Does a High School Diploma Matter? Evidence Using Regression Discontinuity Design

By DENI MAZREKAJ AND SOFIE CABUS*

We study whether education raises productivity (human capital theory) or just reflects it (sorting theory) by estimating returns to a high school diploma. For this purpose, we exploit standardized exit exams students must take in the final year of secondary education to obtain a diploma in the Netherlands. Using a regression discontinuity design on administrative population data, we compare earnings of students who barely passed and barely failed standardized exit exams. The results indicate a positive effect of a high school diploma on earnings of about 0.34 EUR per hour. We interpret this finding as evidence of diploma sorting effects. (JEL I26, J24, J31).

* Deni Mazrekaj (corresponding author): Leuven Economics of Education Research (LEER), KU Leuven, Naamsestraat 69, Leuven, 3000 (e-mail: deni.mazrekaj@kuleuven.be); Sofie Cabus: HIVA Research Institute for Work and Society, KU Leuven, Parkstraat 47, post office box 5300, Leuven, 3000 (e-mail: sofie.cabus@kuleuven.be) and affiliated researcher with Maastricht University, Kapoenstraat 2, Maastricht, 6211 KV. The data for this study are protected by a confidentiality agreement and we are precluded from sharing the data with others. Interested readers can contact the corresponding author for information on how to obtain access to the data. We would be happy to provide assistance and Stata code to replicate the results of this paper. We would like to thank Eric Hanushek, Kristof De Witte and the seminar participants at the London School of Economics and Political Science, KU Leuven, Maastricht University, Politecnico di Milano, University of Lisbon, and University of Antwerp for helpful comments and suggestions. This work was supported by Institute Gak and the European Union's Horizon 2020 program for twinning [grant number 691676]. Deni Mazrekaj acknowledges funding by the Research Foundation Flanders (FWO) as Aspirant [grant number 1172519N]. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

Keywords: Human Capital, Sorting, High School Diploma, Standardized Exit Exams, Regression Discontinuity Design.

1. Introduction

It is well documented that more educated persons tend to earn higher wages and that this education premium has remained relatively stable over time (see review by Psacharopoulos and Patrinos 2018). The reason for this education premium, however, has been a source of considerable debate. Two dominant economic theories have emerged. According to the human capital theory (Becker 1962; Schultz 1961), persons invest in education to raise their productivity. Because of this increased productivity, more educated persons earn higher wages. By contrast, in the sorting theory (Arrow 1973; Spence 1973), education has no impact on productivity. Confronted with imperfect information about workers' productivity, employers may use education as a filter to screen workers into more productive (more educated) and less productive (less educated) groups, paying higher wages to the more educated workers. In turn, more productive persons will choose higher levels of education not to raise, but to signal their productivity.[1] In practice, distinguishing between the human capital and the sorting theory is difficult as both theories predict a positive effect of education on wages.

Whether education raises productivity or just reflects it is an important question for economic policy. The human capital theory and the sorting theory have different policy implications. If the human capital theory holds and education raises productivity, private returns will reflect social returns to education, and policy should aim to increase workers' education levels by removing barriers that prevent persons from acquiring their desired level of education. On the other hand, if the sorting theory holds and education is merely a signal of productivity, aggregate income may remain unchanged even if all workers raise their education levels. This

[1] As suggested by Weiss (1995), we will use the term "sorting" to refer to both screening and signaling of workers. In addition, we abstract from different versions of the two theories such as weak versus strong sorting and pure human capital theory versus human capital theory that allows sorting. In this paper, "human capital theory" refers to the pure human capital theory, and "sorting theory" refers to the pure sorting theory, meaning that one theory excludes the other.

implies that, under sorting, the social returns may be lower than the private returns, calling into question the rationale for public investment in education.

In this paper, we aim to distinguish between the human capital and sorting theories by estimating the earnings returns to a high school diploma. As opposed to years of schooling, a high school diploma is unlikely to affect productivity, as it is essentially only a piece of paper (Clark and Martorell 2014). Therefore, if we were to randomly assign a high school diploma, we could capture the sorting value of a diploma by estimating the difference in earnings between students who obtained a diploma, and students who did not. To approximate the random assignment of diplomas, we exploit the standardized exit exams students must take at the end of the final year of secondary education to obtain a high school diploma in the Netherlands. These exit exams are the last obstacle before graduation. Using a fuzzy Regression Discontinuity (RD) design (Hahn, Todd, and Van der Klaauw 2001), we compare the earnings of students who barely passed and barely failed standardized exit exams. If the assumptions of a fuzzy RD are satisfied, these students are likely to have similar levels of human capital, but different diploma status.

Most previous studies that aimed to distinguish between the human capital and the sorting theory suffer from omitted-variables bias and therefore produce only correlational evidence (see Section 2 for a thorough overview). To solve this selection bias, more recent studies have attempted to find exogenous variation in sorting status of a degree among workers with similar levels of human capital. Tyler, Murnane, and Willett (2000) estimated the sorting value of the General Educational Development (GED) by exploiting different passing standards across the U.S. states in a difference-in-differences framework. Similarly, Jepsen, Mueser, and Troske (2016) used the GED passing threshold in a fuzzy RD design. Other studies used RD designs to study the earnings returns to degree classification such as graduating with honors (Di Pietro 2017; Feng and Graetz 2017; Khoo and Ost 2018).

To the best of our knowledge, only one study estimated the sorting value of a high school diploma quasi-experimentally (Clark and Martorell 2014). Using a similar empirical strategy as in this paper, the authors compared the earnings of students that barely passed and barely failed the high school exit exams in Texas. They found no evidence of diploma sorting effects and no heterogeneity by gender nor race. However, this study suffers from an important limitation. Students take exit exams in Texas for the first time in 10th grade and can retake the exams several times before the 10th and 12th grade. The scores on the 10th grade exit exams are endogenous as the scores could influence the length of schooling or the curriculum in later years. As a result, Clark and Martorell (2014) focused on a small proportion (4.83%) of students that took exit exams at the end of 12th grade. These students have already failed the exam at least once, and often several times. Moreover, they have a lower socioeconomic status than students who took exit exams only once (Clark and Martorell 2014). Therefore, the results in this study are unlikely to hold for the majority of the population.

Using administrative data for the entire population of non-vocational education students in the Netherlands, we are able to solve this problem. In the Netherlands, standardized exit exams take place at the end of the final grade of secondary education for most students[2]. Therefore, we can include all students who have taken standardized exit exams in 12th grade. Moreover, our large administrative data include scores on the exit exams between school year 2005-2006 and school year 2016-2017 for each student. This poses two advantages. First, we can identify

[2] The Netherlands includes three main tracks: pre-university education, general secondary education and vocational education. Exit exams take place at the end of the 12th grade for students in pre-university education and general secondary education. Students in vocational education take standardized exit exams in the 8th grade. Therefore, in this study, we only consider students in pre-university education and general secondary education (74% of students) to avoid selection issues.

students who have retaken exit exams in later school years. Second, we can use initial scores on the standardized exit exams instead of the final scores in a fuzzy regression discontinuity design. This is particularly important, as Jepsen et al. (2016) have shown that using final scores in an RD design leads to biased estimates of the earnings returns to a credential.

At first sight, our results suggest that there is no earnings effect of a high school diploma. However, this is because most students who passed exit exams continued their schooling in post-secondary education. Once we focus on students who immediately entered the labor market upon leaving secondary education, we find a positive effect of a high school diploma on earnings of about 0.34 EUR per hour. Although this effect remains positive regardless of the gender, ethnicity, or track, we find a larger positive effect for boys and students who have completed a program in pre-university education. Therefore, we conclude that a high school diploma is likely to have a positive effect on earnings. We contemplate that this effect can be interpreted as a diploma sorting effect as the students who barely failed and barely passed standardized exit exams have a different diploma status, but are likely to have similar levels of human capital.

The remainder of the article is structured as follows. In Section 2, we offer a comprehensive review of the approaches adopted in the literature to separate the human capital effect from the sorting effect of education. In Section 3, we explain the Dutch education system with an emphasis on high school exit exams. In Section 4, we construct and describe the sample. In Section 5, we formulate the fuzzy RD model based on global polynomial methods and provide evidence that the assumptions underlying this model are likely to hold. In Section 6, we present the results and estimate several alternative specifications, such as a local linear and a local quadratic RD model. The article ends with a discussion of the results and several limitations of the study.

2. Identification of Sorting Effects in the Literature

Many attempts have been made in the literature to separate the productivity-enhancing effect from the sorting effect of education, dating back to 1970s when sorting theory was first introduced by Spence (1973) and Arrow (1973). Nonetheless, the methods applied were often unconvincing and the results mixed. The vast majority of studies is correlational and could not control for productivity differences between workers with higher and lower levels of education that are observable to firms but not to the econometrician. In the earliest approach, Taubman and Wales (1973) estimated earnings regressions within occupational categories and used these regressions to predict earnings of workers if they had been employed in other occupations. They found that many workers in occupations with low educational requirements earn less than their predicted wage in occupations with higher educational requirements. They interpreted this phenomenon as evidence for the sorting hypothesis. In a similar approach, Wiles (1974) proposed to compare the earnings of equally educated workers employed in occupations related and unrelated to their qualifications. If a wage premium is paid to workers employed in occupations related to their qualifications, the sorting hypothesis is rejected. On the other hand, if a worker with a qualification receives a wage premium irrespective of whether the qualification is relevant for the job, the sorting hypothesis is accepted. Using this method, Miller and Volker (1984) found evidence in favor of sorting, while the results in Arabsheibani (1989) supported the human capital theory. Nonetheless, none of these studies could control for potential selection bias arising from workers choosing their occupations.

