (2008). The effects of over-education on earnings in the graduate labour market. *Economics of Education Review*, 27:125–139.

Dolton, P. and Vignoles, A. (2000). The incidence and effects of overeducation in the U.K. graduate labour market. *Economics of Education Review*, 19:179–180.

Drydakis, N. (2009). Sexual orientation discrimination in the labour market. *Labour Economics*, 16:364–372.

Drydakis, N. and Vlassis, M. (2010). Ethnic discrimination in the greek labour market: occupational access, insurance coverage and wage offers. *Manchester School*, 78:201–218.

Duquet, N., Glorieux, I., Laurijssen, I., and Dorsselaer, Y. V. (2006). *Wit krijt schrijft beter.* Garant, Antwerpen.

Dustmann, C. (2004). Parental background, secondary school track choice, and wages. *Oxford Economic Papers*, 56:209–230.

Dustmann, C. and Fabbri, F. (2003). Language proficiency and labour market performance of immigrants in the UK. *Economic Journal*, 113:695–717.

Dustmann, C., Machin, S., and Schonberg, U. (2010). Ethnicity and educational achievement in compulsory schooling. *Economic Journal*, 120:272–297.

Eckstein, Z. and Wolpin, K. (1998). Estimating the effect of racial discrimination on first job wage offers. *Review of Economics and Statistics*, 81:384–392.

Eide, E. and Goldhaber, D. (2005). Grade retention: What are the costs and benefits? *Journal of Education Finance*, 31:195–214.

Eide, E. and Showalter, M. (2001). The effect of grade retention on educational and labor market outcomes. *Economics of Education Review*, 20:563–576.

Erikson, R., Goldthorpe, J., Jackson, M., and Cox, D. (2010). On class differentials in educational attainment. *Proceedings of the National Academy of Sciences of the United States of America*, 102:9730–9733.

European Foundation for the Improvement of Living and Working Conditions (2012). *NEETs Young people not in employment, education or training: Characteristics, costs and policy responses in Europe*. Publications Office of the European Union, Luxembourg.

Euwals, R., Dagevos, J., Gijsberts, M., and Roodenburg, H. (2010). Citizenship and labor market position: Turkish immigrants in Germany and the Netherlands. *International Migration Review*, 44:513–538.

Falter, J., Ferro-Luzzi, G., and Sbergami, F. (2011). The effect of parental background on track choices and wages. *Swiss Journal of Economics and Statistics*, 147:157–180.

Fang, H. and Silverman, D. (2009). Time-inconsistency and welfare program participation: Evidence from the NLSY. *International Economic Review*, 50:1043–1077.

Filippin, A. (2009). Can workers's expectations account for the persistence of discrimination? IZA Discussion Papers 4490.

Frederick, S., Loewenstein, G., and O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, 40:351–401.

Fredriksson, P. and Öckert, B. (2006). Is early learning really more productive? the effect of school starting age on school and labor market performance. IFAU Working Papers 12.

Freeman, R. (1973). Changes in the labor market for black americans, 1948–72. Brookings papers on economic activity.

Fruehwirth, J., Navarro, S., and Takahashi, Y. (2011). How the timing of grade retention affects outcomes: Identification and estimation of time-varying treatment effects. CIBC Centre for Human Capital and Productivity Working Papers 7.

Gamoran, A., Nystrand, M., Berens, M., and LePore, P. (1995). An organizational analysis of the effects of ability grouping. *American Educational Research Journal*, 32:687–715.

Gangl, W. (2003). The structure of labour market entry in Europe: a typological analysis. In Müller, W. and Gangl., M., editors, *Transitions from Education to Work in Europe*. Oxford University Press, Oxford.

Gaure, S., Røed, K., and Zhang, T. (2007). Time and causality: A Monte Carlo assessment of the timing-of-events approach. *Journal of Econometrics*, 141:1159–1195.

Giret, J. and Hatot, C. (2001). Mesurer le déclassement à l'embauche: l'exemple de DUT et de BTS. *Formation Emploi*, 76:59–73.

Goldthorpe, J. (1996). Class analysis and the reorientation of class theory: The case of persisting differentials in educational attainment. *British Journal of Sociology*, 47:481–505.

Goos, M., Manning, A., and Salomons, A. (2009). The polarization of the European labor market. *American Economic Review Papers and Proceedings*, 99:58–63.

