KU LEUVEN

DEPARTMENT OF ECONOMICS

Labour market consequences of a high school diploma

Deni MAZREKAJ, Kristof DE WITTE and Sarah VANSTEENKISTE

FACULTY OF ECONOMICS AND BUSINESS

Labour Market Consequences of a High School Diploma

By DENI MAZREKAJ, KRISTOF DE WITTE AND SARAH VANSTEENKISTE*

[4 July 2018]

This article compares the labour market outcomes of high school dropouts to high school graduates who did not enroll into higher education, but immediately entered the labour market. Using parental educational background as an instrument on a rich administrative dataset in the Flemish Region of Belgium, we find no returns to a high school diploma on average. However, these results hide considerable heterogeneity by gender and educational track. While females and individuals in vocational education may benefit from a diploma, male graduates and students holding a general education diploma may even be worse off on the labour market than dropouts. We show that sectoral heterogeneity acts as an underlying mechanism in the returns to a high school diploma (JEL I21, I28, J24).

Keywords: *Returns to Schooling; Heterogeneity; High School Diploma; School Dropout; Instrumental Variables.*

Word count: 7,558 words.

I. Introduction

Early literature has cast considerable doubt on the way education influences earnings. Attempting to correct for ability bias and measurement error, most articles estimate returns to schooling in an Instrumental Variables (IV) framework using compulsory schooling laws and educational background variables as instruments. Estimates range widely, from large positive returns (Harmon & Walker, 1995; Oreopoulos, 2006; Wang, 2013, Winters, 2015), through moderate positive returns (Angrist & Krueger, 1991; Devereux & Fan, 2011; Sakellariou & Fang, 2016), to no returns to schooling at all (Grenet, 2013; Meghir & Palme, 2005; Oosterbeek & Webbink, 2007; Pischke & von Wachter, 2008; Stephens & Yang, 2014). The conventional practice in the literature is to consider education as a continuous variable measuring the number of years of schooling. Although important, this measure may not be the most adequate way to estimate educational benefits on the labour market. First, the typical approach to construct a years of schooling variable is to assign the usual number of years taken to obtain a degree, disregarding curriculum interruptions and grade retention. This is a strong assumption given that in most OECD countries grade retention is considerable (Miyako & Garcia, 2014), e.g., France (36.9 percent), Belgium (34.9 percent), the Netherlands (26.7 percent), Switzerland (22.8 percent), Italy (16 percent), or the United States (14.2 percent).

Second, years of schooling are often calculated based on surveys or repeated cross-sections such as U.S. Population Censuses or European Labour Force Surveys. This data suffers from non-response and reporting bias. For example, respondents may over-report their attainment, or they may not know if their attainment counts as a qualification (e.g. GED-certificate). Battistin, De Nadai, & Sianesi (2014) estimate that 26.3 percent of dropouts over-report their attainment. This is problematic as measurement error in the years of schooling may overestimate the IV estimates of returns to schooling by as much as 38 percent (Bingley & Martinello, 2017).

Lastly, most educational systems consist of various educational tracks. For policy makers, students and parents, a single average rate of return may not be very informative in this occasion as, e.g., vocational students tend to drop out of high school without a degree at a much higher rate and may have a different rate of return than general education students (Dickson & Harmon, 2011; Balestra & Backes-Gellner, 2017). Given that individuals with an equal number of years of schooling may have very different educational outcomes, it is reasonable to assume that it is

not the years of education that play a significant role in the job-hiring process, but it is the (type of) qualification.

The objective of this article is to circumvent the problems associated with a continuous years of schooling variable, and present novel returns to schooling estimates as measured by a high school diploma that may act as a signal on the labour market (Clark & Martorell, 2014). For this, we employ a rich administrative dataset obtained from the Flemish government in Belgium, which includes all individuals who left secondary school with or without a diploma in school year 2010-2011. These individuals were enrolled in education in school year 2010- 2011, but were not enrolled in education in school year 2011-2012. To estimate the value of a high school diploma on the labour market, we compare labour market outcomes of high school dropouts to labour market outcomes of high school graduates who did not enroll into higher education, but immediately entered the labour market[1]. The focus on graduates without a higher education career is necessary, as these students are more likely to share observed and unobserved characteristics with high school dropouts.

We employ an IV framework using parental education as an instrument for the endogenous diploma attainment. This instrument has widely been used in the returns to schooling literature to control for ability bias (Heckman & Li, 2004; Trostel, Walker, & Woolley, 2002; Wang, 2013, Winters, 2015) on the premise that parental education is a strong predictor of diploma attainment. This is likely to be the case as the literature generally finds large positive intergenerational causal schooling effects (see the comprehensive overview by Holmlund, Lindahl, and Plug, 2011). Nevertheless, parental education should not be directly correlated with labour market outcomes. While the use of parental education to instrument for schooling is controversial on the grounds that it may directly influence earnings through networks, nepotism, or unobserved maternal ability, Hoogerheide, Block, and Thurik (2010) find that moderate violation of the assumption of perfect instrument validity does not affect results when parental education is used as an instrument.

Our contribution is threefold. First, we estimate the returns to a high school diploma and, therefore, circumvent the problems associated with the years of schooling variable listed above.

---

[1] This is around 88 percent of the total population of secondary education school leavers. The other 12 percent continue their studies in higher education.

Unlike commonly used survey data, our administrative dataset is highly reliable[2], and we observe the universe of all students leaving high school in school year 2010-2011. Moreover, our dataset allows us to precisely determine whether and when an individual obtained a diploma, and which education track the individual was enrolled in (general education or vocational education). Therefore, we avoid the upward bias in IV estimates associated with misreported years of schooling, as well as the distortion in the actual quantity of education caused by grade retention. In addition, by using quarterly labour market data, we use a more precise outcome variable than previous studies using annual labour market data. Quarterly labour market data is particularly useful when estimating returns to a high school diploma, as directly after school leaving the labour market prospects of dropouts may be bleak. Consequently, annual data may lead to regression to the mean. We observe quarterly labour market outcomes starting from quarter 3 in 2011 until quarter 4 in 2013, thus directly after graduation or dropout in quarter 2 in 2011.

Second, we estimate returns to schooling by gender and educational track as average estimates of the returns to schooling may hide considerable heterogeneity. Moreover, we provide insights in the mechanism that may explain these heterogeneous effects. In particular, we show that self-selection into employment sectors with varying returns to a high school diploma explain the different returns to schooling for different population subgroups.

Finally, this is the first time returns to schooling have been estimated in Belgium. The Belgian education system is particularly interesting thanks to the combination of a tracking system with a strong vocational orientation while, at the same time, students are initially free to enrol into a track and a school of their choice (there are no catchment areas and tracking is not based on standardized cognitive tests). Given that returns to schooling by country may range from 1.9 to 19.2 percent even when using comparable data (Trostel, Walker, & Woolley, 2002), it is interesting to explore where Belgium fares on the returns to schooling spectrum.

