IZA

DISCUSSION PAPER SERIES

IZA DP No. 7262

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

Stijn Baert
Bart Cockx

March 2013

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

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

Stijn Baert
Ghent University

Bart Cockx
Ghent University,
Université catholique de Louvain, CESifo and IZA

Discussion Paper No. 7262
March 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 7262
March 2013

ABSTRACT

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

This article decomposes the observed gaps in educational attainment and school-to-work transitions between grandchildren of natives and immigrants in Belgium into (i) differences in observed family endowments and (ii) a residual "pure ethnic gap". It 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 observed family endowments. The pure ethnic gap in educational attainment is found to be small if delays are neglected, but substantial if not and for school-to-work transitions. It is shown that more than 20% of the pure ethnic gap in graduating from secondary school without delay originates in tenth grade. Language usage explains only part of the gap in school-to-work transitions for low educated.

JEL Classification: C35, J15, J70

Keywords: dynamic discrete choice, dynamic selection bias, educational attainment, school-to-work transitions, ethnic minorities, discrimination

Corresponding author:

Bart Cockx
Sherppa – Ghent University
Tweekerkenstraat 2
B-9000 Gent
Belgium
E-mail: Bart.Cockx@UGent.be

* We thank Miles Corak, Sabien Dobbelaere, Bas van der Klaauw, Jan van Ours, Dieter Verhaest and the seminar participants at Ghent University, Université catholique de Louvain, VU University Amsterdam, Tilburg University, the Belgian Day for Labor Economists 2011, the IZA European Summer School in Labor Economics 2011, the CESifo Area Conference on Employment and Social Protection 2011, the International Workshop on Applied Economics of Education 2011, the European Network on Transitions in Youth Workshop 2011, the Day of the Scientific Economic Research of the Flemish Society of Economics 2011 and the annual conference of the European Society for Population Economics 2012 for their insightful comments and suggestions, which have helped to improve this study considerably. In addition, we are grateful to the Steunpunt SSL of the Flemish government for the used SONAR data. The authors assume the sole scientific responsibility of the present work.

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 (see, e.g., Arulampalam, 2001; Gregg and Tominey, 2005; Mroz and Savage, 2006). Therefore, not surprisingly, the OECD (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 behavior or by discrimination. However, if these gaps just mirrors different family endowments that result in different levels of educational attainment and therefore in different labor 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 labor 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 labor 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 labor market transitions in isolation from related transitions, such as the decision to enroll 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

[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.

endowments pass the final evaluation at the end of a particular grade and continue schooling only if their unobserved endowments are sufficiently favorable. 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 modeling, 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 is 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 (e.g. 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 paper 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 behavior 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,[2] it is costly if it eventually induces pupils to enter the labor 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 modeling 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.[3]

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 analyzing 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, i.e. passing or failing, and from the "secondary effects" in the educational choices, i.e. 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) realized 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

[2] Most studies actually find that retention has a negative impact on educational achievement (see Xia and Kirby, 2009, for a survey). Only Jacob and Lefgren (2004) and D'Haultfoeuille (2010) find short-term positive effects on educational performance.

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

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 labor 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 labor market perspectives. Therefore, it may matter to control for language usage in an analysis of the determinants of ethnic gaps in schooling and labor market outcomes.

This article is structured in the following way. *Section 2* summarizes the institutional setting: the educational system and the youth labor market in Belgium. The next section describes the dataset and provides descriptive statistics that motivate our analysis. *Section 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. The Institutional Setting: Education and Youth Labor Market[4]

In Belgium the language communities (Flemish and French) are in charge of the organization of the educational system, while labor market regulation is in the period of analysis mostly organized at the national level. Since the data we analyze 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[5] 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

[4] For more details on the educational system, see De Ro (2008).

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

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 behavioral 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 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 specialization, we assimilate it in this study as part of higher (tertiary) education. Students with a secondary school diploma can enroll directly, without any entry exam,[6] into higher (tertiary) education, i.e. 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 to €600 in 2012, depending on the parents' income) in higher education. Twenty-two colleges and seven universities are spread over less than 14,000 km² resulting in a high regional diffusion of providers of tertiary education.

There is no compulsory military service in Flanders and school-leavers enter the labor 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

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

with the registration at the employment office after leaving school. Labor regulation distinguishes between two types of labor 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.

3. Data and Some Facts

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. This database contains exceptionally rich information on both the educational career and the start of the labor 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 labor market status[7] from the moment secondary school is entered,[8] 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.

3.2 Motivating Gaps

Throughout this article, 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", i.e. youths whose grandmother on mother's side possesses the Belgian nationality (8,091 individuals). On the other hand, we consider "immigrants", i.e. youth whose grandmother on mother's side neither has the Belgian nationality nor any other Western[9] nationality (545 individuals). From these sub-populations we select those who resided and went to school in Flanders from the start of