In an attempt to improve the approaches by Taubman and Wales (1973) and by Wiles (1974), many studies emerged that compared the returns to education for a “screened” and an “unscreened” group. The screened group includes workers who are likely to benefit from signaling their ability as determining productivity by employers is costly. The unscreened group includes workers who are unlikely to benefit from signaling their ability as employers can determine productivity at a low cost. In an early attempt, Riley (1979) divided occupations in screened and unscreened groups based on his subjective judgment and found support for the sorting hypothesis. Psacharopoulos (1979), on the other hand, compared the rates of return to education in the private (unscreened) and the public (screened) sectors. He found that returns to education were higher in the private sector, rejecting the sorting hypothesis. Similarly, Wolpin (1977) found higher returns to education among the self-employed (unscreened) than among the privately employed workers (screened). Finally, Albrecht (1981) found that the returns to education were similar between workers who were hired through informal channels such as employee recommendations (unscreened) and workers who were hired through formal channels such as advertisements (screened). These findings provide evidence for the human capital hypothesis. However, similar to the studies that used the Wiles (1974) approach, none of the studies that compared workers in screened and unscreened groups were able to control for potential selection into the two groups. Therefore, the validity of these studies is questionable.

Another strand of literature that produced correlational evidence has focused on employer learning. These studies estimate the returns to education and the returns to ability over time. As employers obtain better information about their employees' real productivity, the returns to education should fall and the returns to ability should rise with experience, if the sorting hypothesis holds. Using this approach, Layard and Psacharopoulos (1974) found that returns to education did not decrease, but increase with experience. Similarly, Lange (2007) estimated that employers learn quickly and most expectation errors in productivity decline by 50 percent within three years. Lange (2007) therefore concluded that sorting is rather limited. By contrast, the findings in Farber and Gibbons (1996), Altonji and Pierret (2001), Arcidiacono, Bayer, and Hizmo (2010), and Kahn and Lange (2014) are consistent with the predictions of the sorting hypothesis. Similar to previous studies, however, these studies are merely correlational and are therefore not able to control for unobserved factors that may be influencing the rate of return to education over time such as on-the-job training. Moreover, falling returns to education and rising returns to ability with experience are also consistent with the human capital theory if skills acquired during education become less relevant for the job over time.

Other correlational studies focused on the measures of education used in the wage regression. Most of these studies estimated sheepskin effects. These effects occur if wages rise faster with extra years of education, when the extra years also include a credential (for instance, the final year of secondary education). Positive sheepskin effects are then interpreted as evidence for the sorting hypothesis. Hungerford and Solon (1987), Belman and Heywood (1991) and Jaeger and Page (1996) found largest returns to education when the additional year of schooling was accompanied by a credential. However, these findings cannot be interpreted as support for the sorting hypothesis, as sheepskin effects are also consistent with the human capital theory (Flores-Lagunes and Light 2010). Human capital models can generate sheepskin effects if good learners are more likely to stay in school long enough to earn a credential, or if students acquire more skills in the years in which they receive a degree rather than in other years of schooling. Therefore, the ability of these studies to distinguish between the two competing theories is limited.

Two related approaches that focused on the measures of education are also worth mentioning. Groot and Oosterbeek (1994) divided actual years of schooling into effective years (number of formally required years to obtain a degree), repeated years, skipped years, inefficient routing years (number of years students switched programs inefficiently), and dropout years (years spent in education without obtaining a degree). Under the sorting hypothesis, more rapid completion of a degree should signal greater ability and should therefore increase earnings. Moreover, years spent in education without obtaining a degree (dropout years) should not increase earnings. On the other hand, Kroch and Sjoblom (1994) included both an absolute and a relative measure of education in a wage regression. The former includes grade level, and the latter the student's rank in the distribution of educational attainment for the entire cohort. If the sorting hypothesis holds, the returns to a student's rank would be positive while the returns to grade level should be zero. Both studies found little support for the sorting hypothesis. Nonetheless, these studies are prone to omitted variables bias and bias arising from measurement error in the different schooling variables.

More advanced empirical approaches have used policy changes to distinguish between the human capital and the sorting hypothesis, albeit not with quasi-experimental methods. Lang and Kropp (1986) proposed that in a sorting model, an increase in the minimum compulsory schooling age would raise educational attainment of students not directly affected by the law. For instance, consider a law which increases minimum compulsory schooling age from 16 to 17 years. If a student who would have left education at the age of 16 is now forced to stay in school until the age of 17, the average ability of all 17-year-old students who remained in school will fall. Consequently, the most able of these students will remain in school until the age of 18. Thus, under sorting, although the compulsory schooling law was meant to affect 16 year-old-students, it also affected high ability 17-year-old students who are now studying longer to signal their ability. Lang and Kropp (1986) provided evidence supporting the sorting hypothesis in the United States. By contrast, using the same approach, Chevalier et al. (2004) found no evidence of sorting in the United Kingdom. An increase in the compulsory schooling age, however, might affect students not directly affected by the law even if no sorting. This can occur if forcing students to stay in school longer teachers them that school is important, or if having weaker students in class lowers the quality of education. Therefore, this approach is unlikely to separate human capital effects from sorting.

A similar approach based on increased university access was proposed by Bedard (2001). She argued that, under the sorting hypothesis, an expansion of the university system would lower the cost of entering a university, but also increase the high school dropout rate. Because of increased university access, high-ability graduates who were previously constrained from entering a university due to high cost would now enroll in a university. This would decrease the average quality of high school graduates and employers would realize this. As a result, low-ability students would have less incentive to pool themselves with the high-school graduates group and would leave school. This is in contrast with the human capital hypothesis that predicts only an upward movement in educational attainment. Bedard (2001) indeed found evidence supporting the sorting hypothesis in the US. However, the decision to drop out from high school, graduate, or enroll into a university is not random. Therefore, the results in Bedard (2001) suffer from selection bias.

It is clear that most studies in the literature produced correlational results, prone to selection bias. Several studies, however, attempted to find exogenous variation in sorting status of a degree among workers with similar levels of human capital. Tyler et al. (2000) estimated the sorting value of the General Educational Development (GED) credential in the United States, by exploiting different passing standards across states. In particular, they compared workers from different states with equal GED scores, but different GED status because of differences in the stringency of the passing standards in their state of residence. The results from their difference-in-differences analysis suggest that GED has a large sorting value for white dropouts, but not for minority dropouts.

Other quasi-experimental studies used regression discontinuity designs. Jepsen et al. (2016) exploited the GED passing threshold in a fuzzy RD design and found sorting effects for men right after graduation, but not for women. Other studies used RD designs to study the returns to degree classification (Di Pietro 2017; Feng and Graetz 2017; Khoo and Ost 2018). These studies generally found positive sorting effects. Finally, in the only study to estimate the sorting value of a high school diploma, Clark and Martorell (2014) used an RD design to compare the earnings of workers who barely passed and barely failed high school exit exams in Texas. They found no evidence of diploma sorting effects.

3. Institutional Setting

3.1. The Dutch Education System

The Dutch education system provides for compulsory education beginning at the age of five and continuing either until the age of 18 or until a younger age if a student has already obtained a high school diploma (Government of the Netherlands 2018c). Primary education lasts for 7 years, from the age of five until the age of 12. Nonetheless, most parents already enroll their children into primary education at the age of four. In the last year of primary education, pupils take a standardized cognitive test measuring their skills in language and mathematics. In general, this test is compulsory for all pupils, although small deviations exist such as for students in special education. This test is not an exam; pupils cannot pass or fail the test. Instead, the goal of the test is (a) to advise pupils' parents on a secondary education track in which they should enroll their children, and (b) to advise schools on the ability of pupils. In addition to the cognitive test, pupils also receive an advice of the primary school teacher on which track to follow in secondary education.

At the age of 12, students enter a tracking system in secondary education. Secondary schools decide on whether to accept students based on the advice of the student's primary school teacher and the cognitive test conducted at the end of primary education. In general, secondary schools do not set their own entrance exams, except in exceptional circumstances, such as dance, music, or sports programs. There are three tracks. Pre-university education (VWO) lasts for six years and prepares students for university education (WO). This track is perceived as the most prestigious track and mainly includes high-performing students. At a university, an academic bachelor's degree program typically lasts for three years. General secondary education (HAVO) lasts for five years and prepares students for higher professional education (HBO) offered at a university college. A professional bachelor's degree program at a university college typically lasts for four years. Students who successfully finished a general secondary education program cannot enroll in a university, but have to complete a higher professional education program first. Furthermore, students who did not obtain a high school diploma cannot enroll into a university nor a university college. Nonetheless, school dropouts can continue studying in an adult education (VAVO) program.