We do not find an effect of a high school diploma on quarterly earnings, on average. However, these results hide considerable heterogeneity by gender. Although a high school diploma is likely

---

[2] It is worth noting that our administrative dataset circumvents the problem of non-random selective samples often encountered in survey data when using parental background variables as instruments to estimate returns to schooling. Typically, parental background in surveys is available only for the individuals whose parents are present in the same household. This may lead to non-random selective samples, as living arrangement is an endogenous choice. Wang (2013) finds that studies estimating returns to schooling using survey data on parental background to instrument for schooling may be plagued by selection bias due to survey data, ranging from 8 to 10.5 percent. To make matters worse, the direction of the bias appears to be unclear.

to increase quarterly earnings of females by around 23 percent, a high school diploma may result in 12 percent lower earnings for males. We explain these results by observing heterogeneity by sector. Different sectors are likely to offer different returns to a high school diploma. In addition, sectors with no returns to a high school diploma also seem to be the sectors that attract most graduates. Similarly, males appear to self-select into sectors with negative returns, while females appear to self-select into sectors with large positive returns to a high school diploma. Furthermore, our results indicate that while a general education diploma appears to yield no returns on the labour market, individuals with a vocational education diploma incur higher quarterly earnings of around 11 percent.

The article proceeds as follows. In section II, we briefly explain the Flemish education system. Sections III and IV explain the methodology, sample construction and the data. In section V, we present the results and offer a potential mechanism explaining these results. Finally, section VI compares the results with the previous literature and provides some limitations of the study.
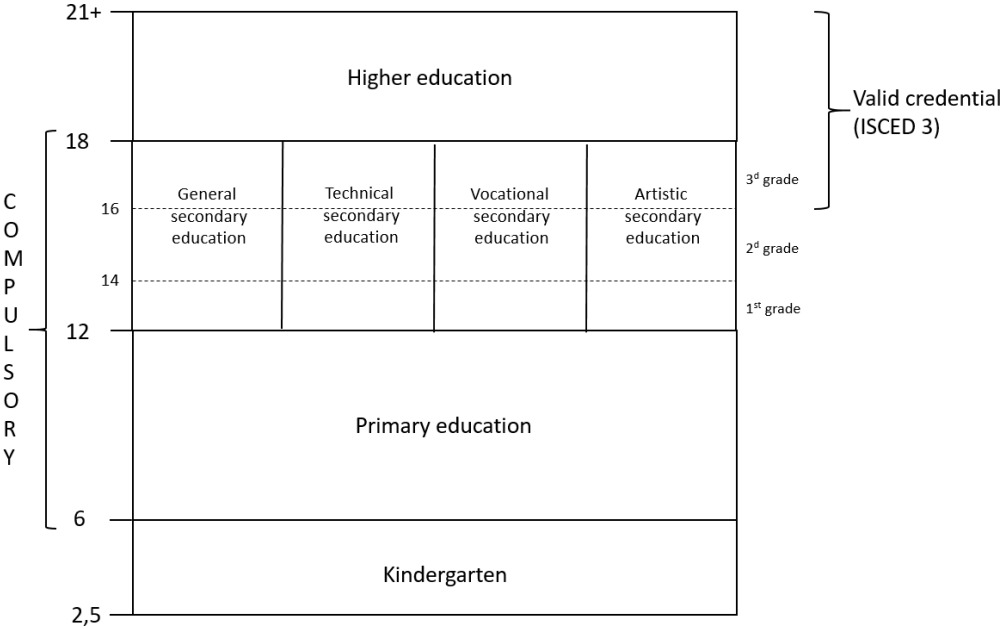
II.   Flemish Education System

FIGURE 1 – FLEMISH EDUCATION SYSTEM
(A valid credential is considered to be a diploma upper secondary education according to the International Standard Classification of Education framework (i.e. qualification ISCED level 3.)

Figure 1 displays the structure of the Flemish education system. Students start compulsory education at the age of six and finish at the age of eighteen. At the end of primary education, which lasts six years (age 12), students enter a tracking system based partly on the advice given by the students' primary school teacher, and partly by the parents' own aspirations (there are no standardized cognitive tests nor catchment areas). Secondary education lasts six years and includes four main tracks. General education aims to prepare students for higher education. Although technical education also intends to armour students with sufficient knowledge to enrol into higher education, the knowledge taught is rather technical and can be used to start a professional career immediately after graduation. By contrast, vocational education focuses on professional life and prepares students for a certain occupation. Lastly, artistic education includes a general education with a focus on arts.

Usually, around 46 percent of the students enrol into general education, around 30 percent into technical education, around 22 percent into vocational education, and around 2 percent into artistic education. The dropout rate is usually highest in vocational education (around 20 percent), followed by artistic education (around 14 percent), technical education (around 8 percent), and general education (3 percent). General education students are most likely to enrol into higher education upon graduation (about 55 percent), followed by technical education students (about 35 percent). It is very uncommon for vocational education and artistic education students to continue their studies in higher education (around 4.5 percent and 2.5 percent of these students enrol into higher education upon graduation, respectively).

III. Empirical Methodology

A. *Two Stage Least Squares (2SLS)*

The structural equation we estimate using Two Stage Least Squares (2SLS) is

$$(1) \quad Y_{it} = \beta_0 + \beta_1 D_{it} + \boldsymbol{\theta}_t + \boldsymbol{\alpha} \boldsymbol{X}_{it} + \epsilon_{it}$$

where Y_{it} is the labour market outcome of individual i in quarter t. We consider four outcomes: log gross quarterly earnings in euros, employment (1 is employed, 0 is unemployed or inactive), full time employment[3] (1 is full time, 0 is part time), and self-employment (1 is self-employed, 0 is not self-employed). Quarterly earnings are observed to fall within narrow intervals. Following Troster, Walker, and Woolley (2002), we therefore perform the analyses on the midpoints of the intervals[4]. The endogenous variable is high school diploma D_{it}, given value of 1 if individual i obtained a high school diploma and 0 otherwise, while $\boldsymbol{\theta}_t$ is a vector of quarter year fixed effects. Depending on the specification, we also include indicators for birth year[5], an indicator for gender (1 is male, 0 is female), and an indicator for origin (1 is foreign origin, 0 is Belgian origin)[6] as part of $\boldsymbol{X}_{it}$ in equation (1)[7].

The corresponding first stage equation is

$$(2) \quad D_{it} = \gamma_0 + \gamma_1 Z_{it} + \boldsymbol{\xi}_t + \boldsymbol{\delta} \boldsymbol{X}_{it} + \upsilon_{it}$$

---

[3] This outcome is estimated solely for the sample of employees.

[4] For instance, observing the quarterly earnings interval [1,000 EUR – 2,000 EUR], we perform the analyses on quarterly earnings of 1,500 EUR. We also used Stewart's (1983) maximum-likelihood interval regression, but little difference is found with the standard OLS estimates (results are available upon request).

[5] We have also estimated these models with a quadratic, and with a quartic in age (age x age2 x age3 x age4). The inclusion of a quadratic and a quartic in age respectively, instead of birth year fixed effects did not significantly influence our findings (the results are available upon request).