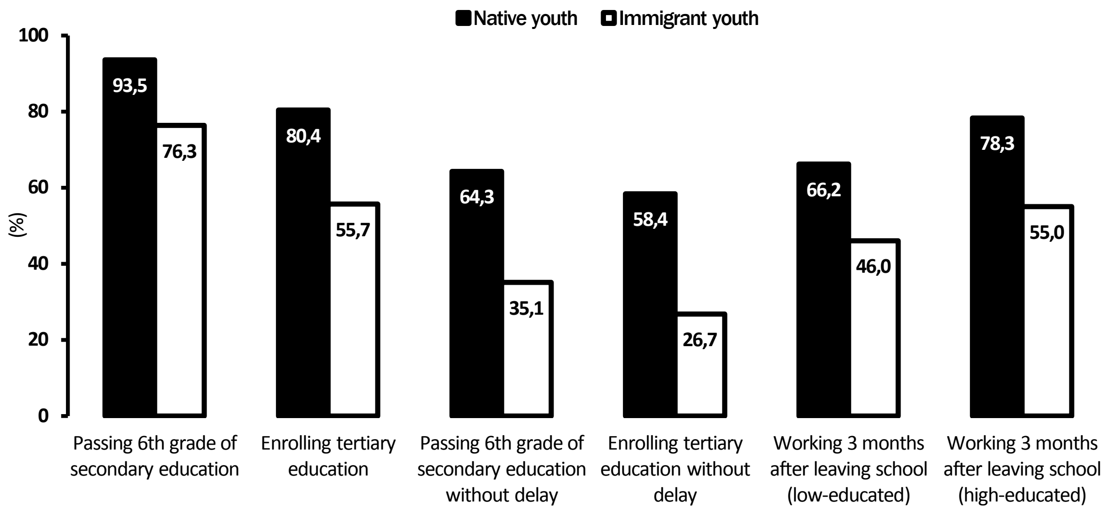
[7] An individual is employed when holding a job of at least one hour a week and during at least one month. Part-time jobs held by students in the vocational track are not considered as employment, but as part of the educational career.

[8] 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.

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

nursery school onwards. Dropping individuals with (i) missing explanatory variables; (ii) inconsistent school registrations and (iii) years of special education (cf. *Section 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 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" or "immigrants" according to the definitions in this section.

Figure 1: Schooling and School-to-Work Transition Outcomes by Ethnic Group



Source: own calculations based on 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.

Figure 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 realizing them (without specifying any potential delay) and realizing them without schooling delay. Those in which schooling delay is left unspecified are usually considered in the literature. However, as argued in the *Introduction*, 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 labor 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.

Second, we report, conditional on observed school attainment, the gap for being employed 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* 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 enroll 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%.

3.3 Explanatory Variables

In this subsection, we describe the explanatory variables used for each modeled 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

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

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 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 labor market success for minority youth.[12]

Table 1 reports descriptive statistics of these variables by ethnicity. These statistics confirm that immigrant youths generally are characterized by more unfavorable 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 labor 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, neighborhood characteristics, indicators of regional labor 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

---

[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 the *Introduction*.

[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.

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 labor market conditions. The unemployment rate of the 24 to 64 year old male population proxies the labor market conditions of the (usually) male breadwinner 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 A–1* and *Table A–2* reported in *Appendix A*.

Table 1: Summary Statistics of the Exogenous Individual Explanatory Variables by Ethnic Groups

	Flemish youth		**Immigrant youth**	
	Mean	Standard deviation	Mean	Standard deviation
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. E.g. 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.

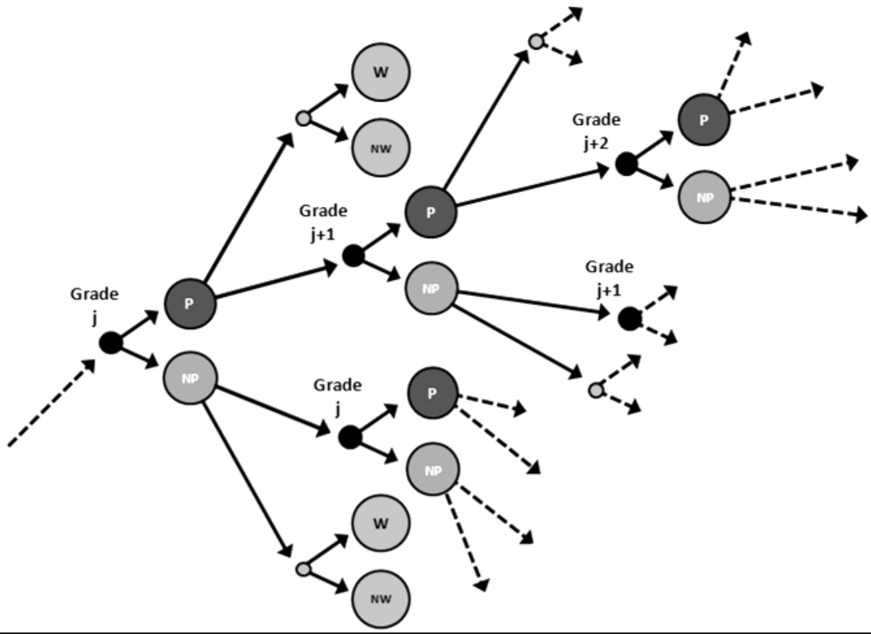
4. Methodology

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 enrolls 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 modeling schooling outcomes as from the start of primary school instead of from the end of compulsory education as researchers 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, i.e. 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 realize this decomposition in *Section 4.2*. We first present the econometric model.

Figure 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).

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* 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*shows a graphical representation of our modeling 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 passing 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 (cf. the *Introduction*). 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, since students with successful schooling achievements possess more favorable 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.

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, i.e. 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

---

[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 (cf. infra) we adopt employment with a permanent contract two years after leaving school as the labor market outcome.