Pre-vocational secondary education (VMBO) lasts for four years and prepares students for secondary vocational education (MBO). This track is typically perceived as the least prestigious track. Pre-vocational secondary education is further divided into four programs with a decreasing level of difficulty. In a theoretical program (TL), students mainly take general subjects. In a combined program (GL), one of the general subjects is replaced by four hours of vocational training. Other general subjects are taught at the same level as in a theoretical program. In a middle-management vocational program (KBL), students receive 12 hours of vocational training, and general subjects are taught at a slightly lower level than in the theoretical program. Finally, in a basic vocational program, students also receive 12 hours of vocational training. However, general subjects are taught at a lower level than in the middle-management vocational program.

After four years of pre-vocational secondary education, students can enroll in secondary vocational education (MBO). This track prepares students for a certain occupation and lasts for up to four years, depending on the level of training. There are four levels. Assistant training (level 1) lasts for 1 year. Basic vocational training (level 2) lasts for one or two years depending on the specific program. Professional training (level 3) lasts for two or three years. Finally, middle-management training (level 4) lasts for three or four years. To enroll in professional and middle-management training, students need to have completed a theoretical, combined, or middle-management vocational program in pre-vocational secondary education. Students who have successfully completed a middle management training (level 4) can continue their studies in higher professional education at a university college. For each level in secondary vocational education, there are two learning pathways. In vocational training (BOL), practical training at a company typically takes up between 20% and 60% of the program duration. In block or day release (BBL), student is employed at a company for more than 60% of the program duration.

3.2. High School Exit Exams and Diplomas in the Netherlands

To obtain a high school diploma, a student must successfully complete a program in either pre-university education, general secondary education, or in basic vocational training (secondary vocational education, level 2). Thus, students who left secondary education before completing a program in one of these three tracks are considered high school dropouts. The dropout rate in school year 2016-2017 was 6.95% (Government of the Netherlands 2018d). Most students drop out from secondary vocational education (9.89%), followed by general secondary education (6.53%). Least students drop out from pre-university education (4.11%). This is unsurprising as pre-university education is the most prestigious track and therefore mainly includes high performing students.

To successfully complete a program in secondary education, all students need to pass the minimum number of required courses set by national law. In the final year of a secondary education track, students conduct two exams per course for the majority of courses: a school exit exam and a standardized exit exam. Nonetheless, some elective courses include only a school exit exam. As the name suggests, school exit exams are set by the schools and therefore differ per school. These exit exams can be oral, practical, or written and are graded by schoolteachers. By contrast, standardized exit exams are standardized national exams compiled by the Dutch Ministry of Education, Culture and Science. One standardized exit exam is administered per course for all pupils in the same track. This exit exam is mostly written[3] and can contain both multiple-choice as open questions. It is either graded by a computer or by appointed external examinators. Standardized exit exams are conducted in pre-university education, general secondary education, and pre-vocational education. Reader should note that secondary vocational education does not include standardized exit exams, but only school exit exams. As mentioned previously, students can only obtain a high school diploma in pre-university education, general secondary education, and basic vocational training (secondary vocational education, level 2). Therefore, students enrolled in pre-vocational education need to continue their studies in at least basic vocational training to qualify for a high school diploma and this track does not include standardized exit exams.

Both school exit exams and standardized exit exams are graded out of possible 10 points. The average number of points on the school exit exam and the standardized exit exam is considered as the final grade for a particular course. None of these scores are made public[4]. Although school exit exams and standardized exit exams are rounded to 1 digit after the comma, final grades are rounded to whole digits. To complete a program, a necessary requirement is that the average passing score on the standardized exit exams for all courses should be at least 5.5 out of possible 10 points. However, this requirement is not sufficient, and several other criteria need to be taken into consideration (see Government of the Netherlands

[3] Some courses in pre-vocational secondary education include a "practical" test in which students go to a separate examination center and perform practical tasks in front of an examinator. These tests are also graded out of possible 10 points.

[4] In Section 7, we dwell deeper on the issue of whether employers can acquire information about the exam scores.

(2018b) for a detailed list of criteria). Therefore, for the purpose of this article, it is important to note that it is possible that a student with an average score on the standardized exit exams of 5.5 or higher does not receive a high school diploma. On the contrary, it is legally not possible that a student with an average score on the standardized exit exams of less than 5.5 still receives a high school diploma.

Before the start of a standardized exit exams, students are informed about the scores on their school exit exams. Thus, although both exit exams are conducted at the end of the final school year and can therefore not determine the curriculum studied, school exit exams precede standardized exit exams. Therefore, standardized exit exams are the last obstacle before program completion. In general, students can take a standardized exit exam in all subjects only once per year, in the second half of May. However, students can retake a standardized exit exam in one course in the second half of June. In exceptional circumstances such as illness, students can also take standardized exit exams in the second half of August. Some students who are not enrolled in a school anymore such as older school dropouts from earlier school years, students in the military and imprisoned students can also take the so-called state exam in August. This exam is equivalent to a standardized exit exam for regular students. In all cases, the highest of the three scores counts. In school year 2016-2017, we observe in our population data (see further) that 99.98% out of 203,310 students took standardized exit exams in May for the first time. Furthermore, 82.90% took standardized exit exams only once and 17.10 retook a standardized exit exam for one of the courses in later periods (June or August).

As mentioned previously, students can only complete a program if they obtain an average score for the standardized exit exams of at least 5.5 out of possible 10 points. If a student does not reach this threshold, this student needs to retake standardized exit exams for all courses again in the next school year, regardless of whether the student passed an individual course. Students can retake standardized exit exams in later school years at most twice. For instance, from the students who took a exit exam in school year 2011-2012 for the first time and failed, 37.49% retook standardized exit exams once in the next school year and 0.14% retook standardized exit exams twice. Moreover, students who failed to obtain a diploma but passed several courses do not receive a certificate as an alternative, except in rare circumstances such as students in the military or imprisoned students who failed the state exam.

4. Data

4.1. Sample Construction

We use administrative education records collected by Statistics Netherlands that cover the entire Dutch population of students. In particular, we observe all educational enrolments starting from school year 2000-2001 until school year 2016-2017. Each student has been given a unique personal identification number, allowing us to follow a student's educational path over the school years. Per student/school year record, we observe the specific program and the track a student was enrolled in (secondary education, higher education and adult education) as well as the high school graduation status coded as *dropped out*, *graduated*, *continued to the next school year*. Moreover, for school years ranging from 2005-2006 until 2016-2017, we also observe (a) whether a student took standardized exit exams, (b) when the student took standardized exit exams (May, June, or August), (c) the score on the standardized exit exams for each time period averaged over the courses, (d) the score on the school exit exams averaged over the courses, and (e) the final score calculated as the average of the school exit exams and the standardized exit exams over all the courses.

Using personal identification numbers, we link these data to population registers that contain demographic characteristics of the entire Dutch population. This allows us to observe the date of birth, gender, neighborhood of residence, and birth country of the student as well as the birth country of the parents. Furthermore, we link these data to labor market information from tax authorities. These data include the job history of the entire Dutch population between 1999 and 2016 as well as gross earnings per job per day, and taxes paid. Therefore, we can observe average earnings per day after leaving high school for most students. Finally, we also construct a measure of a student's socioeconomic status. For this purpose, we use the population registers to link students to their parents. Subsequently, we merge these parental data with tax records allowing us to calculate annual earnings of the parents.

The obvious strengths of our administrative dataset are a very large sample (entire population) and likely little measurement error in the variables of interest compared to survey data. Additionally, we observe the average score on the standardized exit exams per school year, allowing us to identify students who have retaken standardized exit exams in later years. A limitation of our dataset is that we do not observe the score on the standardized exit exam for each course. We solely observe the score on standardized exit exams averaged over the courses. An additional limitation is that we cannot observe long-run outcomes.

To study earnings returns to a high school diploma, we restrict the sample in three ways. First, we solely consider students in pre-university education and in general secondary education. Therefore, we do not consider students who took standardized exit exams in pre-vocational education. This is because of two reasons. First, this track does not lead to a high school diploma. Students who successfully complete a program in the pre-vocational education need to continue their studies to at least basic vocational training (secondary vocational education, level 2) to obtain a high school diploma. However, as mentioned previously, secondary vocational education does not include standardized exit exams. Therefore, at least two years have passed since the moment these students took their initial standardized exit exams and the moment they left secondary education. As a result, students with an equal average score on the standardized exit exams may have accumulated a different amount of human capital in the years after the standardized exit exam. The second reason is that standardized exit exams are equal for all students within the same track. Therefore, comparing scores within a track is warranted. For students in pre-vocational education, however, the average score on the standardized exit exams often serves as a mechanism to choose which level in secondary vocational education to follow, leading to selection effects. As a result, including a variable for both the average score on the standardized exit exams and a variable for the track would lead to post-treatment bias (Rosenbaum, 1984). In sum, the results in this paper do not allow us to draw conclusions about the value of a high school diploma for students in vocational education. As these students have a higher dropout rate (see above) than students in the other two tracks, it is unclear whether our results can be generalized for this population of students.