Green, F. and McIntosh, S. (2007). Is there a genuine under-utilization of skills among the over-qualified? *Applied Economics*, 39:427–439.

Green, F. and Zhu, Y. (2010). Overqualification, job dissatisfaction, and increasing dispersion in the returns to graduate education. *Oxford Economic Papers*, 62:740–763.

Greene, J. and Winters, M. (2007). Revisiting grade retention: An evaluation of Florida's test-based promotion policy. *Education Finance and Policy*, 2:319–340.

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

Gregg, P. and Tominey, E. (2005). The wage scar from male youth unemployment. *Labour Economics*, 12:487–509.

Groeneveld, S. and Hartog, J. (2004). Overeducation, wages and promotions within the firm. *Labour Economics*, 11:701–714.

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

Guiso, L., Monte, F., Sapienza, P., and Zingales, L. (2008). Culture, gender, and math. *Science*, 320:1164–1165.

Hagy, A. and Staniec, J. (2002). Immigrant status, race, and institutional choice in higher education. *Economics of Education Review*, 21:381–392.

Hámori, S. (2007). The effect of school starting age on academic performance in Hungary. Budapest Working Papers on the Labour Market No. 2007/2.

Hartlaub, V. and Schneider, T. (2012). Educational choice and risk aversion: How important is structural vs. individual risk aversion? SOEP Paper on Multidisciplinary Panel Data Research 433.

Hartog, J. (2012). Overeducation and earnings: where are we, where should we go? *Economics of Education Review*, 19:131–147.

Haveman, R. and Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature*, 33:1829–1878.

Heckman, J. (1981a). The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In Manski, C. and McFadden, D., editors, *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press, Cambridge.

Heckman, J. (1981b). Statistical models for discrete panel data. In Manski, C. and McFadden, D., editors, *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press, Cambridge.

Heckman, J. (1998). Detecting discrimination. *Journal of Economic Perspectives*, 12:101–116.

Heckman, J. (2011). The American family in black and white: A post-racial strategy for improving skills to promote equality. *Daedalus*, 140:70–89.

Heckman, J. and Navarro, S. (2007). Dynamic discrete choice and dynamic treatment effects. *Journal of Econometrics*, 136:341–396.

Heckman, J. and Siegelman, P. (1993). *The Urban Institute Audit Studies: Their Methods and Findings*. Urban Institute, Washington DC.

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

Heckman, J., Urzua, S., and Vytlacil, E. (2006). Understanding instrumental variables in models with essential heterogeneity. *Review of Economics and Statistics*, 88:389–432.

Hellerstein, J., Neumark, D., and Troske, K. (2002). Market forces and sex discrimination. *Journal of Human Resources*, 37:353–380.

Hennessey, M., Rumrill, P., Fitzgerald, S., and Roessler, R. (2008). Disadvantagement-related correlates of career optimism among college and university students with disabilities. *Work, a Journal of Prevention Assessment and Rehabilitation*, 30:483–492.

Holzer, H. (1987). Job search by employed and unemployed youth. *Industrial and Labor Relations Review*, 40:601–611.

Jacob, B. and Lefgren, L. (2004). Remedial education and student achievement: A regressor-discontinuity analysis. *Review of Economics and Statistics*, 86:226–244.

Jacob, B. and Lefgren, L. (2009). The effect of grade retention on high school completion. *American Economic Journal: Applied Economics*, 1:33–58.

Jimerson, S. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review*, 30:420–437.

Kaas, L. and Manger, C. (2012). Ethnic discrimination in Germany's labour market: a field experiment. *German Economic Review*, 13:1–20.

Keane, M. and Wolpin, K. (1997). The career decisions of young men. *Journal of Political Economy*, 105:473–522.

Korpi, T. and Tåhlin, M. (2009). Education mismatch, wages, and wage growth: overeducation in Sweden 1974–2000. *Labour Economics*, 16:183–193.

Lancaster, T. (1990). *The Econometric Analysis of Transition Data*. Cambridge University Press, Cambridge.

Leuven, E. and Oosterbeek, H. (2011). Overeducation and mismatch in the labor market. In Hanushek, E., Machin, S., and Woessman, L., editors, *Handbook of the Economics of Education*. North Holland, Amsterdam.

Machado, J. and Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, 20:445–465.

Mahuteau, S. and Juanankar, P. (2008). Do migrants get good jobs in Australia? the role of ethnic networks in job search. *Economic Record*, 84:115–130.