[6] This is calculated based on four characteristics: (1) current nationality, (2) nationality of birth, (3) father's nationality, and (4) mother's nationality. If the answer to one of these traits is 'foreign' and not 'Belgian', the individual is given value of one on the origin dummy. Replacing this variable with a dummy given value of 1 if the individual speaks a language other than Dutch with his parents and 0 otherwise did not significantly influence the results (the results are available upon request).

[7] The results (available upon request) are robust to the inclusion of school and municipality fixed effects.

where Z_{it} represents the parental educational background instrument. It is given value of 1 if an individual's mother[8] has obtained at least a high school diploma and 0 otherwise. ξ_t is a vector of quarter year fixed effects, while X_{it} estimated in some specifications includes controls (birth year indicators, gender, origin) as in equation (1).

Parental educational background serves as an instrument in our 2SLS specification (as used in other returns to schooling studies such as Heckman & Li, 2004; Trostel, Walker, & Woolley, 2002; Wang, 2013, Winters, 2015), hereby assuming that the instrument is strong. Weak instruments would not only lead to biased point estimates, but also incorrect confidence intervals (Staiger and Stock 1997). In the results presented below, we provide the first stage Kleibergen-Paap (2006) rk Wald F statistic in each table. All first-stage F statistics are well above the conventional weak instruments thresholds of both Stock and Yogo (2006) as Montiel Olea and Pflueger (2013)[9].

B. Restricting The Sample

We solely focus on students in fulltime secondary education. To avoid endogeneity issues arising from unobserved heterogeneity bias, students in part-time education, special education, post-secondary education and adult education have been removed from the analysis. Moreover, our data do not include students who obtained a high school diploma and subsequently enrolled into higher education. Given that the large majority (around 88 percent) of school leavers from secondary education in Flanders do not enroll into higher education, students with a high school diploma who do enroll into higher education arguably have different observed and unobserved characteristics, such that biased estimates would be obtained. Consequently, we solely compare labour market outcomes of high school dropouts with labour market outcomes of high school graduates who did not enroll into higher education, but who entered the labour market immediately upon graduation. Furthermore, individuals who have been inactive[10] during the entire period (from quarter 3 in 2011 until quarter 4 in 2013) have been removed from the

---

[8] Our dataset only includes mother's education. We have no information about other often used background variables, such as father's or spousal education.

[9] For completeness, we have also estimated all specifications using Limited Information Maximum Likelihood (LIML). The results are very similar to the results presented below (available upon request), giving further support that our instrument is likely to be strong.

[10] These individuals were primarily occupied with either housekeeping or volunteering.

analysis as these individuals were neither in employment nor unemployment. Finally, we consider individuals between the age of 18 and 24, consistent with the Eurostat (2017) definition of school dropout.

IV. Data

Our administrative dataset, originating from the Flemish government, consists of all individuals who have left education in school year 2010-2011. These individuals are either dropouts or graduates who were enrolled in education in school year 2010-2011, but were no longer enrolled in school year 2011-2012[11]. To distinguish between high school dropouts and high school graduates, we employ the definition constructed by Eurostat (2017). They define high school dropouts as individuals aged 18 to 24 who have not completed upper secondary education[12] and who were no longer in education and training. The labour market outcomes are observed each quarter after school leaving in 2011 until 2013. This leads to a balanced dataset with each individual's labour market outcomes observed ten times, from quarter three in 2011 (the individual was still enrolled into education in the first two quarters) until quarter four in 2013. Descriptive statistics are presented in **Table 1**.

---

[11] Individuals who later returned to education cannot be identified and are considered to have left education in school year 2010-2011.

[12] Upper secondary education is defined according to the International Standard Classification of Education framework (i.e. ISCED level 3).

TABLE 1 – DESCRIPTIVE STATISTICS

	Sample of school leavers in school year 2010-2011						
	Mean	Min	Max				
Endogenous variable: Diploma (1=yes, 0=no)	0.56	0	1				

	Graduates			Dropouts			Diff.
	Mean	Min	Max	Mean	Min	Max	
Instrument:							
Mother's education (1=at least upper secondary education, 0=otherwise)	0.69	0	1	0.50	0	1	**0.192** (0.003)
Control variables:							
Gender (1=male, 0=female)	0.66	0	1	0.63	0	1	**0.022** (0.003)
Origin (1=foreign, 0=Belgian)	0.16	0	1	0.40	0	1	**-0.239** (0.003)
Age	19.16	18	24	18.80	18	24	**0.356** (0.007)
Outcomes (quarterly):							
Gross earnings (EUR)	6,734	0	25,000	3,781	0	25,000	**2,952** (300.48)
Log gross earnings	4.54	2.71	5.52	4.46	2.71	5.52	**0.087** (0.002)
Employed (1=employed, 0=unemployed or inactive)	0.73	0	1	0.44	0	1	**0.284** (0.003)
Full time employed (1=full time, 0=part time)	0.67	0	1	0.43	0	1	**0.233** (0.004)
Self employed (1=self employed, 0=not self employed)	0.04	0	1	0.02	0	1	**0.025** (0.001)
Observations		58,070			45,020		103,090

Notes. Significance at the 5 percent level is denoted by bold coefficients. Robust standard errors are in parentheses.

The final dataset includes 103,090 observations (10,309 individuals observed in ten consecutive quarters after school leaving); 5,807 graduates and 4,502 dropouts. A little less than half of the students (44 percent) in our sample left education without a high school diploma. One should remember that we are comparing high school dropouts to high school graduates who did not enroll into higher education, but instead immediately entered the labour market. Given that we are not including graduates who enrolled into higher education, this high percentage of high school dropouts is not surprising. About 66 percent of our sample are males and 16 percent of the sample are individuals with foreign origin. Furthermore, most individuals left education at 18 or 19 years old. Using a simple t-test, **Table 1** suggests that graduates have a significantly

higher parental background than dropouts do. Graduates seem to be about 19 percentage points more likely to have a mother who obtained at least a high school diploma. At first sight, graduates are more likely to be both employed and self-employed. Moreover, once employed, graduates are more likely to have a full time job and to earn more. Evidently, these results are merely descriptive given that abler students are more likely to obtain a diploma and perform better on the labour market regardless. In the next section, we attempt to correct for this ability bias by instrumenting diploma with parental educational background.

V. Results

A. The Effect of a High School Diploma on Log Quarterly Earnings

Table 2 shows the endogenous OLS and causal 2SLS effect of a high school diploma on log quarterly earnings. To rule out labour force participation effects, the models were also estimated by gender[13] (Oosterbeek & Webbink, 2007).

