PALKANSAAJIEN TUTKIMUSLAITOS
LÖNTAGARNAS FORSKNINGSINSTITUT
LABOUR INSTITUTE FOR ECONOMIC RESEARCH

TYÖPAPEREITA | WORKING PAPERS 329

The long shadow of high stakes exams: Evidence from discontinuities*

Artturi Björk[1] & Hannu Karhunen[2,3,4]

* We would like to acknowledge the financial support from The Finnish Cultural Foundation, the Strategic Research Council at the Academy of Finland (grant numbers 303686 and 303689) and comments from Marko Terviö, Niku Määttänen, Tuomo Suhonen, Roope Uusitalo, Tuomas Pekkarinen, Julia Salmi, Tuomas Laiho and Andrey Zhukov and participants at the annual meeting of The Finnish Economic Association, the annual meeting of The European Association of Labour Economists, Helsinki Graduate School of Economics lunch seminar and Labour Institute for Economic Research seminar. We would also like to thank the Jyväskylä University School of Business and Economics, The Matriculation Examination Board and Statistics Finland for proving the microdata for this study.

[1] Aalto yliopisto
[2] Labour Institute for Economic Research
[3] Statistics Finland
[4] Jyväskylä University School of Business and Economics

Abstract

We estimate the effect of receiving a higher grade on a high school exit exam on labor market and education outcomes. Identification comes from comparing students on di˙erent sides of grade cutoffs. Being above a cutoff in an exam leads to (i) an increase in quality of education, but no change in years of schooling, (ii) an increase in yearly earnings that peaks between 1 and 5% at age 48, but no change in employment. At most 60% of the increase in earnings is explained by better education opportunities.

Keywords: High school exit exam, Regression discontinuity
JEL-classification: I21, I26, J24

Tiivistelmä

Tässä tutkimuksessa tarkastellaan englannin ja ruotsin kielen ylioppilaskirjoitusten arvosanojen vaikutusta koulutus- ja työmarkkinatulemiin. Tutkimuksessa hyödynnetään arvosanarajojen yhteyteen syntyvää pisteiden epäjatkuvuuskohtaa, jonka ansiosta voimme verrata juuri korkeamman arvosanan saaneita henkilöitä niihin henkilöihin, jotka jäävät niukasti arvosanarajan alle. Tulosten mukaan korkeampi arvosana johtaa korkeampaan koulutukseen muttei vaikuta kouluvuosien määrään. Korkeamman arvosanan havaitaan myös vaikuttavan positiivisesti henkilön ansioihin aina jopa 50 ikävuoteen asti, mutta arvosanalle ei löydetä työllisyysvaikutusta. Arviomme mukaan voimme selittää enintään 60 prosenttia tulovaikutuksesta koulutuksella. Tulokset myös tukevat johtopäätöstä, että ylioppilaskirjoitusten tuloksia tulisi hyödyntää nykyisiä arvosanoja hienojakoisemmalla jaottelulla.

Avainsanat: Ylioppilaskirjoitukset, regressioepäjatkuvuusmenetelmä

Palkansaajien tutkimuslaitos
Labour Institute for Economic Research
Pitkänsillanranta 3 A
00530 Helsinki

Työpapereita | Working Papers 329
ISBN 978-952-209-185-7 (pdf)
ISSN 1795-1801 (pdf)

www.labour.fi

Helsinki 2019

1 Introduction

Many countries have an educational system where access to higher education is rationed based on performance in upper-secondary education. In the United States nationwide standardized tests SAT and ACT are used in determining access joint with high school grade point average (GPA) and other factors. The benefit of using tests in determining access to higher education is that they are more objective than other measures such as the GPA or an admission letter and typically the results of such tests are difficult to manipulate. The downside to tests is that the results might have more random variation. Testing for attributes that cause success in or high returns from higher education can be difficult.

We study the exit exams of Finnish secondary education track that prepares students for post-secondary studies (from now on *high school*) to learn about the effect of success in a high stakes test on labor market outcomes through the life-cycle. The exams are scored on a relatively fine scale with integers from zero to 300. Information about the performance in the exams is released as six possible grades for each exam. The grades are determined by cutoffs that are set after the scoring. During our sample period Finnish higher education programs grant access based on their own entry exams (60% weight on average) and the national high school exit exams (40%). We use a regression discontinuity design and compare education and labor market outcomes of similarly skilled students just below and above grade cutoffs. This allows us to identify the effect of the grade independent of the skill of the student.