[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.

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 consequence 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 realized (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 primary 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:

$$\hat{c}^{O_g}_{g,t} = c \in C^{O_g} \quad if \quad \omega^{O_g}_c < U^{O_g}_{g,t,c} \le \omega^{O_g}_{c+1}, \tag{1}$$

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 $\omega^{O_g}_c$ and $\omega^{O_g}_{c+1}$ are threshold utilities that determine the ordered choice ($\omega^{O_g}_0 \equiv -\infty$ and $\omega^{O_g}_{n^{O_g}+1} \equiv +\infty$).[18] As advocated by, e.g., Heckman (1981) and adopted by other authors, we

[18] In the case of a binary choice the threshold $\omega^{O_g}_1$ 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.

approximate this $U_{g,t,c}^{O_g}$ by a linear index:

$$U_{g,t,c}^{O_g} = \alpha_g^{O_g} + \boldsymbol{Z'_t}\boldsymbol{\beta^{O_g}} + \gamma^{O_g} V_t + \nu_{g,t,c}^{O_g}, \tag{2}$$

where $\alpha_g^{O_g}$ is a parameter that depends on the grade in which the outcome type O_g is considered,[19,20] $\boldsymbol{Z'_t}$ is a $1 \times M$ vector representing the M number of (possibly time-varying) observed strictly exogenous variables, $\boldsymbol{\beta^{O_g}}$ 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 characterized 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 generalize 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:

$$\nu_{g,t,c}^{O_g} = \delta^{O_g}\eta + \varphi^{O_g} V_t \eta + \varepsilon_{g,t,c}^{O_g}, \tag{3}$$

in which δ^{O_g} and φ^{O_g} are outcome type specific coefficients and $\varepsilon_{g,t,c}^{O_g}$ 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 (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, i.e. the number of years of delay at the start of primary school, is free of selection. This means that $\boldsymbol{Z'_t}$ should be independent of η for all g, t and

[19] The parameter corresponding to the first grade that can be observed within the outcome type is taken as the reference grade. It is normalized 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).

choice sets C^{O_g}. Note that this does not mean that *conditional on past choices* beyond the start of primary school $\boldsymbol{Z'_t}$ is independent of η, since, as mentioned in the *Introduction,* dynamic sorting will induce negative correlation between favorable observed determinants of the educational outcomes that we consider, and the unobserved η. This is because pupils with unfavorable observed endowments experience successful educational outcomes only if these unfavorable endowments are compensated for by favorable 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 improve in this respect on the existing literature by starting modeling 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 (i.e. no delay) during the period of mandatory schooling.[22]

We assume that $\varepsilon_{g,t,c}^{O_g}$ 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:

$$Pr(\hat{c}_{g,t}^{O_g} = c \mid \boldsymbol{Z'_t}, V_t, g, O_g, \eta; \boldsymbol{\theta}) = \frac{exp(\omega_{c+1}^{O_g} - \alpha_g^{O_g} - \boldsymbol{Z'_t}\boldsymbol{\beta}^{\boldsymbol{O_g}} - \delta^{O_g}\eta - \varphi^{O_g}V_t\eta)}{1 + exp(\omega_{c+1}^{O_g} - \alpha_g^{O_g} - \boldsymbol{Z'_t}\boldsymbol{\beta}^{\boldsymbol{O_g}} - \delta^{O_g}\eta - \varphi^{O_g}V_t\eta)} - \frac{exp(\omega_{c}^{O_g} - \alpha_g^{O_g} - \boldsymbol{Z'_t}\boldsymbol{\beta}^{\boldsymbol{O_g}} - \delta^{O_g}\eta - \varphi^{O_g}V_t\eta)}{1 + exp(\omega_{c}^{O_g} - \alpha_g^{O_g} - \boldsymbol{Z'_t}\boldsymbol{\beta}^{\boldsymbol{O_g}} - \delta^{O_g}\eta - \varphi^{O_g}V_t\eta)}, \quad (4)$$

in which we denote the vector of unknown parameters by $\boldsymbol{\theta}$. The likelihood contribution $\ell_i(\boldsymbol{Z'_{it}}, V_{it}, \eta; \boldsymbol{\theta})$ for any sampled individual, *conditional on the unobservable* η, is then constructed by the product of the probabilities of the school and labor market outcomes as expressed by (4) realized in each time period *t* between the start of primary school, and the labor 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 characterized 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:

---

[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.

$$p_k(\lambda) = \frac{exp(\lambda_k)}{\sum_{j=1}^{K} exp(\lambda_j)} \quad with \quad k = 1,2,...K, \boldsymbol{\lambda} \equiv [\lambda_1, \lambda_2, ..., \lambda_K]' \; and \; \lambda_1 = 0 \cdot \tag{5}$$

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

$$\ell_i(\boldsymbol{Z'_{it}}, V_{it}; \boldsymbol{\theta}, \boldsymbol{\lambda}) = \sum_{k=1}^{K} p_k(\boldsymbol{\lambda}) \ell_i(\boldsymbol{Z'_{it}}, V_{it}, \eta_k; \boldsymbol{\theta}). \tag{6}$$

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 maximized 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 minimizes an information criterion. In our case (cf. *Section 4.3*) the optimal choice minimizes both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