TABLE 2 – THE EFFECT OF A HIGH SCHOOL DIPLOMA ON LOG GROSS QUARTERLY EARNINGS
FOR 18 TO 24 YEAR OLDS FROM QUARTER 3 IN 2011 TO QUARTER 4 IN 2013

Sample of school leavers in school year 2010-2011	Full sample		Males		Females	
Endogenous variable: Diploma (1 is high school diploma)	(mean log earnings dropouts = 4.46) [3,781 EUR]		(mean log earnings dropouts = 4.51) [4,015 EUR]		(mean log earnings dropouts = 4.36) [3,375 EUR]	
	(1)	(2)	(3)	(4)	(5)	(6)
OLS	**0.088**	**0.071**	**0.093**	**0.085**	**0.052**	**0.037**
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
2SLS	**0.126**	0.009	-0.011	**-0.091**	**0.222**	**0.230**
	(0.012)	(0.016)	(0.015)	(0.022)	(0.025)	(0.035)
First stage:	[mean diploma = 0.56]		[mean diploma = 0.57]		[mean diploma = 0.55]	
Mother's education[a]	**0.173**	**0.125**	**0.172**	**0.126**	**0.130**	**0.095**
	(0.004)	(0.004)	(0.005)	(0.005)	(0.007)	(0.007)
F (first stage instruments)	1853.89	986.58	1188.69	648.11	332.94	179.60
Fixed effects:						
Birth year	No	Yes	No	Yes	No	Yes
Year x quarter	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls[b]	No	Yes	No	Yes	No	Yes
Observations	59,461	59,461	39,583	39,583	19,878	19,878

Notes. Significance at the 5 percent level is denoted by bold coefficients. Robust standard errors are in parentheses.
[a] Instrument mother's education is coded as 1 is individual's mother has obtained at least a high school diploma, 0 is otherwise.
[b] Control variables include gender (1 is male, 0 is female) and origin (1 is foreign origin, 0 is Belgian origin).

The endogenous OLS results in **Table 2** indicate that the returns to a high school diploma are positive in both the full sample as for males and females separately, regardless of the inclusion of controls in the model. However, as these results might be plagued by ability bias, we estimate

---

[13] We have also estimated the models by origin. However, the conclusions from the full sample generally appear to hold for both the individuals of Belgian origin as for the individuals of foreign origin. One exception is that a diploma does not appear to affect employment of Belgian individuals. These results are presented in **Table A1** and **Table A2** in the Appendix.

all specifications with 2SLS using mother's education as an instrument. This instrument appears to be strong, given the high F-statistics and given that individuals' probability of obtaining a high school diploma increases by 9.5 to 17 percentage points (from a mean of around 56 percent) if their mother has obtained at least a high school diploma. Columns (1) and (2) indicate that, once we control for birth year, gender and origin, there is no significant effect of a high school diploma on quarterly earnings14. However, these results hide considerable heterogeneity by gender. Column (6) shows that a high school diploma is likely to increase quarterly earnings of females by 23 percent (from an average of 3,375 euros). Males, on the other hand, appear to be harmed by a high school diploma. More specifically, column (4) suggests that a high school diploma reduces males' quarterly earnings by 9 percent (from an average of 4,015 euros). These negative returns to a high school diploma may be explained by compensating wage differentials. Less educated workers in jobs with less-desirable traits may be compensated with higher wages compared to more educated workers in enjoyable jobs (Oreopoulos & Salvanes, 2011). We elaborate on this in Section D (Mechanisms).

B. *The Effect of a High School Diploma on Employment*

TABLE 3 – THE EFFECT OF A HIGH SCHOOL DIPLOMA ON EMPLOYMENT FOR 18 TO 24 YEAR OLDS FROM QUARTER 3 IN 2011 TO QUARTER 4 IN 2013

| Endogenous variable: High school diploma (1=yes) | Sample of school leavers in school year 2010-2011 | | | | Observations |
	OLS (1)	2SLS[a] (2)	OLS (3)	2SLS[a] (4)	
Outcomes for the full sample:					
1. Employed	**0.284**	**0.276**	**0.246**	**0.079**	103,090
(mean dropouts = 0.44)	(0.003)	(0.015)	(0.003)	(0.022)	
2. Full time employed[b]	**0.235**	**0.485**	**0.201**	**0.309**	59,775
(mean dropouts = 0.43)	(0.004)	(0.024)	(0.004)	(0.032)	
3. Self employed	**0.025**	**0.048**	**0.028**	**0.064**	103,090
(mean dropouts = 0.02)	(0.001)	(0.006)	(0.001)	(0.008)	
Outcomes for males:					
4. Employed	**0.306**	**0.221**	**0.276**	0.010	66,650
(mean dropouts = 0.45)	(0.004)	(0.019)	(0.004)	(0.032)	
5. Full time employed[b]	**0.245**	**0.354**	**0.223**	**0.262**	39,744
(mean dropouts = 0.52)	(0.005)	(0.027)	(0.005)	(0.038)	
6. Self employed	**0.023**	**0.028**	**0.028**	**0.047**	66,650
(mean dropouts = 0.02)	(0.001)	(0.008)	(0.001)	(0.012)	
Outcomes for females:					
7. Employed	**0.241**	**0.311**	**0.187**	**0.169**	36,440
(mean dropouts = 0.43)	(0.005)	(0.024)	(0.005)	(0.032)	
8. Full time employed[b]	**0.181**	**0.398**	**0.153**	**0.378**	20,031
(mean dropouts = 0.28)	(0.007)	(0.041)	(0.007)	(0.059)	
9. Self employed	**0.028**	**0.076**	**0.028**	**0.088**	36,440
(mean dropouts = 0.01)	(0.002)	(0.008)	(0.002)	(0.011)	
Fixed effects:					
Birth year	No	No	Yes	Yes	
Year x quarter	Yes	Yes	Yes	Yes	
Additional controls[c]	No	No	Yes	Yes	

Notes. Significance at the 5 percent level is denoted by bold coefficients. Robust standard errors are in parentheses.
[a] 2SLS estimates are obtained with the instrument 'mother's education' coded as 1 is individual's mother has obtained at least a high school diploma, and 0 is otherwise. The first stage in each model is significantly different from zero and the Kleibergen- Paap (2006) rk Wald F statistic is far above the conventional thresholds of Stock and Yogo (2006), and Montiel Olea and Pflueger (2013).
[b] Full time employment (rows 2, 5, and 8) is estimated for the sample of employees.
[c] Control variables include gender (1 is male, 0 is female) and origin (1 is foreign origin, 0 is Belgian origin).

   Table 3 presents the effect of a high school diploma on employment outcomes. The results from the preferred specification in column (4) show that, on average, a high school diploma appears to increase the probability of employment by 7.9 percent, full time employment once employed by 30.9 percent, and self-employment by 6.4 percent. Although these positive results seem to hold for females, the effect of a high school diploma on employment for males, shown

in row 4 column (4), does not appear to significantly differ from zero. Furthermore, there appear to be differences between males and females for full time employment (26.2 percent versus 37.8 percent, respectively) and self-employment (4.7 percent versus 8.8 percent).