We study the population of Finnish high school exit exam takers from 1982 to 2014 (n $\sim$ 1 million). We use administrative data on the scores and grades from English as a second language (starting at age 9) and Swedish as a third language (starting at age 12) exams. Most exit exam takers take both exams which brings the total number of exams we study to $\sim$ 2 million. Some students repeat the exams and students from the Swedish speaking minority (6% of the population) typically take only the English exam. We include the repeaters and the Swedish speaking minority for a total of $\sim$ 2.2 million exams. The data on education, employment and earnings outcomes come from registries.

Our main finding is that the effect of crossing a grade cutoff on earnings is increasing in age. The impact of being above a cutoff on yearly earnings is between €-80 and €0 or between -1 and 0% of the mean at age 20. Ten years after the exam the effect is between €-150 and €150 (-0.5 and 0.5%). The effect reaches its maximum at age 48 between €600 and €1900 (1.5 and 4.5%). This effect is robust to the choice of bandwidth, functional form and the cohorts we include in the analysis.

Since the exit exam grades are used in determining access to post-secondary education, the most obvious mechanism how better high school grades cause higher earnings is through education. However there is hardly a change in the quantity of education at the cutoff: scoring above it increases schooling by between 4 and 18 days. We do find support for an increase in the selectivity of education at the cutoff. We measure the selectivity of an post-secondary education program (subject-institution pair) by the exit exam results of its students, which we anchor to earnings. Crossing a cutoff increases the mean of the anchored exit exam results in students post-secondary education program by 0.1%. Median earnings at a student's post-secondary program is due to both the earnings potential of the admitted students and the value of the education. Because students are positively selected to education, the effect of the cutoff on the median earnings of the post-secondary program gives an upper bound on the effect of the cutoff on earnings that works through education. The median earnings of a student's education program jump at the cutoff. The effect is increasing in age similarly to the effect on student's own earnings. The effect on program's median earnings is 60% of the effect on student's own earnings.

Previous studies (Ebenstein, Lavy, and Roth 2016; Canaan and Mouganie 2018; Tyler, Murnane, and Willett 2000) have shown that success in a test at the end of upper secondary education has large positive effects on earnings in the first ten year period after the test, but we don't know about longer term outcomes. Studying longer term outcomes in conjunction with short term outcomes can help us differentiate between proposed ways in which success at the test can affect later life outcomes. The hypothesis that a test produces information about productivity that is rewarded in the labor market is consistent with larger positive effects from success in the test soon after the test, which diminish with age as the true productivity of the student is learned. Small or even negative effects soon after the test and increasing effects later would be consistent with the hypothesis that success in the test increases access to human capital building education. Our contribution is to study labor market outcomes up to thirty years after the test and to bound the share of the effect explained by the increase in education opportunities afforded by higher grades.

Tyler, Murnane, and Willett (2000) consider the GED in the US, which is an alternative to a high school certification for drop outs. They use between state variation in passing standards to identify the effect of being GED certified, controlling for selection into the GED. They find that wages for whites are 10 to 20% higher 5 years after the test, but they find no effect for minority students.

Our paper is closely related to two recent papers on the effect of high school exit exams on labor market outcomes that use random variation in the results of the exam to disentangle the effect of the success in the exam from the underlying ability.

Ebenstein, Lavy, and Roth (2016) study the effect of success in high school exit exams in Israel using random variation in exam performance coming from sandstorms. They find large positive effects from success in the exit exam on education and earnings ten years after the exam. This effect is increasing in skill, which is consistent with their posited mechanism of success in the exit exam granting access to competitive education. They find that a one standard deviation increase in the score of the exit exam causes a 13% increase in earnings at age 30.

Canaan and Mouganie (2018) study the effect of marginally passing the French high school exit exams, which rests on the same assumption that we make i.e. that unobserved heterogeneity is continuous across the cutoff and hence controlled by the running variable. Contrary to us, they have manipulation across the cutoff and hence cannot use the data close to the cutoffs that would be most informative without manipulation. They look at short run effects of the pass/fail cutoff and find that crossing the cutoff increases the quality of post secondary education and raises self reported monthly earnings by 12% 10 years after the exam. They attribute the increase in wages to an increase in the quality of post-secondary education.