As a second sample restriction, we study cohorts who took their initial standardized exit exams between school year 2007-2008 and school year 2012-2013. Before 2007-2008, we do not observe whether a student took standardized exit exams multiple times. Similarly, after, 2012-2013, we cannot study labor market outcomes. Consider, for instance, a student who took standardized exit exams for the first time in school year 2013-2014 and failed two times. This student is allowed to try one more time, in 2015-2016. If the student passes the exit exams and obtains a high school diploma, this student will enter the labor market in 2017[5]. However, we observe labor market outcomes until the year 2016. Therefore, we restrict the sample to observe students' labor market outcomes one year after leaving school. Lastly, we remove a small percentage of students (0.24%) with missing values on the outcomes, variables of interest, or one of the covariates. The

[5] Technically, a student can also enter the labor market at the end of 2016.

final sample includes 435,768 students who took their initial standardized exit exams between school year 2007-2008 and school year 2012-2013. We observe the outcomes of these students one year after leaving school.

4.2. Variable Construction

Outcome Variables. We aim to distinguish between the human capital and the sorting hypothesis by estimating the returns to a high school diploma. Therefore, the outcome of interest is productivity, proxied by the logarithm of the average net earnings per hour in the first year after graduation. We calculate this variable in four steps. First, we calculate total net earnings per year as the sum of gross earnings over all jobs within a year minus the amount of taxed paid. Second, we multiply the number of days worked by a part-time factor (for instance, 0.5 means employed at 50% and 1 means employed full time) to obtain the number of Full Time Equivalents (FTE). Next, we divide the net earnings per year by the number of FTE worked and by 8 (usual number of working hours per day). In the last step, we take the logarithm of this variable. In supplemental analyses available online (Table S1), we also reproduced our results using the net earnings level (not log), net earnings per year instead of per hour, log of gross earnings per hour, and log of gross earnings per year.

Standardized Exit Exam Score. As mentioned previously, students obtain three exit exam scores at the end of secondary education: the average score on the school exit exams, the average score on the standardized exit exams, and the final score defined as the average of the average score on the school exit exams and the average score on the standardized exit exams. We have chosen for the average standardized exit exam score as the exam score of interest given that it is nationally determined, externally graded and includes a clear threshold of 5.5 out of 10 points, which students need to obtain to receive a high school diploma. School exit exams, on the other hand, are set by schools, teacher graded and thus could be manipulated, and do not include a clear cut off. Similarly, the final score also does not contain a clear cut off and it would ignore the minimum requirement students need to attain on the standardized exit exams. Furthermore, to account for students retaking the standardized exit exams, we follow Jepsen et al. (2016) and solely consider the average score on the initial standardized exit exams a student took. We construct two variables. The first is a continuous variable for the average score on the initial standardized exit exams a student took, centered to be zero at passing threshold. The second variable is an indicator given value of 1 if a student scored at or above the threshold and 0 if the student scored below the threshold.

High School Diploma, Track, and Year of Exam. Our data include an indicator for high school dropout per school year. It includes three categories: dropped out, graduated, or continued to the next school year. Therefore, we measure high school diploma as an indicator given value of 1 if a student graduated in the school year of taking the standardized exit exams, and 0 if a student dropped out. If students continued their studies in high school, we follow them for two more school years as they are allowed to retake standardized exit exams twice. If students still continued their studies in high school after these two school years without graduating, they are counted as school dropouts (this is less than 0.01% of students). Furthermore, we include a variable for the track student took standardized exit exams in, because standardized exit exams are equal for all students within the same track. However, they may differ across tracks. We code this variable as an indicator given value of 1 for pre-university education and value of 0 for general secondary education. Finally, we also include indicators for the year students took their initial standardized exit exams to account for exams differing by year.

Control Variables. Although not strictly necessary for our empirical strategy (see further), we also construct several control variables. We include an indicator for gender (1 is boy, 0 is girl), a continuous variable for age at initial standardized exit exams, and age squared. In addition, we include three variables as a measure of socioeconomic background. First, we construct a variable for child's ethnicity based on parental birth country. It is given value of 1 if at least one parent was born outside the Netherlands and value of 0 if both parents were born in the Netherlands. Second, we include indicators for neighborhood of residence. Third, we include a continuous measure for parental annual net income in 10,000 euros one year before taking standardized exit exams. We calculate this variable as the total annual net income of the father and the mother. Both neighborhood of residence as parental income are measured one year before taking standardized exit exams to avoid potential endogeneity arising from post-treatment bias (Rosenbaum 1984).

4.3. Descriptive Statistics

Descriptive statistics are presented in **Table 1**. We observe that 7.9% of the students dropped out from high school after their last standardized exit exams. This dropout rate is very similar to the dropout rate in the official Dutch statistics presented earlier. We further observe in **Table 1** that about 79% passed the initial standardized exit exams. From the students that failed the initial standardized exit exams, about 60% retook at least one standardized exit exam either in the same school year or in one of the following school years. About 73% went to post-secondary education immediately after the last standardized exit exam. If we compare students who immediately entered the labor market after secondary education and therefore did not enroll into post-secondary education with the full sample, several important differences are worth mentioning. First, students who immediately entered the labor market are typically more disadvantaged than students in the full sample. They are more likely to be from foreign ethnicity, to be younger, to have younger parents, and less parental income. They are also more likely to be boys who typically have a higher dropout rate. Consequently, they also have a higher dropout rate than students in the full sample. These findings will have an important consequence on our analysis as students who enter post-secondary education typically do not earn a wage. Therefore, our results will be biased downward.

TABLE 1 – DESCRIPTIVE STATISTICS

	Full sample	Students not in post-secondary education	Difference (T-test)
Gender (1=boy, 0=girl)	0.489	0.446	0.043***
Ethnicity (1=foreign, 0=Dutch)	0.165	0.191	-0.026***
Age at first standardized exit exam	17.423	17.990	-0.567***
Age of the mother	45.215	48.659	-3.444***
Age of the father	48.119	51.298	-3.179***
Parental annual income in 10,000 EUR	3.671	4.182	0.511***
Passed initial standardized exit exams (1=yes)	0.785	0.661	-0.124***
Retook standardized exit exam (1=yes)[b]	0.598	0.423	0.175***
High school diploma (1=diploma, 0=dropout)	0.921	0.617	0.304***
Net earnings per hour in EUR[c]	2.984	3.961	-0.977***
Number of students	435,768	118,446	

Notes. Standard deviations are in parentheses.
[a] Share of students who retook at least one standardized exit exam from the total number of students who failed the initial standardized exit exams.
[b] These variables were measured one year after the last standardized exit exam.
*** Significant at the 1% level.

5. Empirical Strategy

5.1. Fuzzy Regression Discontinuity Design

The goal of the paper is to compare earnings of students with and without a high school diploma among students with similar levels of human capital. For this purpose, we use the average score on the standardized exit exams as a running variable in a fuzzy regression discontinuity design (Hahn et al. 2001). The fuzziness of the RD design arises from two sources. First, it is possible that a student with an average score on the standardized exit exams of 5.5 or higher does not receive a high school diploma due to an insufficient score on the school exit exams. Second, although legally students cannot obtain a diploma if they did not reach the threshold on the standardized exit exams on average, it is possible to retake one standardized exit exam in a second period within the same school year, and to retake all standardized exit exams in later school years. Given that we use the average score on the initial standardized exit exams as a running variable as recommended by Jepsen et al. (2016), students can obtain an insufficient average score on the initial standardized exit exams, but then still pass once they retake the exams. Therefore, the fuzziness of our approach arises from students who passed the standardized exit exams and still drop out of high school as well as from students who failed the standardized exit exams and still obtained a diploma.

Intuitively, we study the earnings difference between those students who barely passed and barely failed on the high school standardized exit exams. More formally, Hahn at al. (2001) show that fuzzy RD design can be formulated as a parametric Instrumental Variables (IV) model estimated by Two-Stage Least Squares (2SLS). Therefore, we use passing status on the standardized exit exams as an instrumental variable for diploma status in models that control for various polynomial orders of the average standardized exit exam scores (Clark and Martorell 2014). The outcome equation is formulated as follows:

$$(1)\quad Y_i = \beta_0 + \beta_1 DIPLOMA_i + \beta_2 TRACK_i + \beta_3 YEAR_i + g(p_i) + \boldsymbol{\delta X}_i + \epsilon_i$$

The corresponding first stage equation is formulated as follows:

$$(2)\quad D_i = \alpha_0 + \alpha_1 PASS_i + \alpha_2 TRACK_i + \alpha_3 YEAR_i + f(p_i) + \boldsymbol{\gamma X}_i + \varepsilon_i$$

In Equations (1) and (2), Y_i represents a labor market outcome, i.e. log net earnings per hour for student i, $DIPLOMA_i$ is the high school diploma status (1=diploma, 0=dropout), $PASS_i$ is an indicator for passing the standardized exit exams (1=passed, 0=failed), $TRACK_i$ is an indicator for the track (1=pre-university education, 0=general secondary education), $YEAR_i$ is an indicator for the year a student took the initial standardized exit exams, $\boldsymbol{X}_i$ is a set of control variables as specified above, and p_i is the average score on the initial standardized exit exams, centered at 0 for the threshold score if 5.5 out of 10 points. Furthermore, $g(p_i)$ captures the relationship between the log earnings and the average score on the initial standardized exit exams and $f(p_i)$ captures the relationship between the high school diploma status and the average score initial on the standardized exit exams.