C. Heterogeneity by Educational Track

TABLE 4 – THE EFFECT OF A HIGH SCHOOL DIPLOMA ON LABOUR MARKET OUTCOMES FOR 18 TO 24 YEAR OLDS FROM QUARTER 3 IN 2011 TO QUARTER 4 IN 2013 BY EDUCATIONAL TRACK

	Sample of school leavers in school year 2010-2011			
	General track (ASO) (mean diploma = 0.53)		Vocational track (BSO) (mean diploma = 0.39)	
Endogenous variable:	OLS	2SLS[a]	OLS	2SLS[a]
High school diploma (1=yes)	(1)	(2)	(3)	(4)
Outcomes:				
	(mean dropouts = 4.45)		(mean dropouts = 4.46)	
	[2,932 EUR]		[3,967 EUR]	
	observations = 3,006		observations = 18,363	
1. Log quarterly earnings	0.010	-0.024	**0.050**	**0.113**
	(0.010)	(0.050)	(0.003)	(0.047)
	(mean dropouts = 0.35)		(mean dropouts = 0.46)	
	observations = 7,100		observations = 34,520	
2. Employed	**0.119**	**-0.370**	**0.222**	**0.167**
	(0.013)	(0.116)	(0.005)	(0.071)
	(mean dropouts = 0.38)		(mean dropouts = 0.42)	
	observations = 3,008		observations = 18,517	
3. Full time employed	**0.219**	0.192	**0.146**	0.082
	(0.020)	(0.109)	(0.007)	(0.095)
	(mean dropouts = 0.02)		(mean dropouts = 0.02)	
	observations = 7,100		observations = 34,520	
4. Self employed	0.001	**-0.145**	**0.028**	**0.157**
	(0.003)	(0.041)	(0.001)	(0.026)
Fixed effects:				
Birth year	Yes	Yes	Yes	Yes
Year x quarter	Yes	Yes	Yes	Yes
Additional controls[c]	Yes	Yes	Yes	Yes

Notes. Significance at the 5 percent level is denoted by bold coefficients. Robust standard errors are in parentheses.
[a] 2SLS estimates are obtained with the instrument 'mother's education' coded as 1 is individual's mother has obtained at least a high school diploma, and 0 is otherwise. The first stage in each model is significantly different from zero and the Kleibergen- Paap (2006) rk Wald F statistic is far above the conventional thresholds of Stock and Yogo (2006), and Montiel Olea and Pflueger (2013).
[b] Full time employment (rows 2, 5, and 8) is estimated for the sample of employees.
[c] Control variables include gender (1 is male, 0 is female) and origin (1 is foreign origin, 0 is Belgian origin).

Table 4 presents the effect of a high school diploma on four labour market outcomes by educational track[14]. The results in column (2) indicate that there are no returns to a diploma for individuals with a diploma from general education who do not enrol in higher education. Thus, it appears that if general education students do not continue their studies, they do not benefit from their high school diploma in terms of earnings on the labour market. In other words, these students seem to be penalized on the labour market. The likely reason for this is that students from general education are expected to enrol into higher education. Consequently, not enrolling into higher education may serve as a negative signal to potential employers. Moreover, general education graduates even seem to be worse off on the labour market than high school dropouts in terms of employment. Nonetheless, once employed, these individuals appear to have a higher chance of being full time employed (marginally significant at the ten percent level). Furthermore, high school graduates from the general track have a 14 percent lower probability to engage into self-employment than high school dropouts. On the other hand, results in column (4) suggest that individuals with a vocational education diploma incur positive returns of around 11 percent. Moreover, they are more likely to be both employed as self-employed. However, it appears that vocational education graduates do not benefit from their diploma in terms of having full time employment.

---

[14] We present the results for only two educational tracks instead of four. This is in part because artistic education includes too little observations to have a meaningful power of the analysis, and in part because our instrument in the sample of only technical education students produced a Kleibergen-Paap (2006) Wald rk F statistic which did not exceed the Montiel Olea and Pflueger (2013) threshold at 5 percent level.

D. Mechanism

TABLE 5 – THE EFFECT OF A HIGH SCHOOL DIPLOMA ON LABOUR MARKET OUTCOMES FOR 18
TO 24 YEAR OLDS FROM QUARTER 3 IN 2011 TO QUARTER 4 IN 2013 BY SECTOR, AND
SELECTION INTO SECTORS BY DIPLOMA AND BY GENDER

Endogenous variable: High school diploma (1=yes)	Sample of school leavers in school year 2010-2011		
	Returns to a diploma (1)	% of dropouts who chose the sector (2)	% of graduates who chose the sector (3)
Sectors[a]:			
1. Manufacturing (N=8,176)	-0.039	9.37%	17.41%
(mean earnings=11,476 EUR)	(0.059)		
(mean earnings dropouts=10,780 EUR)			
2. Construction (N=6,792)	-0.009	7.79%	14.52%
(mean earnings=10,569 EUR)	(0.041)		
(mean earnings dropouts=10,194 EUR)			
3. Transport (N=1,959)	**0.132**	3.25%	3.71%
(mean earnings=9,405 EUR)	(0.065)		
(mean earnings dropouts=8,875 EUR)			
4. Health (N=4,017)	**0.218**	8.76%	4.30%
(mean earnings=9,381 EUR)	(0.040)		
(mean earnings dropouts=8,719 EUR)			
5. Support service (N=14,876)	**-0.058**	<u>37.72%</u>	<u>21.90%</u>
(mean earnings=9,176 EUR)	(0.027)		
(mean earnings dropouts=8,991 EUR)			
6. Government (N=3,416)	**0.208**	7.08%	4.42%
(mean earnings=9,023 EUR)	(0.033)		
(mean earnings dropouts=8,472 EUR)			
7. Catering (N=4,253)	0.015	11.66%	5.82%
(mean earnings=8,788 EUR)	(0.043)		
(mean earnings dropouts=8,669 EUR)			
8. Trade (N=11,592)	**0.174**	21.49%	20.79%
(mean earnings=8,161 EUR)	(0.049)		
(mean earnings dropouts=7,664 EUR)			
Method[b]	2SLS		
Fixed effects:			
Birth year	Yes		
Year x quarter	Yes		
Additional controls[c]	Yes		

Notes. Significance at the 5 percent level is denoted by bold coefficients. Highest frequencies are underlined in columns (2) and (3). Robust standard errors are reported in parentheses in model (1).

[a] Sectors are defined as in the Nomenclature statistique des Activités économiques dans la Communauté Européenne (NACE).

[b] 2SLS estimates are obtained with the instrument 'mother's education' coded as 1 is individual's mother has obtained at least a high school diploma, and 0 is otherwise. The first stage in each model is significantly different from zero and the Kleibergen- Paap (2006) rk Wald F statistic is far above the conventional thresholds of Stock and Yogo (2006), and Montiel Olea and Pflueger (2013).

[c] Control variables include gender (1 is male, 0 is female) and origin (1 is foreign origin, 0 is Belgian origin).

It is possible to explain the absence of returns to a high school diploma by estimating, on the one hand, the sectoral heterogeneity in returns to a high school diploma, and on the other hand, the way different subgroups select into sectors with different returns to a diploma. These results are estimated in **Table 5** for the eight biggest sectors in our dataset as defined in the 'Nomenclature statistique des Activités économiques dans la Communauté Européenne' (NACE). We observe in column (1) that returns to a high school diploma are absent in manufacturing, construction and catering, while they are negative in support services (these are occupations such as employment agencies, travel agencies and security). On the other hand, returns to a high school diploma are positive in trade, transport, government and health sectors. Moreover, frequencies in columns (2) and (3) show that dropouts (37.72 percent) and graduates (21.90 percent) mainly choose to work in the support services sector which yields negative returns to a high school diploma. This is also the case for males as can be seen from the frequencies displayed in **Figure A2** in the Appendix. Females, on the other hand, mostly choose to work in the trade sector which yields high returns to a high school diploma. Thus, it appears that graduates and males self-select into sectors with low or negative returns to a high school diploma. This provides a potential explanation for the results in the previous section in which we estimate no returns to a high school diploma on average, negative returns to a high school diploma for males, and positive returns to a high school diploma for females.