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 realize 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 $\boldsymbol{Z_N}$ and $\boldsymbol{Z_I}$ 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 $\hat{\boldsymbol{\theta}}_{\boldsymbol{N}}$ and $\hat{\boldsymbol{\theta}}_{\boldsymbol{I}}$ 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:

$$log\left(\frac{E_{\boldsymbol{Z_N}}\ Pr\left[\hat{c}_{g,t}^{O_g}=c\middle|\boldsymbol{Z}_{\boldsymbol{N}};\hat{\boldsymbol{\theta}}_{\boldsymbol{N}}\right]}{E_{\boldsymbol{Z_I}}\ Pr\left[\hat{c}_{g,t}^{O_g}=c\middle|\boldsymbol{Z}_{\boldsymbol{F}};\hat{\boldsymbol{\theta}}_{\boldsymbol{I}}\right]}\right), \tag{7}$$

where $Pr\left[\hat{c}_{g,t}^{O_g}=c\middle|.;.\right]$ is the probability that the particular outcome $\hat{c}_{g,t}^{O_g}=c$ [23] is realized according to the model simulation and $E_{\boldsymbol{Z}_N}$ and $E_{\boldsymbol{Z}_I}$ the expectations over the distributions of $\boldsymbol{Z}_{\boldsymbol{N}}$ respectively $\boldsymbol{Z}_{\boldsymbol{I}}$. Note that the existing literature expresses the gap in terms of the absolute difference in the probabilities of realization 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 *Appendix B*.

We now propose the following decomposition of the predicted ethnic gap, as expressed by the log odds ratio in *Equation (7)* into the sum of an "explained" and a "residual pure ethnic gap":

$$log\left(\frac{E_{\boldsymbol{Z_N}}\ Pr\left[\hat{c}_{g,t}^{O_g}=c\middle|\boldsymbol{Z}_{\boldsymbol{N}};\hat{\boldsymbol{\theta}}_{\boldsymbol{N}}\right]}{E_{\boldsymbol{Z_I}}\ Pr\left[\hat{c}_{g,t}^{O_g}=c\middle|\boldsymbol{Z}_{\boldsymbol{I}};\hat{\boldsymbol{\theta}}_{\boldsymbol{I}}\right]}\right)=log\left(\frac{E_{\boldsymbol{Z_N}}\ Pr\left[\hat{c}_{g,t}^{O_g}=c\middle|\boldsymbol{Z}_{\boldsymbol{N}};\hat{\boldsymbol{\theta}}_{\boldsymbol{N}}\right]}{E_{\boldsymbol{Z_I}}\ Pr\left[\hat{c}_{g,t}^{O_g}=c\middle|\boldsymbol{Z}_{\boldsymbol{I}};\hat{\boldsymbol{\theta}}_{\boldsymbol{N}}\right]}\right)+log\left(\frac{E_{\boldsymbol{Z_I}}\ Pr\left[\hat{c}_{g,t}^{O_g}=c\middle|\boldsymbol{Z}_{\boldsymbol{I}};\hat{\boldsymbol{\theta}}_{\boldsymbol{N}}\right]}{E_{\boldsymbol{Z_I}}\ Pr\left[\hat{c}_{g,t}^{O_g}=c\middle|\boldsymbol{Z}_{\boldsymbol{I}};\hat{\boldsymbol{\theta}}_{\boldsymbol{I}}\right]}\right). \tag{8}$$

The first term on the right-hand side of (8) is the gap that can be explained by differences in the observed endowments $\boldsymbol{Z}_{\boldsymbol{N}}$ and $\boldsymbol{Z}_{\boldsymbol{I}}$ evaluated by using the parameters as estimated on the native sample, $\hat{\boldsymbol{\theta}}_{\boldsymbol{N}}$. The last term in *Equation (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

[23] We may consider outcomes that are not conditioned on any particular time period *t* or grade *g*. In that 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 $Pr\left[\hat{c}_{g,g}^{O_g}=1\middle|.;.\right]$, 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 (8)* and in the benchmark empirical

literature to be negligible (Cameron and Heckman, 2001; Belzil and Poinas, 2010) or reduced substantially (Colding, 2006; Colding *at 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. However, if we consider the realization 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 realized if at earlier stages educational outcomes were successful: educational attainments realize 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 (9)*:[25]

analysis, we focus on the gap for youth with typical immigrant characteristics, i.e. at the lower end of the socioeconomic scale. We implement the alternative decomposition as a sensitivity analysis (cf. infra).

[25] Note that we do not condition the choice on *t*, meaning that we implicitly average over *t* (cf. footnote 23).