To enable comparisons with other seminal papers in this literature (Clark and Martorell 2014; Jepsen et al. 2016), we use global polynomial methods to model the relationship between the log earnings and the standardized exit exam scores on the one hand, and diploma status and the standardized exit exam scores on the other. Gelman and Imbens (2018) advise against the use of higher order polynomials in RD designs. Therefore, in the main specifications, we will use polynomials of order one and two with and without interactions consistent with Jepsen et al. (2016). In supplemental analysis available online (Table S2), we also reproduced our results using higher order polynomials. Furthermore, we also interact test scores with the indicator for passing standardized exit exams to allow different functional forms on the two sides of the threshold. As recommended by Lee and Lemieux (2010), we use the same functional form for both the first-stage and the second-stage (outcome) equation. One possible concern is that global polynomial methods may give large weights to observations far away from the threshold (Gelman and Imbens 2018). Therefore, in Section 6.4., we reanalyze the data using local linear and local quadratic methods with a triangular kernel for various bandwidths.

5.2. Validity Checks

If the fuzzy RD assumptions hold, the variation in treatment near the threshold is randomized as though from a randomized experiment for the group of compliers (Lee and Lemieux 2010). Intuitively, these assumptions include: (a) a sizeable discontinuity in the exam scores (said otherwise, passing standardized exit exams should be a *strong instrument* for diploma attainment), (b) the passing status should only influence log earnings through a high school diploma (*exclusion restriction*), and (c) there should be no defiers (*monotonicity*). We consider each of these assumptions in turn.

Strong Instrument. In **Figure 1**, we show diploma status as a function of the average score on the initial standardized exit exams. As bin size, we consider the full range of scores as suggested by "optimal" bin sizes constructed using *rdplot* command in Stata 14 (Calonico, Cattaneo, and Titiunik 2014). We observe a large jump in diploma attainment at the threshold suggesting that the instrument is strong. Moreover, in Section 6.1., we present the empirical results for the first stage equation including the Kleibergen and Paap (2006) *F*-statistic. All first stage *F*-statistics are well above the weak instrument thresholds of Montiel Olea and Pflueger (2013).

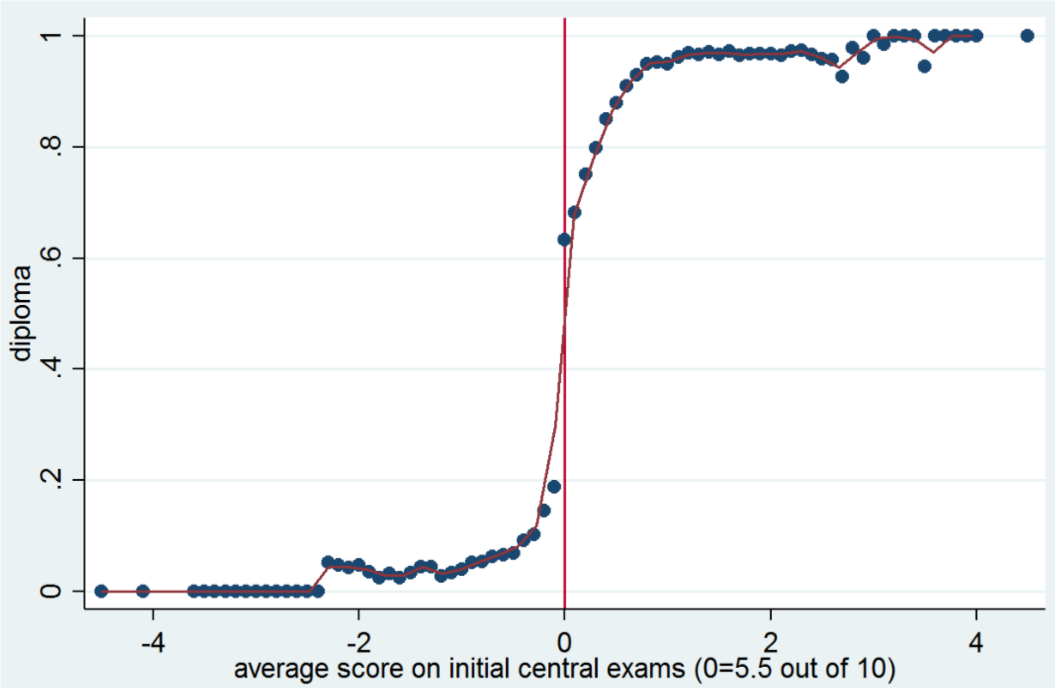

FIGURE 1 – DIPLOMA STATUS BY INITIAL STANDARDIZED EXIT EXAM SCORES. *Notes.* Diploma is an indicator given value of 1 if a student obtained a high school diploma and a value of 0 if a student dropped out from high school. Dots represent the share of students who obtained a diploma by initial standardized exit exam scores. The full range of scores is depicted on the x-axis. This bin size was suggested to be the "optimal" bin size by the *rdplot* command in Stata 14 (Calonico et al. 2014).

Exclusion Restriction. Ultimately, the exclusion restriction is untestable. However, we conduct several checks that suggest that the exclusion restriction is likely to hold in our setting. First, students should not be able to precisely manipulate the average score on the initial standardized exit exams. This assumption is very likely to hold in our setting as standardized exit exams are graded either by a computer or by external examinators, not by school teachers. In addition, we use the average score on the standardized exit exams, and students take a standardized exit exam for the majority of courses. Therefore, it is very unlikely that a student's intensity of studying could precisely manipulate this average score. To confirm this, we show the density of standardized exit exam scores in **Figure 2** as suggested by McCrary (2008). We find that the density of the standardized exit exam scores appears to be continuous near the threshold. Formally, we also use the manipulation test suggested by Catteneo, Jansson, and Ma (2018) and do not reject the null hypothesis of no manipulation (*p*-value is 0.488).

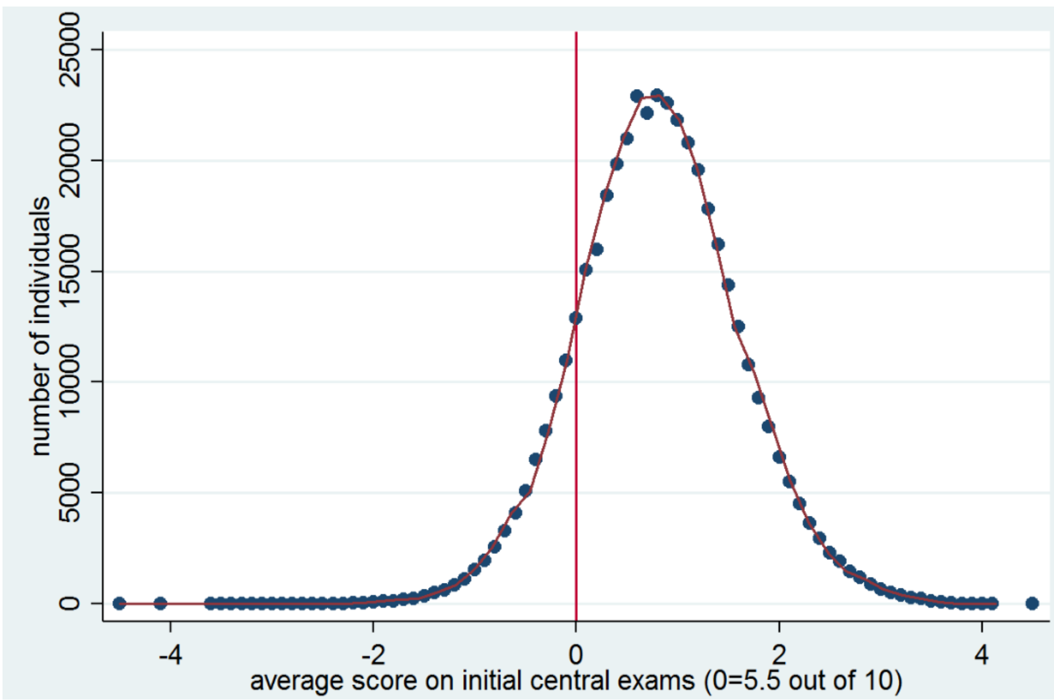

FIGURE 2 – DISTRIBUTION OF INITIAL STANDARDIZED EXIT EXAM SCORES.
Notes. Dots represent the number of students who obtained a given initial standardized exit exam score. The full range of scores is depicted on the x-axis. This bin size was suggested to be the "optimal" bin size by the *rdplot* command in Stata 14 (Calonico et al. 2014).