Table 5 also reveals another important point. It seems that, in general, returns to a high school diploma are higher in sectors with lower mean earnings. To further analyse this, we use Instrumental Variables Quantile Regression (IVQR) proposed by Frölich and Melly (2013). This enables us to assess the impact of a high school diploma on each quantile of earnings, conditional on other covariates. This additional analysis confirms our findings. We observe higher returns to a high school diploma at the lower quantiles of the conditional earnings distribution (also observed in Balestra and Backes-Gellner, 2017). The results are presented in **Table A3** in the Appendix.

VI. Conclusion

This article estimated the returns to a high school diploma using administrative data of school leavers in school year 2010-2011 in the Flemish region of Belgium. To our knowledge, this is the first time returns to schooling have been calculated in Belgium. To account for selection bias, we employed the IV framework using mother's educational background as an instrument. Our results suggest no returns to a high school diploma on average, with high returns for females of around 23 percent and negative effects for males of around 12 percent. We find that the self-selection into employment sectors acts as a mechanism explaining the results. Different sectors are likely to offer different returns to a high school diploma. In addition, sectors with no returns to a high school diploma also seem to be the sectors that attract most graduates. Similarly, males appear to self-select into sectors with negative returns, while females appear to self-select into sectors with large positive returns to a high school diploma. Furthermore, we also show that although individuals with a vocational diploma are likely to benefit on the labour market, individuals with a diploma general education who do not enrol in higher education may experience no consequences in terms of earnings.

These results are in sharp contrast with the traditional survey data estimates of the returns to a high school diploma of around 26 percent in the United Kingdom (Battistin, De Nadai, & Sianesi, 2014), between 22 and 25 percent in Norway (Aakvik, Salvanes, & Vaage, 2010), between 28 and 66 percent in Austria (Fersterer & Winter-Ebmer, 2003), and around 42 percent in Italy (Brunello & Miniaci, 1999). However, our results correspond to the only other study on returns to a high school diploma that uses administrative data. In particular, Clark and Martorell (2014) compare students who barely failed and barely passed a high school exit exam in a Regression Discontinuity Design (RDD). They find no returns to a high school diploma in Texas.

It should be noted, however, that there are several limitations to our study. First, our results are averaged over the population of compliers. These are students who obtained a diploma due to their mother having at least a high school diploma and who would not have done so if their mother were a high school dropout. Although the Local Average Treatment Effect (LATE) estimates should converge towards the Average Treatment Effect (ATE) estimates as the number of compliers becomes an increasingly large proportion of the sample (Oreopoulos, 2006), these results may not apply for students who would or would have not obtained a diploma regardless

of their mother's education. Second, we only consider private returns to schooling defined as the increase in individual earnings resulting from a diploma. We make no statement about the social returns to education defined as the increase in total earnings resulting from an increase in average diploma attainment in the entire Flemish region. These social returns to schooling may be lower (signalling theory) or even higher (human capital theory) than the private ones (Acemoglu & Angrist, 2000). Lastly, we only consider the effect of education on pecuniary outcomes. However, education can also have an effect on nonpecuniary outcomes such as health, crime, marriage rates, political participation and others (Oreopoulos & Salvanes, 2011). These outcomes are just as (if not more) important than earnings. Future research should address these issues.
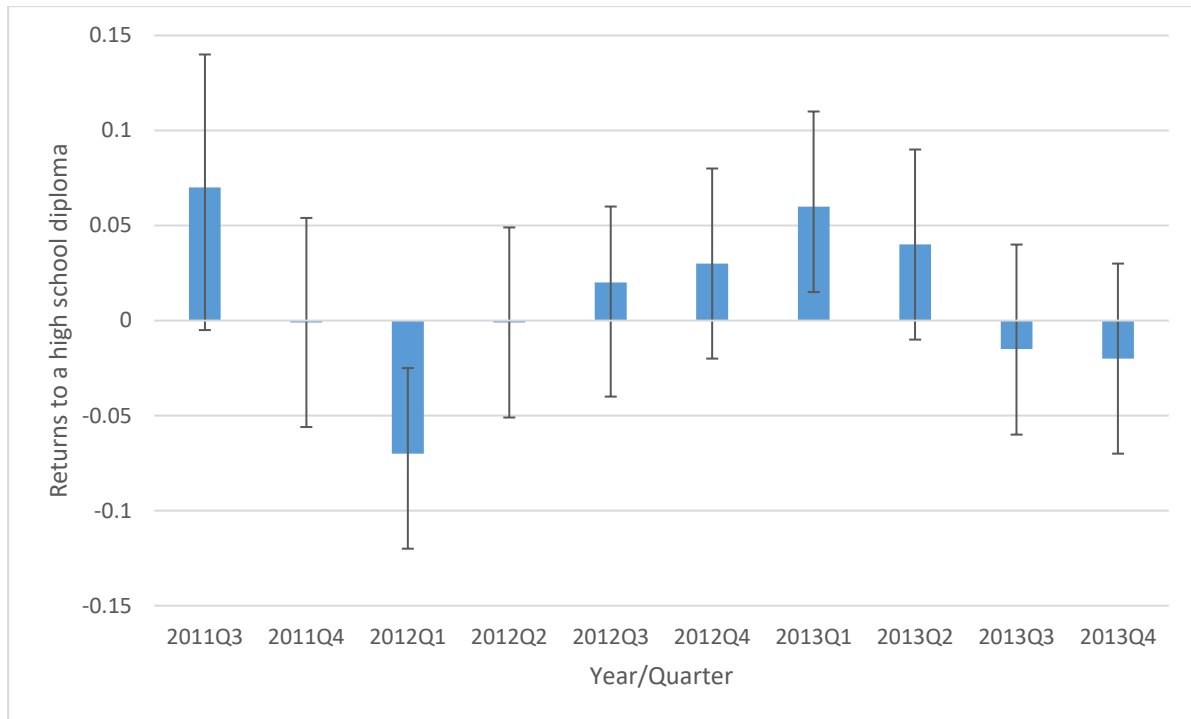
FIGURE A1 – THE EFFECT OF A HIGH SCHOOL DIPLOMA ON LOG GROSS QUARTERLY EARNINGS FOR 18 TO 24 YEAR OLDS FROM QUARTER 3 IN 2011 TO QUARTER 4 IN 2013 BY QUARTER (None of the coefficients is significantly different from zero at the five percent significance level (based on robust standard errors). Models are estimated by 2SLS with the instrument 'mother's education' coded as 1 is individual's mother has obtained at least a high school diploma, and 0 is otherwise. The first stage in each model is significantly different from zero and the Kleibergen-Paap (2006) rk Wald F statistic is far above the conventional thresholds of Stock and Yogo (2006), and Montiel Olea and Pflueger (2013). Control variables include gender (1 is male, 0 is female), origin (1 is foreign origin, 0 is Belgian origin), and birth year indicators.)