$$
\begin{aligned}
log\left(\frac{E_{\boldsymbol{Z_I}}\, Pr\left[\hat{c}_5^3 = 1 \middle| \boldsymbol{Z_I};\hat{\boldsymbol{\theta}}_{\boldsymbol{N}}\right]}{E_{\boldsymbol{Z_I}}\, Pr\left[\hat{c}_5^3 = 1 \middle| \boldsymbol{Z_I};\hat{\boldsymbol{\theta}}_{\boldsymbol{I}}\right]}\right) &= log\left(\frac{E_{\boldsymbol{Z_I}}\, Pr\left[\hat{c}_4^3 = 1 \middle| \boldsymbol{Z_I};\hat{\boldsymbol{\theta}}_{\boldsymbol{N}}\right]}{E_{\boldsymbol{Z_I}}\, Pr\left[\hat{c}_4^3 = 1 \middle| \boldsymbol{Z_I};\hat{\boldsymbol{\theta}}_{\boldsymbol{I}}\right]}\right) \\
&+ log\left(\frac{E_{\boldsymbol{Z_I}}\, Pr\left[\hat{c}_4^4 = 1 \middle| \boldsymbol{Z_I};\hat{\boldsymbol{\theta}}_{\boldsymbol{N}}\right] \Big/ E_{\boldsymbol{Z_I}}\, Pr\left[\hat{c}_4^3 = 1 \middle| \boldsymbol{Z_I};\hat{\boldsymbol{\theta}}_{\boldsymbol{N}}\right]}{E_{\boldsymbol{Z_I}}\, Pr\left[\hat{c}_4^4 = 1 \middle| \boldsymbol{Z_I};\hat{\boldsymbol{\theta}}_{\boldsymbol{I}}\right] \Big/ E_{\boldsymbol{Z_I}}\, Pr\left[\hat{c}_4^3 = 1 \middle| \boldsymbol{Z_I};\hat{\boldsymbol{\theta}}_{\boldsymbol{I}}\right]}\right). \qquad (9) \\
&+ log\left(\frac{E_{\boldsymbol{Z_I}}\, Pr\left[\hat{c}_5^3 = 1 \middle| \boldsymbol{Z_I};\hat{\boldsymbol{\theta}}_{\boldsymbol{N}}\right] \Big/ E_{\boldsymbol{Z_I}}\, Pr\left[\hat{c}_4^4 = 1 \middle| \boldsymbol{Z_I};\hat{\boldsymbol{\theta}}_{\boldsymbol{N}}\right]}{E_{\boldsymbol{Z_I}}\, Pr\left[\hat{c}_5^3 = 1 \middle| \boldsymbol{Z_I};\hat{\boldsymbol{\theta}}_{\boldsymbol{I}}\right] \Big/ E_{\boldsymbol{Z_I}}\, Pr\left[\hat{c}_4^4 = 1 \middle| \boldsymbol{Z_I};\hat{\boldsymbol{\theta}}_{\boldsymbol{I}}\right]}\right)
\end{aligned}
$$

Notice first that this gap cannot realize 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 (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 (9)*. Finally, the source of the gap can originate from not having passed fifth 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 (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 generalize this procedure for other outcomes and longer sequences of outcomes. This is what we do in the empirical application.

5. Results

We estimate the econometric model separately for native and immigrant youth. As mentioned in *Section 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 A–3* in *Appendix A* 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 paper 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 4.2.* In *Section 5.3* we specifically focus on the role that language plays in this composition. Finally, we briefly present some sensitivity analyses in *Section 5.4.*

5.1 Goodness of Fit

We simulate the fraction of both native and immigrant youth who realize a variety of schooling and labor market outcomes. The difference between the first two columns of *Table 2* (or the two first panels of *Figure 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 labor market outcome probabilities. A positive number means that native youth is more 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 3.2*. *Column B* describes the corresponding gaps based on the simulations that were described in *Section 4.2*. The main outcomes of interest, introduced as motivating gaps in *Section 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.

5.2 The Role of Family Endowments in Explaining the Gaps

Column C of *Table 2* presents the pure ethnic gap for a range of outcomes as obtained by conducting the counterfactual simulations outlined in *Section 4.2,* i.e. 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.

[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.

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 realized 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 probability from 0.20 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 the *Introduction*. 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* the pure ethnic gaps are further decomposed. First, consider the 7% pure ethnic gap of successfully completing secondary school. By composing this gap by analogy with the lines of *Equation (9)* in *Section 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 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.

We now consider the same two educational outcomes, but restrict these outcomes to be realized without schooling delay. As already mentioned in *Section 3.2,* this restriction substantially increases the (total) observed ethnic gaps for these outcomes. The log odds ratio for completing secondary education increases from 0.20 to 0.60, and for enrolling in tertiary education from 0.37 to 0.78. 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 focusing 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 standardized 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* 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.[27] 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 labor 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 of the field experiment conducted by Baert *et al.* (2013) for school-leavers in Flanders. This correspondence test indicates that in Flanders Turkish school graduates have to send out 44% more job applications in order to get the same amount of job interviews as their Flemish clones and points therefore to discrimination as a relevant explanation of the important pure ethnic gap in the labor market outcome that we find.

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 labor market outcome gaps is provided by, e.g., Hennessey *et al.* (2008), Filippin (2009) and Zaiceva and Zimmermann (2010). Second, a recent literature deals with the role of ethnic networks in explaining labor market outcome gaps (Winters, 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 neighborhoods (Colding, 2006; Colding *et al.*, 2009; Dustmann *et al.*, 2010). Third, in the absence of specific teaching incentive programs for disadvantaged or immigrant

[27] We only present labor market outcomes conditional on the educational level, because not all respondents completed education at the time the survey was conducted at age 23 (cf. *Section 3.2*).

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.