Second, we seek for potential other discontinuities that may reduce confidence in our model. Nonetheless, in **Figure 1**, we do not observe any discontinuities other than the discontinuity at the passing threshold of 5.5 out of possible 10 points. We formally test this by estimating our models for various placebo exit exam scores such as the median exam score below the threshold and the median exam score above the threshold. None of the coefficients were significant at the 5% level (see supplemental material available online, Table S3). Finally, we perform a balancing test on the covariates. Theoretically, it is not necessary to include covariates in an RD model as the variation in treatment near the threshold is as good as randomized if the crucial RD assumptions hold (Lee and Lemieux 2010). Nonetheless, covariates can be used to test whether there is a discontinuity in variables other than the treatment at the passing threshold. If this was the case, students on the left of the threshold would not be similar to students on the right of the threshold, invalidating our RD design. In particular, we run the following model on each covariate:

$$(3) \quad X_i = \theta_0 + \theta_1 PASS_i + \theta_2 TRACK_i + \theta_3 YEAR_i + h(p_i) + \mu_i$$

In Equation (3), X_i represents a covariate, e.g. gender for student i, $PASS_i$ is an indicator for passing standardized exit exams (1=passed, 0=failed), $TRACK_i$ is an indicator for the track (1=pre-university education, 0=general secondary education), $YEAR_i$ is an indicator for the year a student took the initial standardized exit exams, and h_i is the average score on the initial standardized exit exams centered at 0 for the threshold score if 5.5 out of 10 points. Finally, $h(p_i)$ captures the relationship between the covariate and the average score on the initial standardized exit exams. We use polynomials of order one and two with and without interactions as in the main analyses (higher order polynomials yield analogous results). Results shown in **Table 2** indicate that the discontinuity is not significant in any of the covariates at the passing threshold. This indicates that students near the threshold are similar. In sum, although the exclusion restriction ultimately cannot be tested, our analyses suggest that it is likely to hold in our setting.

TABLE 2 – TESTS OF COVARIATE BALANCE

	Discontinuity estimate			
	(1)	(2)	(3)	(4)
Gender (1=boy, 0=girl)	0.005	0.004	0.005	0.006
	(0.007)	(0.007)	(0.008)	(0.007)
Ethnicity (1=foreign, 0=Dutch)	0.002	0.001	0.004	-0.002
	(0.004)	(0.004)	(0.006)	(0.004)
Age at first standardized exit exam	0.002	0.002	0.004	0.003
	(0.010)	(0.009)	(0.009)	(0.007)
Age of the mother	0.028	0.011	0.021	0.020
	(0.061)	(0.055)	(0.078)	(0.081)
Age of the father	0.014	0.012	0.011	0.012
	(0.072)	(0.081)	(0.080)	(0.077)
Parental net annual income	0.041	0.038	0.067	0.032
	(0.043)	(0.044)	(0.058)	(0.068)
Polynomial order	linear	linear interaction	quadratic	quadratic interaction
Number of students	435,768	435,768	435,768	435,768

Notes. Robust standard errors are in parentheses. All specifications include an indicator for the track (1=pre-university education, 0=general secondary education).

Monotonicity. In a fuzzy regression discontinuity design, the results solely hold for *compliers* near the threshold (Lee and Lemieux 2010). In our study, these are the students who obtained a high school diploma because they barely passed the standardized exit exams, and who would not have done so if they barely failed the standardized exit exams. By contrast, the results in this paper do not allow us to draw conclusions about the value of a high school diploma for *always-takers*: students who would have obtained a diploma regardless of the passing status, by for instance retaking the standardized exit exams. In our data, we find that 9.4% of students who first failed standardized exit exams, eventually obtained a diploma by retaking them. Similarly, our results also do not hold for *never-takers*: students who would not have obtained a diploma regardless of whether they barely passed or barely failed the standardized exit exams, by for instance failing the school exit exams. In our data, we find that 11,5% of students who passed standardized exit exams, actually dropped out. Finally, our model assumes the absence of *defiers*: students who obtained a high school diploma because they barely failed standardized exit exams, and who would not have done so if they barely passed standardized exit exams. Evidently, this behavior would be counterintuitive. Furthermore, according to the Dutch law, students who did not pass standardized exit exams cannot obtain a high school diploma. Therefore, we conclude that the monotonicity assumption is likely to hold.

6. Results

In this section, we present the estimates of the effect of a high school diploma on earnings. First, we report estimates of the effect of passing the initial standardized exit exams on diploma attainment and on earnings. Then, we report estimates of the effect of high school diploma on earnings. In addition, we also report heterogeneous effects based on gender and ethnicity. We end this section by showing that the results are similar when using local linear and local quadratic methods instead of global polynomial methods. All the models have been estimated with covariates included to improve precision of the estimates. Nonetheless, models without covariates yield very similar results.

6.1. The Effect of Passing Initial Standardized Exit Exams on High School Diploma and on Earnings

The effect of passing the initial standardized exit exams on diploma attainment (*first stage*) is reported in **Table 3** (Panel A). As expected from **Figure 1**, we find that students who passed standardized exit exams are much more likely to graduate from high school. This effect is robust to the polynomial order and ranges from 0.229 to 0.529 percentage points. Moreover, all the *F*-statistics are large, suggesting that our instrument is strong. Goodness-of-fit statistics indicate that the second model with linear interaction as polynomial order is the preferred specification.

TABLE 3 – THE EFFECT OF PASSING INITIAL STANDARDIZED EXIT EXAMS ON DIPLOMA ATTAINMENT AND EARNINGS

Panel A: Outcome: Diploma Attainment	(1)	(2)	(3)	(4)
Passed (1=yes, 0=no)	0.529***	0.363***	0.383***	0.229***
	(0.002)	(0.003)	(0.003)	(0.006)
Polynomial order	linear	linear interaction	quadratic	quadratic interaction
Number of students	435,768	435,768	435,768	435,768
F-statistic	258.584	279.339	275.550	284.334
Panel B: Outcome Log Net Earnings	(5)	(6)	(7)	(8)
Passed (1=yes, 0=no)	0.002	0.003	0.003	0.005
	(0.007)	(0.007)	(0.008)	(0.009)
Polynomial order	linear	linear interaction	quadratic	quadratic interaction
Number of students	435,768	435,768	435,768	435,768

Notes. Robust standard errors are in parentheses. Log net earnings are calculated per hour in the year immediately after students finished their last standardized exit exam. All specifications include an indicator for the track (1=pre-university education, 0=general secondary education) as well as additional covariates (gender, ethnicity, age at initial standardized exit exams, age at initial standardized exit exams squared, year of initial standardized exit exam, age of the father, age of the mother, neighborhood of residence, and parental income).
*** Significant at the 1 percent level.

In **Figure 3**, we show earnings by the initial mean standardized exit exam scores. There appears to be a positive association between earnings and initial standardized exit exam scores. This is unsurprising as the initial standardized exit exam scores are predictive of diploma attainment and we expect a positive association between a high school diploma and earnings. Nonetheless, there appears to be no discontinuity at the passing threshold.

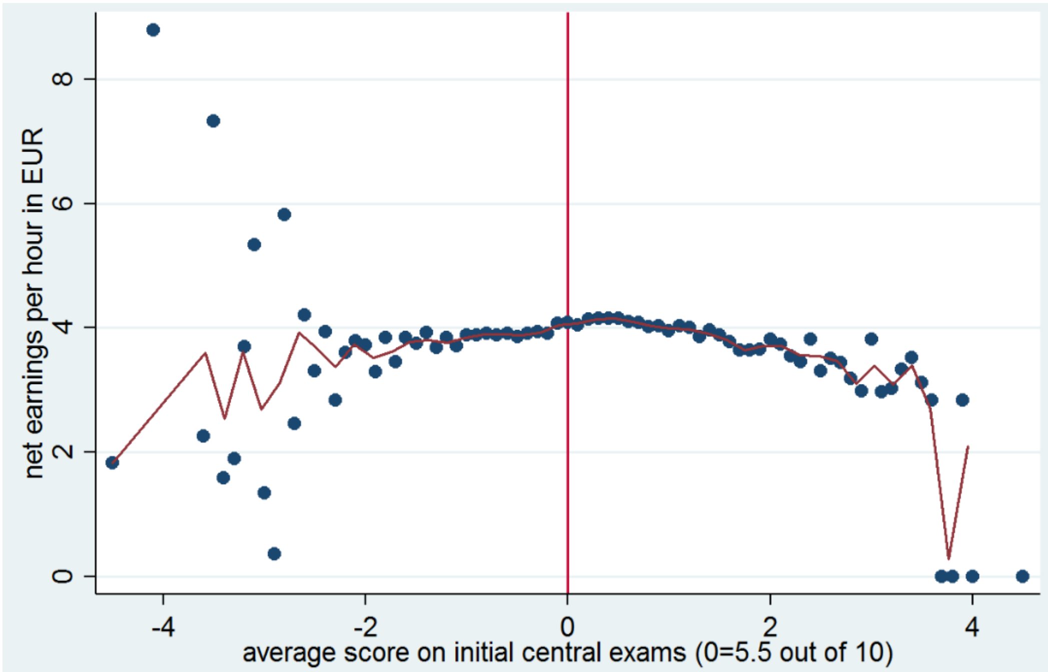

FIGURE 3 – EARNINGS BY INITIAL STANDARDIZED EXIT EXAM SCORES.
Notes. Log net earnings are calculated per hour in the year immediately after students finished their last standardized exit exam. Dots represent the mean log earnings by initial standardized exit exam scores. The full range of scores is depicted on the x-axis. This bin size was suggested to be the "optimal" bin size by the *rdplot* command in Stata 14 (Calonico et al. 2014).