Sample of school leavers in school year 2010-2011 Endogenous variable: Diploma (1 is high school diploma)	Full sample (mean log earnings dropouts = 4.46) [3,781 EUR]		Belgian origin (mean log earnings dropouts = 4.46) [4,196 EUR]		Foreign origin (mean log earnings dropouts = 4.45) [3,150 EUR]	
	(1)	(2)	(3)	(4)	(5)	(6)
OLS	**0.088**	**0.071**	**0.088**	**0.072**	**0.066**	**0.066**
	(0.002)	(0.002)	(0.003)	(0.002)	(0.005)	(0.004)
2SLS	**0.126**	0.009	**0.139**	0.010	0.029	-0.007
	(0.012)	(0.016)	(0.019)	(0.020)	(0.028)	(0.027)
First stage:	[mean diploma = 0.56]		[mean diploma = 0.64]		[mean diploma = 0.34]	
Mother's education[a]	**0.173**	**0.125**	**0.126**	**0.113**	**0.169**	**0.168**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.009)	(0.009)
F (first stage instruments)	1853.89	986.58	782.74	647.90	341.80	353.63
Fixed effects:						
Birth year	No	Yes	No	Yes	No	Yes
Year x quarter	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls[b]	No	Yes	No	Yes	No	Yes
Observations	59,461	59,461	47,645	47,645	11,816	11,816

Notes. Significance at the 5 percent level is denoted by bold coefficients. Robust standard errors are in parentheses.
[a] Instrument mother's education is coded as 1 is individual's mother has obtained at least a high school diploma, 0 is otherwise.
[b] Control variables include gender (1 is male, 0 is female) and origin (1 is foreign origin, 0 is Belgian origin).

	Sample of school leavers in school year 2010-2011				
Endogenous variable:	OLS	2SLS[a]	OLS	2SLS[a]	Observations
High school diploma (1=yes)	(1)	(2)	(3)	(4)	
Outcomes for Belgian origin students:					
1. Employed	**0.262**	**0.099**	**0.248**	0.030	76,090
(mean dropouts = 0.48)	(0.004)	(0.025)	(0.004)	(0.028)	
2. Full time employed[b]	**0.202**	**0.467**	**0.186**	**0.317**	47,868
(mean dropouts = 0.48)	(0.005)	(0.038)	(0.005)	(0.040)	
3. Self employed	**0.030**	**0.071**	**0.034**	**0.081**	76,090
(mean dropouts = 0.01)	(0.001)	(0.009)	(0.001)	(0.010)	
Outcomes for foreign origin students:					
4. Employed	**0.245**	**0.188**	**0.240**	**0.174**	27,000
(mean dropouts = 0.38)	(0.006)	(0.040)	(0.006)	(0.039)	
5. Full time employed[b]	**0.241**	**0.335**	**0.247**	**0.283**	11,907
(mean dropouts = 0.34)	(0.009)	(0.053)	(0.009)	(0.051)	
6. Self employed	**0.011**	**0.038**	**0.013**	**0.033**	27,000
(mean dropouts = 0.02)	(0.002)	(0.014)	(0.002)	(0.014)	
Fixed effects:					
Birth year	No	No	Yes	Yes	
Year x quarter	Yes	Yes	Yes	Yes	
Additional controls[c]	No	No	Yes	Yes	

Notes. Significance at the 5 percent level is denoted by bold coefficients. Robust standard errors are in parentheses.
[a] 2SLS estimates are obtained with the instrument 'mother's education' coded as 1 is individual's mother has obtained at least a high school diploma, and 0 is otherwise. The first stage in each model is significantly different from zero and the Kleibergen- Paap (2006) rk Wald F statistic is far above the conventional thresholds of Stock and Yogo (2006), and Montiel Olea and Pflueger (2013).
[b] Full time employment (rows 2, 5, and 8) is estimated for the sample of employees.
[c] Control variables include gender (1 is male, 0 is female) and origin (1 is foreign origin, 0 is Belgian origin).
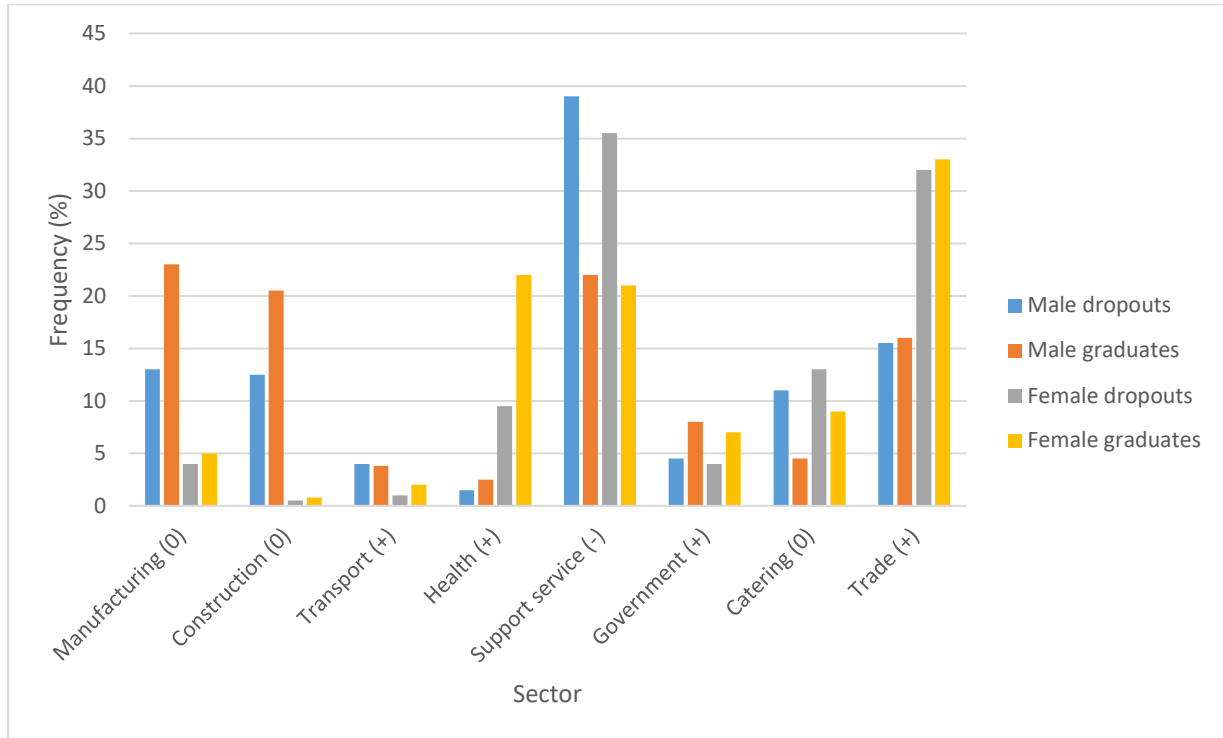
FIGURE A2 – PERCENTAGE OF DROPOUTS AND GRADUATES WHO CHOSE A PARTICULAR SECTOR BY GENDER

(Returns to a high school diploma are presented between brackets (0 is no returns, + is positive returns, and – is negative returns). Sectors are defined as in the Nomenclature statistique des Activités économiques dans la Communauté Européenne (NACE).)