5.3 Gap Closing Role for Language?

Column E of *Table 2* presents evidence on language spoken at parental home as a source of schooling and first labor market gaps between native and immigrant youth. In the spirit of Cameron and Heckman (2001) these gaps are obtained by estimating the following ratio:

$$log\left(\frac{E_{\boldsymbol{Z}_{\boldsymbol{I}}^{*}}\ Pr\left[\hat{c}_{g,t}^{O_g}=c\middle|\boldsymbol{Z}_{\boldsymbol{I}}^{*};\hat{\boldsymbol{\theta}}_{\boldsymbol{I}}\right]}{E_{\boldsymbol{Z}_{\boldsymbol{I}}}\ Pr\left[\hat{c}_{g,t}^{O_g}=c\middle|\boldsymbol{Z}_{\boldsymbol{I}};\hat{\boldsymbol{\theta}}_{\boldsymbol{I}}\right]}\right), \tag{10}$$

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

From *Column E* of *Table 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 (3) of the 14 (11) points predicted (total) 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 (5) 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.

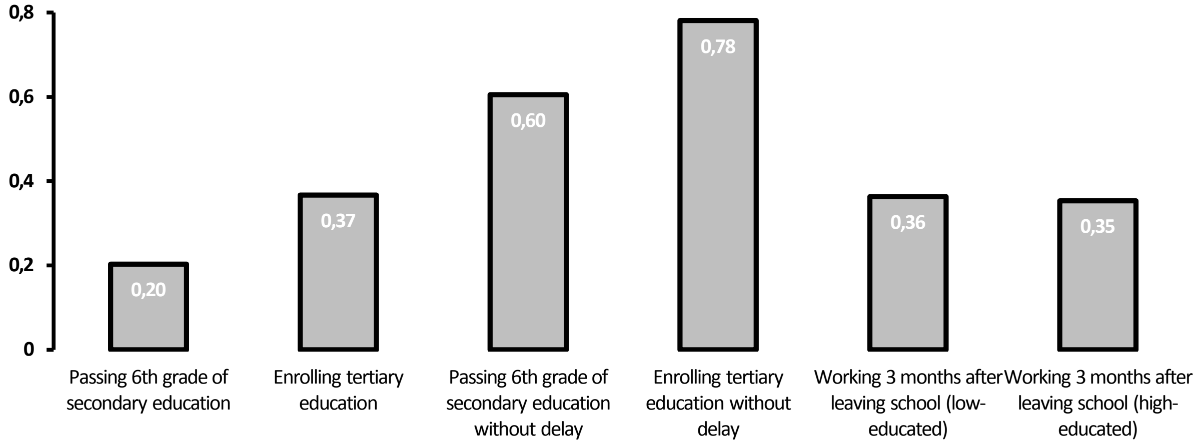
Table 2: Simulation results

	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]
Employed 3 m. 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]
Employed 3 m. after LS | LS with only SE6 degree	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]
Employed 3 m. after LS | LS with only SE6 degree 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]**
Employed 3 m. after LS | LS with TE1 - TE4 degree 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]**
Employed 3 m. after LS | LS with TE3 or TE4 degree 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]

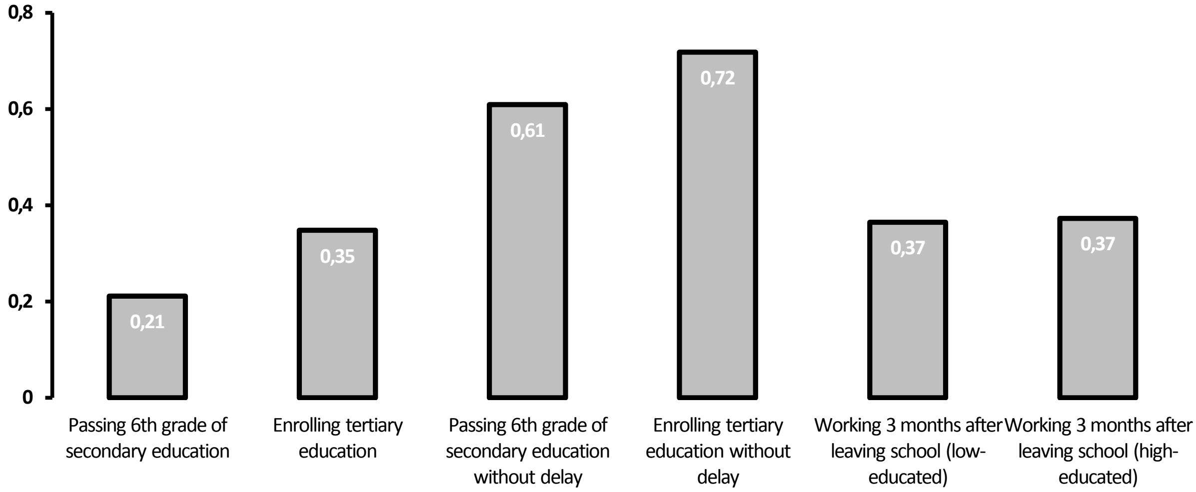
The described gaps are log odds ratios of the observed and simulated fractions of the native respectively the immigrant youth. Some abbreviations are used: PE (primary education), SE1 (1^{st} grade of secondary education), SE2 (2^{nd} grade of secondary education), SE3 (3^{rd} grade of secondary education), SE4 (4^{th} grade of secondary education), SE5 (5^{th} grade of secondary education), SE6 (6^{th} grade of secondary education), TE (tertiary education), TE1 (1^{th} grade of tertiary education), TE2 (2^{th} grade of tertiary education), TE3 (3^{th} grade of tertiary education), TE4 (4^{th} grade of tertiary education), D (delay), LS (leaving school) and m. (months). 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 3: Decomposition of observed schooling and labor market gaps