Manacorda, M. (2012). The cost of grade retention. *The Review of Economics and Statistics*, 94:596–606.

Manning, A. (2003). The real thin theory: Monopsony in modern labour markets. *Labour Economics*, 10:105–131.

Mavromaras, K. and McGuinness, S. (2012). Overskilling dynamics and education pathways. *Economics of Education Review*, 31:619–628.

Mavromaras, K., McGuinness, S., O'Leary, N., Sloane, P., and Wei, Z. (2013). Job mismatches and labour market outcomes: Panel evidence on Australian university graduates. *Economic Record*. In press.

McCormick, B. (1990). A theory of signalling during job search, employment efficiency, and stigmatised jobs. *Review of Economic Studies*, 57:299–313.

McEwan, P. and Shapiro, J. (2008). The benefits of delayed primary school enrollment:

Discontinuity estimates using exact birth dates. *Journal of Human Resources*, 43:1–29.

McGinnity, F., Nelson, J., Lynn, P., and Quinn, E. (2009). *Discrimination in Recruitment: Evidence from a Field Experiment.* The Equality Authority, Dublin.

McGuinness, S. (2006). Overeducation in the labour market. *Journal of Economic Surveys*, 20:387–418.

Mroz, T. (1999). Discrete factor approximations in simultaneous equation models: Estimating the impact of a dummy endogenous variable on a continuous outcome. *Journal of Econometrics*, 92:233–274.

Mroz, T. and Picone, G. (2011). A multiple state duration model with endogenous treatment. HEDG Working Paper 11/19.

Mroz, T. and Savage, T. (2006). The long-term effects of youth unemployment. *Journal of Human Resources*, 41:259–293.

Neumark, D. (2012). Detecting discrimination in audit and correspondence studies. *Journal of Human Resources*, 47:1128–1157.

Nonneman, W. (2012). School achievement and failure of immigrant children in Flanders. Working papers of the Faculty of Applied Economics of the University of Antwerp 2012008.

OECD (2004). *Learning for tomorrow's world—First results from PISA 2003.* OECD, Paris.

OECD (2007). *Des emplois pour les jeunes.* OECD, Paris.

OECD (2008). *Jobs for Immigrants. Labour Market Integration in France, Belgium, the Netherlands and Portugal.* OECD, Paris.

OECD (2010). *Off to a good start? Jobs for youth.* OECD, Paris.

OECD (2011). *Grade Retention during Compulsory Education in Europe: Regulations and Statistics.* OECD, Paris.

Ortiz, E. and Dehon, C. (2008). What are the factors of success at university? a case study in Belgium. *Cesifo Economic Studies*, 54:121–148.

Pager, D. (2007). The use of field experiments for studies of employment discrimination: contributions, critiques, and directions for the future. *Annals of the American Academy of Political and Social Science*, 609:104–133.

Peoples, J. and Saunders, L. (1993). Trucking deregulation and the black/white wage gap. *Industrial and Labor Relations Review*, 47:23–35.

Petit, P. (2007). The effects of age and family constraints on gender hiring discrimination: A field experiment in the French financial sector. *Labour Economics*, 14:371–391.

Pissarides, C. (1994). Search unemployment with on-the-job search. *Review of Economic Studies*, 61:457–475.

Pollmann-Schult, M. and Büchel, F. (2004). Career prospects of overeducated workers in West Germany. *European Sociological Review*, 20:321–331.

Pozzoli, D. (2009). The transition to work for Italian university graduates. *Labour*, 23:131–169.

Riach, P. and Rich, J. (2002). Field experiments of discrimination in the market place. *Economic Journal*, 112:480–518.

Riach, P. and Rich, J. (2004). Deceptive field experiments of discrimination: Are they ethical? *Kyklos*, 57:447–470.

Robst, J. (1995). Career mobility, job match and overeducation. *Eastern Economic Journal*, 21:539–550.

Rooth, D.-O. (2007). Evidence of unequal treatment in hiring against obese applicants: A field experiment. IZA Discussion Paper Series 2775.

Rosén, A. (2003). Search, bargaining, and employer discrimination. *Journal of Labor Economics*, 21:807–829.

Rubb, S. (2003). Overeducation: a short or long run phenomenon for individuals? *Economics of Education Review*, 22:389–394.

Rubb, S. (2006). Educational mismatches and earnings: extensions of occupational mobility theory and evidence of human capital depreciation. *Education Economics*, 14:135–154.