TABLE A3 – THE EFFECT OF A HIGH SCHOOL DIPLOMA ON LOG GROSS QUARTERLY EARNINGS FOR 18 TO 24 YEAR OLDS FROM QUARTER 3 IN 2011 TO QUARTER 4 IN 2013 ACROSS THE CONDITIONAL EARNINGS DISTRIBUTION

Endogenous variable: High school diploma (1=yes)	Sample of school leavers in school year 2010-2011				
	20th percentile (1)	40th percentile (2)	50th percentile (3)	60th percentile (4)	80th percentile (5)
Returns to a high school diploma	**0.143** (0.012)	**0.125** (0.012)	0.000 (0.014)	0.000 (0.014)	0.000 (0.015)
Fixed effects:					
Birth year	Yes	Yes	Yes	Yes	Yes
Year x quarter	Yes	Yes	Yes	Yes	Yes
Additional controls[a]	Yes	Yes	Yes	Yes	Yes

Notes. Significance at the 5 percent level is denoted by bold coefficients. Standard errors are in parentheses. Instrumental Variables Quantile Regression estimates are obtained with the instrument "mother's education" coded as 1 is individual's mother has obtained at least a high school diploma, and 0 is otherwise.

[a] Control variables include gender (1 is male, 0 is female) and origin (1 is foreign origin, 0 is Belgian origin).

24

REFERENCES

Aakvik, A., Salvanes, K. G., & Vaage, K. (2010). Measuring Heterogeneity in the Returns to Education in Norway Using Educational Reforms. *European Economic Review, 54(4)*, 483-500.

Acemoglu, D., & Angrist, J. D. (2000). How Large are Human-Capital Externalities? Evidence from Compulsory-Schooling Laws. *NBER Macroeconomics Annual, 15*, 9-59.

Angrist, J. D., & Krueger, A. B. (1991). Does Compulsory School Attendance Affect Schooling and Earnings? *Quarterly Journal of Economics, 106(4)*, 979-1014.

Balestra, S., & Backes-Gellner, U. (2017). Heterogeneous returns to education over the wage distribution: Who profits the most? *Labour Economics, 44*, 89-105.

Battistin, E., De Nadai, M., & Sianesi, B. (2014). Misreported schooling, multiple measures and returns to educational qualifications. *Journal of Econometrics, 181(2)*, 136-150.

Bingley, P., & Martinello, A. (2017). Measurement Error in Income and Schooling and the Bias of Linear Estimators. *Journal of Labor Economics, 35(4)*, 1117–1148.

Brunello, G., & Miniaci, R. (1999). The economic returns to schooling for Italian men. An evaluation based on instrumental variables. *Labour Economics, 6(4)*, 509-519.

Clark, D., & Martorell, P. (2014). The Signaling Value of a High School Diploma. Journal of *Labor Economics, 122(2)*, 282-318.

Dickson, M., & Harmon, C. (2011). Economic returns to education: What We Know, What We Don't Know, and Where We Are Going – Some brief pointers. *Economics of Education Review, 30(6)*, 1118-1122.

Devereux, P. J., & Fan, W. (2011). Earnings returns to the British education expansion. *Economics of Education Review, 30(6)*, 1153-1166.

Eurostat. (2017, May 1). Early Leavers From Education and Training. Luxembourg.

Fersterer, J., & Winter-Ebmer, R. (2003). Are Austrian returns to education falling over time? *Labour Economics, 10(1)*, 73-89.

Frölich, M., & Melly, B. (2013). Unconditional Quantile Treatment Effects Under Endogeneity. *Journal of Business & Economic Statistics, 31(3)*, 346-357.

Grenet, J. (2013). Is Extending Compulsory Schooling Alone Enough to Raise Earnings? Evidence from French and British Compulsory Schooling Laws. *Scandinavian Journal of Economics, 115(1)*, 176-210.

Harmon, C., & Walker, I. (1995). Estimates of the Economic Return to Schooling for the United Kingdom. *American Economic Review, 85(5)*, 1278-1286.

Heckman, J. J., & Li, X. (2004). Selection Bias, Comparative Advantage and Heterogeneous Returns to Education: Evidence from China in 2000. *Pacific Economic Review, 9(3)*, 155-171.

Holmlund, H., Lindahl, M., & Plug, E. (2011). The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods. *Journal of Economic Literature, 49(3)*, 615-651.

Hoogerheide, L., Block, J. H., & Thurik, R. (2012). Family background variables as instruments for education in income regressions: A Bayesian analysis. *Economics of Education Review, 31(5)*, 515-523.

Ikeda, M., & Garcia, E. (2014). Grade repetition: A comparative study of academic and non-academic consequences. *OECD Journal: Economic Studies, 2013(1)*, 269-315.

Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics, 133(1)*, 97-126.

Meghir, C., & Palme, M. (2005). Educational Reform, Ability, and Family Background. *American Economic Review, 95(1)*, 414-424.

Olea Montiel, J. L., & Pflueger, C. (2013). A Robust Test for Weak Instruments. *Journal of Business & Economic Statistics, 31(3)*, 358-369.

Oosterbeek, H., & Webbink, D. (2007). Wage Effects of an Extra Year of Basic Vocational Education. *Economics of Education Review, 26*, 408-419.

Oreopoulos, P. (2006). Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *American Economic Review, 96(1)*, 152-175.

Oreopoulos, P., & Salvanes, K. G. (2011). Priceless: The Nonpecuniary Benefits of Schooling. *Journal of Economic Perspectives, 25(1)*, 159-184.

Pischke, J.-S., & von Wachter, T. (2008). Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation. *Review of Economics and Statistics, 90(3)*, 592-598.

Sakellariou, C., & Fang, Z. (2016). Returns to schooling for urban and migrant workers in China: a detailed investigation. *Applied Economics, 48(8)*, 684-700.

Staiger, D., & Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica, 65(3)*, 557-586.

Stephens, M., & Yang, D.-Y. (2014). Compulsory Education and the Benefits of Schooling. *American Economic Review, 104(6)*, 1777-1792.

Stewart, M. B. (1983). On Least Squares Estimation when the Dependent Variable is Grouped. *Review of Economic Studies, 50(4)*, 737-753.

Stock, J. H., & Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In J. H. Stock, & D. W. Andrews, *Identification and inference for Econometric Models: Essays in Honor of Thomas Rothenberg* (pp. 80-108). New York: Cambridge University Press.

Wang, L. (2013). Estimating returns to education when the IV sample is selective. *Labour Economics, 21*, 74-85.