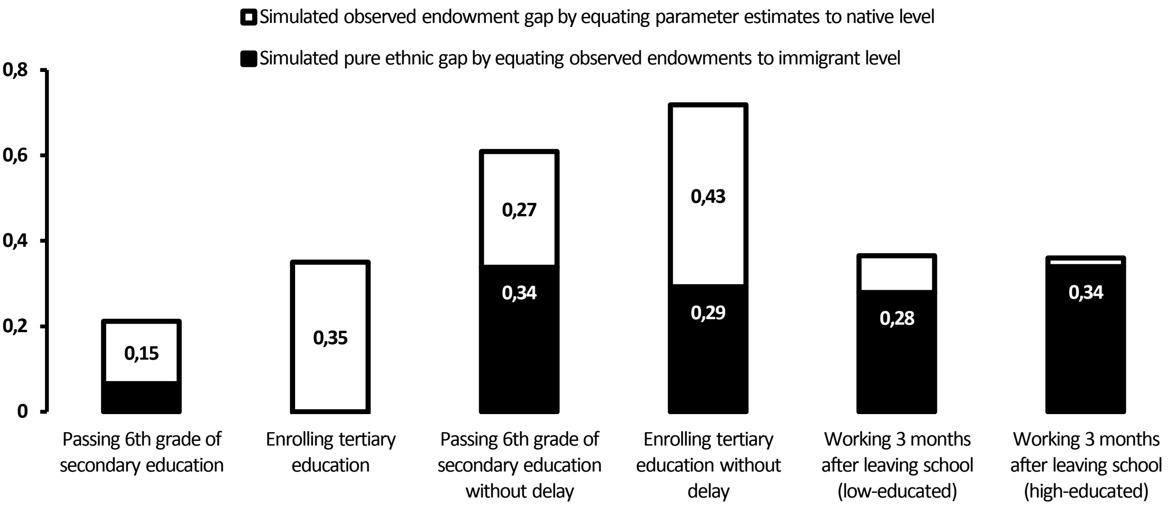
Panel A: observed total gaps (log odds ratios: native/immigrant)



Panel B: predicted total gaps (log odds ratios: native/immigrant)



Panel C: decomposed gaps (log odds ratios: native/immigrant)



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* 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.

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*, 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 addition, 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.

Second, we re-estimated our model replacing our indicator of labor market success by an alternative one, i.e. 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 A–4* in *Appendix A*. Even if the predictions for the alternative labor 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 labor 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 labor 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 labor 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. On the one hand, we restricted the immigrant population to the immigrant respondents with a Belgian nationality. In contrast with 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 get, for both the total and the pure ethnic gaps, simulation results that are very similar to the ones using our baseline definition. On the other hand, 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.

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 labor 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, born in Belgium or immigrated prior to age three.

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 labor 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 realize 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) realized 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 labor market indicator that measures permanent employment two years after graduation is used. In that case the native school-leaver is even 80%[30] 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 labor 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

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

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.

Appendix A: Additional tables

Table A–1: Unemployment rate (UR) in Flanders for males aged 15–64 (1982–2003)							
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 A–2: Unemployment rate (UR) in Flanders for the aged 15–24 (1982–2003)							
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 A–3: Model selection (information criteria values)

K	**Native 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 A-4: 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]
Permanent contract 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]
Permanent contract 2y. after LS | LS with only SE6 degree	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]
Permanent contract 2y. after LS | LS with only SE6 degree 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]**
Permanent contract 2y. after LS | LS with TE1 - TE4 degree and ≤ 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]**
Permanent contract 2y. after LS | LS with TE3 or TE4 degree and ≤ 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) and y. (years). 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.

Appendix B: 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, i.e. to $\boldsymbol{Z}_N$ respectively $\boldsymbol{Z}_I$;
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 realized;
4. Sum over all the possible paths to predict the R probabilities of realization 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.

References

ALDASHEV, A., J. GERNANDT, and S. L. THOMSEN (2009): "Language Usage, Participation, Employment and Earnings Evidence for Foreigners in West Germany with Multiple Sources of Selection," *Labour Economics*, 16, 330–341.

ANGRIST, J. D., and A. B. KRUEGER (1991): "Does Compulsory School Attendance Affect Schooling and Earnings," *Quarterly Journal of Economics*, 106, 979–1074.

ARULAMPALAM, W. (2001): "Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wages," *Economic Journal*, 111, 585–606.

BAERT, S., B. COCKX, N. GHEYLE, and C. VANDAMME (2013): "Do Employers Discriminate Less If Vacancies Are Difficult to Fill? Evidence from a Field Experiment," *IZA Discussion Papers*, 7145.

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

BELZIL, C., and F. POINAS (2010): "Education and Early Career Outcomes of Second-Generation Immigrants in France," *Labour Economics*, 17, 101–110.

BOUDON, R. (1974): "Education, Opportunity, and Social Inequality: Changing Prospects in Western Society," Wiley, New York.

CAMERON, S. V., and J. J. HECKMAN (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. V., and J. J. HECKMAN (2001): "The Dynamics of Educational Attainment for Black, Hispanic and White Males," *Journal of Political Economy*, 109, 455–499.

CARNEIRO, P., K. HANSEN, and J. J. HECKMAN (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.