Ryan, P. (2001). The school-to-work transition: A cross-national perspective. *Journal of Economic Literature*, 39:34–92.

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

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

Skott, P. (2006). Wage inequality and overeducation in a model with efficiency wages. *Canadian Journal of Economics*, 39:94–123.

SONAR (2000a). *Hoe maken de jongeren de overgang van school naar werk? Basisrapportering Cohorte 1978 (eerste golf)*. Steunpunt WAV, Brussels.

SONAR (2000b). *Jongeren in transitie*. Steunpunt WAV, Brussels.

SONAR (2005). *Hoe maken de jongeren de overgang van school naar werk? Basisrapportering Cohorte 1980 (eerste golf)*. Steunpunt WAV, Brussels.

Strøm, B. (2004). Student achievement and birthday effects. Paper presented at the CESifo-Harvard University/PEPG Conference on “Schooling and Human Capital in the Global Economy: Revisiting the Equity-Efficiency Quandary”, CESifo Conference Center, Munich.

Taubman, P. and Wales, T. (1974). *Higher Education and Earnings: College as an Investment and Screening Device*. NBER Books, Massachusetts.

Thurow, L. (1972). Education and economic inequality. *Public Interest*, 28:66–81.

Tsang, M. (1987). The impact of underutilization of education on productivity: a case study of the U.S. Bell Companies. *Economics of Education Review*, 4:93–104.

Van de gaer, E., Pustjens, H., Van Damme, J., and De Munter, A. (2006). Tracking and the effects of school-related attitudes on the language achievement of boys and girls. *British Journal of Sociology of Education*, 27:293–309.

van der Klaauw, B. and van Vuuren, A. (2010). Job search and academic achievement. *European Economic Review*, 54:298–320.

Van Houtte, M. (2004). Tracking effects on school achievement: A quantitative explanation in terms of the academic culture of school staff. *American Journal of Education*, 110:354–388.

Van Houtte, M., Demanet, J., and Stevens, P. (2012). Self-esteem of academic and vocational students: Does within-school tracking sharpen the difference? *Acta Sociologica*, 55:73–89.

van Ours, J. and Veenman, J. (2003). The educational attainment of second-generation immigrants in the Netherlands. *Journal of Population Economics*, 16:739–753.

van Smoorenburg, M. and van der Velden, R. (2000). The training of school-leavers, complementarity or substitution? *Economics of Education Review*, 19:207–217.

VDAB (2009a). *Allochtonen op de Vlaamse arbeidsmarkt.* VDAB Studiedienst, Brussels.

VDAB (2009b). *Analyse Vacatures 2009. Knelpuntberoepen.* VDAB Studiedienst, Brussels.

Verhaest, D. and Omey, E. (2009). Objective over-education and worker well-being: A shadow price approach. *Journal of Economic Psychology*, 30:469–481.

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

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

Verhaest, D. and Omey, E. (2013). The relationship between formal education and skill acquisition in young workers' first jobs. *Manchester School*, 81:638–659.

Verhaest, D. and van der Velden, R. (2013). Cross-country differences in graduate overeducation. *European Sociological Review*, 29:642–653.

Weichselbaumer, D. (2003). Sexual orientation discrimination in hiring. *Labour Economics*, 10:629–642.

Weichselbaumer, D. (2004). Is it sex or personality. the impact of sex stereotypes on discrimination in applicant selection. *Eastern Economic Journal*, 20:159–186.

Winters, P., de Janvry, A., and Sadoulet, E. (2001). Family and community networks in Mexico–U.S. migration. *Journal of Human Resources*, 36:159–184.

Wood, M., Hales, J., Purdon, S., Sejersen, T., and Hayllar, O. (2009). A test for racial discrimination in recruitment practice in british cities. DWP Research Reports 607.

Xia, N. and Kirby, N. (2009). Retaining students in grade. a literature review of the effects of retention on students' academic and nonacademic outcomes. Nber working papers.

Yamauchi, F. and Tanabe, S. (2008). Nonmarket networks among migrants: evidence from metropolitan Bangkok, Thailand. *Journal of Population Economics*, 21:649–664.

Zaiceva, A. and Zimmermann, K. (2010). Do ethnic minorities 'stretch' their time? UK household evidence on multitasking. *Review of Economics of the Household*, 9:181–206.

Zenou, Y. (2011). Spatial versus social mismatch: The strength of weak ties. CEPR Discussion Papers 8244.