The estimates in **Table 3** (Panel B) confirm this conclusion. We find that the effect of passing the initial standardized exit exams on log earnings is insignificant, regardless of the specification. These results are consistent with other similar studies (Clark and Martorell 2014; Jepsen et al. 2016) who also found no discontinuity effects at the passing threshold for earnings, but large effects in the first stage.

6.2. The Effect of a High School Diploma on Earnings

The instrumental variables estimates of the effect of a high school diploma on net earnings are presented in **Table 4**. It appears from Panel A that a high school diploma does not have a causal effect on net earnings immediately after leaving school. Regardless of the specification, the estimates are small and not significantly different from zero.

TABLE 4 – THE EFFECT OF A HIGH SCHOOL DIPLOMA ON EARNINGS

Panel A: Full sample	(1)	(2)	(3)	(4)
Diploma (1=yes, 0=no)	0.029 (0.053)	0.013 (0.010)	0.009 (0.009)	0.008 (0.021)
Polynomial order	linear	linear interaction	quadratic	quadratic interaction
Number of students	435,768	435,768	435,768	435,768
Panel B: Not in Post-Secondary Education	(5)	(6)	(7)	(8)
Diploma (1=yes, 0=no)	0.559*** (0.046)	0.340*** (0.049)	0.297*** (0.052)	0.223*** (0.085)
Polynomial order	Linear	linear interaction	quadratic	quadratic interaction
Number of students	118,446	118,446	118,446	118,446

Notes. Robust standard errors are in parentheses. Outcome in both panels is log net earnings per hour in the year immediately after students finished their last standardized exit exam. All specifications include an indicator for the track (1=pre-university education, 0=general secondary education) as well as additional covariates (gender, ethnicity, age at initial standardized exit exams, age at initial standardized exit exams squared, year of initial standardized exit exam, age of the father, age of the mother, neighborhood of residence, and parental income).

A possible concern, however, is that students who obtained a diploma are more likely to invest in skill acquisition and therefore more likely to enroll into post-secondary education. This is indeed the case in our sample as high school dropouts cannot, by law, enroll into higher education. We observe that about 73% of the students enrolled into post-secondary education immediately after either obtaining a diploma (students went to higher education) or after dropping out (students went to adult education). Therefore, consistent with Jepsen et al. (2016), we present the results with these students excluded in Panel B. We observe that the earnings effect of a diploma is likely to be positive. Among students who did not enroll into post-secondary education, students with a high school diploma are likely to earn a wage premium of 0.2 EUR to 0.5 EUR per hour immediately after leaving school. Assuming that a working day of 8 hours and 21 working days per month, this represents about 400 EUR to 1,000 EUR per year, a rather large effect.

6.3. Heterogeneous Effects

It is possible that the earnings effect of a high school diploma varies across different types of students. Therefore, in **Table 5**, we present the results by gender, ethnicity, and by track. We include the full sample in Panel A and we exclude students who enrolled into post-secondary education immediately after the last standardized exit exam in Panel B. To conserve space, we only present the results with standardized exit exam scores modelled as a linear interaction as this specification yielded best performance based on goodness-of-fit statistics. Nonetheless, as for the main effects, alternative specifications yield very similar results. For the full sample, our results suggest that a high school diploma does not affect earnings for any of the subgroups. By contrast, once we exclude students in post-secondary education, we observe a positive effect for all subgroups. We find no heterogenous effects by ethnicity. On the other hand, we find heterogenous effects by gender and by track. The effect appears to be higher for boys than for girls and higher for students in pre-university education than students in general secondary education.

TABLE 5 – THE EFFECT OF A HIGH SCHOOL DIPLOMA ON EARNINGS, BY GENDER, ETHNICITY, AND TRACK

	Boys	Girls	Foreign	Dutch	Pre-university education	General secondary education
Panel A: Full Sample						
Diploma (1=yes, 0=no)	0.008	-0.012	0.004	0.005	0.008	0.009***
	(0.010)	(0.017)	(0.018)	(0.014)	(0.013)	(0.023)
Number of students	52,885	65,561	22,599	95,847	46,673	71,773
Panel B: Not in Post-Secondary Education						
Diploma (1=yes, 0=no)	0.564***	0.176***	0.364***	0.326***	0.429***	0.290***
	(0.070)	(0.067)	(0.115)	(0.056)	(0.090)	(0.053)
Number of students	52,885	65,561	22,599	95,847	46,673	71,773

Notes. Robust standard errors are in parentheses. Outcome in both panels is log net earnings per hour in the year immediately after students finished their last standardized exit exam. Specifications by gender and by origin include an indicator for the track (1=pre-university education, 0=general secondary education). All specifications include additional covariates (gender, ethnicity, age at initial standardized exit exams, age at initial standardized exit exams squared, year of initial standardized exit exam, age of the father, age of the mother, neighborhood of residence, and parental income).
*** Significant at the 1% level

6.4. Alternative Specifications

Until now, we have used an approach based on global polynomial methods to estimate the causal effect of a high school diploma on earnings. Nonetheless, as suggested by Lee and Lemieux (2010), both global as local polynomial methods should be used when conducting an RD analysis. Therefore, in this section we perform a local linear and a local quadratic regression with a triangular kernel. In these specifications, the choice of bandwidth is crucial. Consistent with prior studies (Clark and Martorell 2014; Jepsen et al. 2016), we use the "optimal" bandwidth by Imbens and Kalyanaraman (2012), but also explore the robustness of our results to a variety of other bandwidths. Each specification includes the full set of covariates. Nonetheless, our results are robust to the exclusion of covariates. **Table 6** presents the results. In Panel A, we observe the results for the full sample and in Panel B, we observe the results for students who did not enroll in post-secondary education. Regardless of the specification, the results are very similar to the results using global polynomial methods, with no effect in the full sample and a positive effect of about 0.35 EUR to 0.43 EUR per hour.

TABLE 6 – THE EFFECT OF A HIGH SCHOOL DIPLOMA ON EARNINGS USING LOCAL POLYNOMIAL METHODS

	IK bandwidth (± 0.497)	± 0.25	± 0.75	± 1.00	± 1.25
Panel A: Full Sample					
Local linear regression					
Diploma (1=yes, 0=no)	0.001	0.001	0.004	0.009	0.010
	(0.011)	(0.012)	(0.014)	(0.025)	(0.031)
Local quadratic regression					
Diploma (1=yes, 0=no)	0.002	0.002	0.003	0.009	0.009
	(0.010)	(0.009)	(0.009)	(0.015)	(0.016)
Number of students	435,768	435,768	435,768	435,768	435,768
Panel B: Not in Post-Secondary Education					
Local linear regression					
Diploma (1=yes, 0=no)	0.351***	0.348***	0.370***	0.381***	0.399***
	(0.008)	(0.011)	(0.013)	(0.015)	(0.019)
Local quadratic regression					
Diploma (1=yes, 0=no)	0.355***	0.350***	0.385***	0.401***	0.429***
	(0.013)	(0.015)	(0.018)	(0.020)	(0.025)
Number of students	118,446	118,446	118,446	118,446	118,446

Notes. Robust standard errors are in parentheses. Specifications by gender and by origin include an indicator for the track (1=pre-university education, 0=general secondary education). All specifications include additional covariates (gender, ethnicity, age at initial standardized exit exams, age at initial standardized exit exams squared, year of initial standardized exit exam, age of the father, age of the mother, neighborhood of residence, and parental income).

7. Discussion

Using a fuzzy regression discontinuity design, we compared students who barely passed and barely failed standardized exit exams in the final year of secondary education. At first, we found no earnings effect of a high school diploma, both for the full sample as for different subgroups. However, we observed that most students who passed the exams actually enrolled in post-secondary education. Once we focused on students who immediately entered the labor market after secondary education, we found a positive effect of a high school diploma on earnings of about 0.34 EUR per hour. Although this effect remains positive regardless of the gender, ethnicity, or track, we found a larger positive effect for boys and students who have completed a program in pre-university education. Therefore, we conclude that a high school diploma is likely to have a positive effect on earnings. This finding is in line with previous correlational studies that estimated sheepskin effects (Belman and Heywood 1991; Hungerford and Solon 1987; Jaeger and Page 1996). However, it is in contrast with the only study that estimated causal sorting effects of a high school diploma (Clark and Martorell 2014). The difference lies likely in the different populations under study. Whereas Clark and Martorell (2014) focused on students who already failed exit exams at least once, we focused on all students in non-vocational education who did not enroll into post-secondary education, but immediately entered the labor market.

To interpret our findings as sorting effects instead of human capital effects, several conditions under which sorting effects can occur are worth mentioning. As noted by Clark and Martorell (2014), diplomas will have a sorting effect if three conditions apply. First, diplomas should contain information about relevant productivity differences in a competitive labor market. We found a positive association between having a high school diploma and earnings of about 25%. Therefore, employers are likely to use this information when screening the workers.