CHISWICK, B. R. (2008): "The Economics of Language: An Introduction and Overview," *IZA Discussion Papers*, 3568.

COLDING, B. (2009): "A dynamic analysis of educational progression of children of immigrants," *Labour Economics*, 13, 479–492.

COLDING, B., L. HUSTED, and H. HUMMELGAARD (2009): "Educational progression of second-generation immigrants and immigrant children," *Economics of Education Review*, 28, 434–443.

CONSTANT, A. F., A. KRAUSE, U. RINNE, and K. F. ZIMMERMANN (2010): "Economic preferences and attitudes of the unemployed: Are natives and second generation migrants alike?" *International Journal of Manpower*, 32, 825–851.

DE RO, J. (2008): "Education in Flanders. A broad view on the Flemish education landscape," Agency for Educational Communication Publications, Brussels.

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

DUSTMANN, C., and F. FABBRI (2003): "Language Proficiency and Labour Market Performance of Immigrants in the UK," *Economic Journal*, 113, 695–717.

DUSTMANN, C., S. MACHIN, and U. SCHONBERG (2010): "Ethnicity and Educational Achievement in Compulsory Schooling," *Economic Journal*, 120, 272–297.

ECKSTEIN, Z., and K. I. WOLPIN (1998): "Estimating the Effect of Racial Discrimination on First Job Wage Offers," *Review of Economics and Statistics*, 81, 384–392.

ERIKSON, R., J. H. GOLDTHORPE, M. JACKSON, and D. R. COX (2010): "On class differentials in educational attainment," *Proceedings of the National Academy of Sciences of the United States of America*, 102, 9730–9733.

EUWALS, R., J. DAGEVOS, M. GIJSBERTS, and H. ROODENBURG (2010): "Citizenship and Labor Market Position: Turkish Immigrants in Germany and the Netherlands," *International Migration Review*, 44, 513–538.

FILIPPIN, A. (2009): "Can Workers's Expectations Account for the Persistence of Discrimination?" *IZA Discussion Papers*, 4490.

FRUEHWIRTH, J. C., S. NAVARRO, and Y. TAKAHASHI (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 Paper Series*, 7.

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

HAGY, A. P., and J. F. O. STANIEC (2002): "Immigrant Status, Race, and Institutional Choice in Higher Education," *Economics of Education Review*, 21, 381–392.

HECKMAN, J. J. (1981): "Statistical models for Discrete Panel Data," In: *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge.

HECKMAN, J. J. (2011): "The American Family in Black & White: A Post-Racial Strategy for Improving Skills to Promote Equality," *Daedalus*, 140, 70–89.

HECKMAN, J. J., and B. SINGER (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52, 271–320.

HECKMAN, J. J., and S. NAVARRO (2007): "Dynamic Discrete Choice and Dynamic Treatment Effects," *Journal of Econometrics*, 136, 341–396.

HECKMAN, J. J., S. URZUA, and E. VYTLACIL (2006): "Understanding Instrumental Variables in Models with Essential Heterogeneity", *Review of Economics and Statistics*, 88, 389–432.

HENNESSEY, M. L., P. D. RUMRILL, S. FITZGERALD, and R. ROESSLER (2008): "Disadvantagement-Related Correlates of Career Optimism among College and University Students with Disabilities," *Work – a Journal of Prevention Assessment & Rehabilitation*, 30, 483–492.

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

KEANE, M. P., and K. I. WOLPIN (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105, 473–522.

MACHADO, J. A. F., and J. MATA (2005): "Counterfactual decomposition of changes in wage distributions using quantile regression," *Journal of Applied Econometrics*, 20, 445–465.

MAHUTEAU, S., and P. N. JUANANKAR (2008): “Do Migrants get Good Jobs in Australia? The Role of Ethnic Networks in Job Search,” *Economic Record*, 84, 115–130.

MROZ, T. A., and T. H. SAVAGE (2006): “The Long-Term Effects of Youth Unemployment,” *Journal of Human Resources*, 41, 259–293.

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*, 20012008.

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.

ORTIZ, E. A., and C. DEHON (2008): “What Are the Factors of Success at University? A Case Study in Belgium,” *Cesifo Economic Studies*, 54, 121–148.

POZZOLI, D. (2009): “The Transition to Work for Italian University Graduates,” *Labour*, 23, 131–169.

RYAN, P. (2001): “The School-to-Work Transition: A Cross-National Perspective,” *Journal of Economic Literature*, 39, 767–767.

VAN DER KLAAUW, B., and A. V. VUUREN (2010): “Job search and academic achievement,” *European Economic Review*, 54, 298–320.

VAN OURS, J. C., and J. VEENMAN (2003): “The Educational Attainment of Second-Generation Immigrants in the Netherlands,” *Journal of Population Economics*, 16, 739–753.

WINTERS, P., A. DE JANVRY, and E. SADOULET (2001): “Family and Community Networks in Mexico-U.S. Migration,” *Journal of Human Resources*, 36, 159–184.

XIA, N., and N. KIRBY (2009): “Retaining students in grade. A literature review of the effects of retention on students’ academic and nonacademic outcomes,” *NBER Technical Reports*.

YAMAUCHI, F., and S. TANABE (2008): “Nonmarket networks among migrants: evidence from metropolitan Bangkok, Thailand,” *Journal of Population Economics*, 21, 649–664.

ZAICEVA, A., and K. F. ZIMMERMANN (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.