Second, employers should observe diplomas and be able to verify them if necessary. This is especially likely in the Netherlands where diploma attainment can easily be verified through an online diploma register maintained by the Dutch Ministry of Education, Culture and Science (Government of the Netherlands, 2018a). This register includes all the obtained diplomas and can only be accessed by the student who obtained the diploma. Thus, employers just have to ask workers to contact the register and provide a confirmation that they have indeed obtained a diploma. Lastly, firms should not obtain the information about productivity differences from other sources. In the Netherlands, students do not obtain a certificate if they fail standardized exit exams. Although employers can ask workers for their individual grades or test workers themselves, this is unlikely to happen in practice due to cost concerns. Most companies request the high school diploma status and the secondary education track. In sum, we interpret the positive earnings effect of a high school diploma as a diploma sorting effect.

Although our research design could address most of the issues in the literature, this study is not without limitations. First, we did not include students in vocational education as they take standardized exit exams several years before graduation. Given that vocational students typically stem from a disadvantaged socioeconomic background and are especially likely to drop out of high school, it is unclear whether our results also hold for these students. Second, we could only study short-term effects of a high school diploma. It would be interesting to study whether sorting effects disappear after several years as employers learn more about the workers. Finally, as any study using a fuzzy regression discontinuity design, our results solely apply for the population of compliers at the threshold. Addressing these limitations can provide new potential avenues for further research.

REFERENCES

Albrecht, James W. 1981. "A Procedure for Testing the Signalling Hypothesis." *Journal of Public Economics* 15: 123-132.

Altonji, Joseph G., and Charles R. Pierret. 2001. "Employer Learning and Statistical Discrimination." *Quarterly Journal of Economics* 116 (1): 313-350.

Arabsheibani, Gholamreza. 1989. "The Wiles Test Revisited." *Economics Letters* 29: 361-364.

Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo. 2010. "Beyond Signaling and Human Capital: Education and the Revelation of Ability." *American Economic Journal: Applied Economics* 2 (4): 76-104.

Arrow, Kenneth J. 1973. "Higher Education as a Filter." *Journal of Public Economics* 2: 193-216.

Becker, Gary. 1962. “Investment in Human Capital: A Theoretical Analysis.” *Journal of Political Economy* 70 (5): 9-49.

Bedard, Kelly. 2001. "Human Capital versus Signaling Models: University Access and High School Dropouts." *Journal of Political Economy* 109 (4): 749-775.

Belman, Dale, and John S. Heywood. 1991. "Sheepskin Effects in the Returns to Education: An Examination of Women and Minorities." *Review of Economics and Statistics* 73 (4): 720-724.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust data-driven inference in the regression-discontinuity design." *Stata Journal* 14 (4): 909-946.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2018. "Manipulation testing based on density discontinuity." *Stata Journal* 18 (1): 234-261.

Chevalier, Arnaud, Colm P. Harmon, Ian Walker, and Yu Zhu. 2004. "Does Education Raise Productivity, Or Just Reflect It?" *Economic Journal* 114 (499): F499-F517.

Clark, Damon, and Paco Martorell. 2014. "The Signaling Value of a High School Diploma." *Journal of Political Economy* 122 (2): 282-318.

Di Pietro, Giorgio. 2017. "Degree classification and recent graduates' ability: Is there any signalling effect?" *Journal of Education and Work* 30 (5): 501-514.

Farber, Henry S., and Robert Gibbons. 1996. "Learning and Wage Dynamics." *Quarterly Journal of Economics* 111 (4): 1007-1047.

Feng, Andy, and Georg Graetz. 2017. "A question of degree: The effects of degree class on labor market outcomes." *Economics of Education Review* 61: 140-161.

Flores-Lagunes, Alfonso, and Audrey Light. 2010. "Interpreting Degree Effects in the Returns to Education." *Jounal of Human Resources* 45 (2): 439-467.

Gelman, Andrew, and Guido Imbens. 2018. "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs." *Journal of Business & Economic Statistics* 1-11.

Government of the Netherlands. 2018a. *Diplomaregister werkgevers [Diploma register employers]*. December 20. Accessed December 20, 2018. https://duo.nl/zakelijk/diploma/diplomaregister-werkgevers.jsp.

—. 2018b. *Eindexamenbesluit VO [Final exam decision secondary education]*. March 15. Accessed December 12, 2018. https://wetten.overheid.nl/BWBR0004593/2018-03-15/0.

—. 2018c. *Leerplicht en kwalificatieplicht [Compulsory schooling and qualification requirement]*. December 12. Accessed December 12, 2018. https://www.rijksoverheid.nl/onderwerpen/leerplicht/leerplicht-en-kwalificatieplicht.

—. 2018d. *Maatregelen tegen voortijdig schoolverlaten [Measures against school dropout]*. 12 December. Accessed December 12, 2018. https://www.rijksoverheid.nl/onderwerpen/vsv/minder-voortijdig-schoolverlaters.

Groot, Wim, and Hessel Oosterbeek. 1994. "Earnings Effects of Different Components of Schooling: Human Capital Versus Screening." *Review of Economics and Statistics* 76 (2): 317-321.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69 (1): 201-209.

Hungerford, Thomas, and Gary Solon. 1987. "Sheepskin Effects in the Returns to Education." *Review of Economics and Statistics* 69 (1): 175-177.

Imbens, Guido, and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79: 933-959.

Jaeger, David A., and Marianne E. Page. 1996. "Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education." *Review of Economics and Statistics* 78 (4): 733-740.

Jepsen, Christopher, Peter Mueser, and Kenneth Troske. 2016. "Labor Market Returns to the GED Using Regression Discontinuity Analysis." *Journal of Political Economy* 124 (3): 621-649.

Kahn, Lisa B., and Fabian Lange. 2014. "Employer Learning, Productivity, and the Earnings Distribution: Evidence from Performance Measures." *Review of Economic Studies* 81 (4): 1575-1613.

Khoo, Pauline, and Ben Ost. 2018. "The effect of graduating with honors on earnings." *Labour Economics* 55: 149-162.

Kleibergen, Frank, and Richard Paap. 2006. "Generalized Reduced Rank Tests Using the Singular Value Decomposition." *Journal of Econometrics* 133 (1): 97-126.

Kroch, Eugene A., and Kriss Sjoblom. 1994. "Schooling as Human Capital or a Signal: Some Evidence." *Journal of Human Resources* 29 (1): 156-180.

Lambropoulos, Haris S. 1992. "Further Evidence on the Weak and Strong Versions of the Screening Hypothesis in Greece." *Economics of Education Review* 11 (1): 61-65.

Lang, Kevin, and David Kropp. 1986. "Human Capital Versus Sorting: The Effects of Compulsory Attendance Laws." *Quarterly Journal of Economics* 101 (3): 609-624.

Lange, Fabian. 2007. "The Speed of Employer Learning." *Journal of Labor Economics* 25 (1): 1-35.

Layard, Richard, and George Psacharopoulos. 1974. "The Screening Hypothesis and the Returns to Education." *Journal of Political Economy* 82 (5): 958-998.

Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48: 281-355.

Lee, Kiong-Hock. 1980. "Screening, Ability, and the Productivity of Education in Malaysia." *Economics Letters* 5: 189-193.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics* 142: 698-714.

Miller, Paul W., and Paul A. Volker. 1984. "The Screening Hypothesis: An Application of the Wiles Test." *Economic Inquiry* 22 (1): 121-127.

Montiel Olea, José Luis, and Carolin Pflueger. 2013. "A Robust Test for Weak Instruments." *Journal of Business & Economic Statistics* 31 (3): 358-369.

Psacharopoulos, George. 1979. "On the Weak Versus the Strong Version of the Screening Hypothesis." *Economics Letters* 4: 181-185.

Psacharopoulos, George, and Harry Anthony Patrinos. 2018. "Returns to investment in education: a decennial review of the global literature." *Education Economics* 26 (5): 445-458.

Riley, John G. 1979. "Testing the Educational Screening Hypothesis." *Journal of Political Economy* 87 (5 (Part 2)): S227-S252.

Rosenbaum, Paul R. 1984. "The Consequences of Adjustment for a Concomitant Variable That Has Been Affected by the Treatment." *Journal of the Royal Statistical Society (Series A)* 147 (5): 656-666.

Schultz, Theodore W. 1961. "Investment in Human Capital." *American Economic Review* 51 (1): 1-17.

Spence, Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics* 87 (3): 355-374.

Taubman, Paul J., and Terence J. Wales. 1973. "Higher Education, Mental Ability, and Screening." *Journal of Political Economy* 81 (1): 28-55.

Tucker, Irvin B. 21. "Evidence on the Weak and the Strong Versions of the Screening Hypothesis in the United States." *Economics Letters* 21: 391-394.

Tyler, John H., Richard J. Murnane, and John B. Willett. 2000. "Estimating the Labor Market Signaling Value of the GED." *Quarterly Journal of Economics* 115 (2): 431-468.

Weiss, Andrew. 9. "Human Capital vs. Signalling Explanations of Wages." *Journal of Economic Perspectives* 9 (4): 133-154.

Wiles, Peter. 1974. "The Correlation Between Education and Earnings: The External-Test-Not-Content Hypothesis (ETNC)." *Higher Education* 3 (1): 43-57.

Wolpin, Kenneth I. 1977. "Education and Screening." *American Economic Review* 67 (5): 949-958.