2 Institutional Setup

2.1 Upper Secondary Education in Finland

The upper secondary education in Finland is divided to academic (*lukio*) and vocational (*ammattikoulu*) tracks. This paper concerns the academic track of upper secondary education, which we call *high school* from now on to save space. In recent years roughly half of the cohort has attended high school and half the vocational track. The vocational track prepares students for professions such as hair dresser or builder and most students do not continue their studies further.

High school prepares students for studies in higher education. The duration of high school was fixed to three years until 1994. Since then, students have been able to choose to finish it in two to four years. 80% of students who started in autumn of 2012 had finished by the end of 2015 (Loukkola 2017). There is a set of core courses that are compulsory to all students and in addition students get to choose from optional courses. The subjects include natural and social sciences, mathematics, arts and languages. One course includes 38 hours of instruction in addition to self study. Graduating high school requires passing exit exams (*ylioppilaskirjoitukset*) in addition to passing the core courses and 75 courses in total. At high school graduation students receive a Matriculation Examination Certificate with a grade for each of the passed exit exams and a high school diploma, which has an integer grade point average and the number of courses for each of the subjects the student studied. Section 2.1 shows that the number of high school graduates and their share of the cohort has more than doubled from birth year 1950 to 1990.

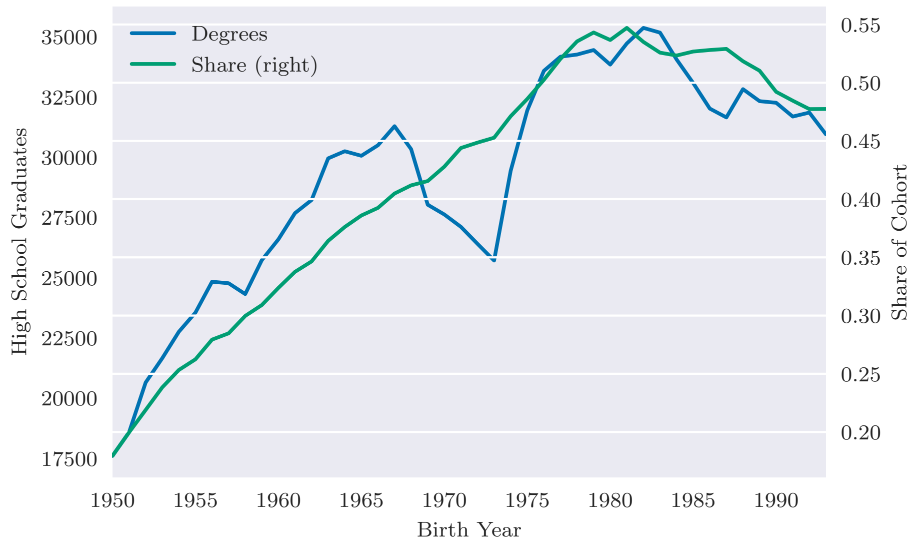

Figure 1: High school graduates by year of birth

2.2 Exit Exams

High school exit exams are held twice a year. The exams in March are ten times more popular than the ones in September. The popularity of the March exams is explained by the timing of application to higher education.

Each subject has its own exam. The available subjects are native language (either Finnish or Swedish), science (the students choose which questions to answer from any of the humanities and natural sciences that are thought in high school), basic and advanced mathematics and foreign languages on two proficiency levels, A and B. In principle level A exam is for the first foreign language started at age 9 and level B for the second foreign language started at age 12. Until 2005 the required exit exams were either science or basic or advanced mathematics, native language, two foreign languages in total, at least one at A-level and including the other official language (Swedish for native Finnish speakers and vice versa). From 2005 the compulsory exams have been native language and three out of science, mathematics and languages.

The exams are prepared by the Matriculation Exam Board. Each subject is a separate exam and it's administered on the same day across the country. The exams are scored centrally by the Matriculation Exam Board's assistants who are usually either high school teachers or post-secondary educators of the subject of the exam. The answer sheets that the assistants receive contain the students name, her school, her answers and a preliminary score by the

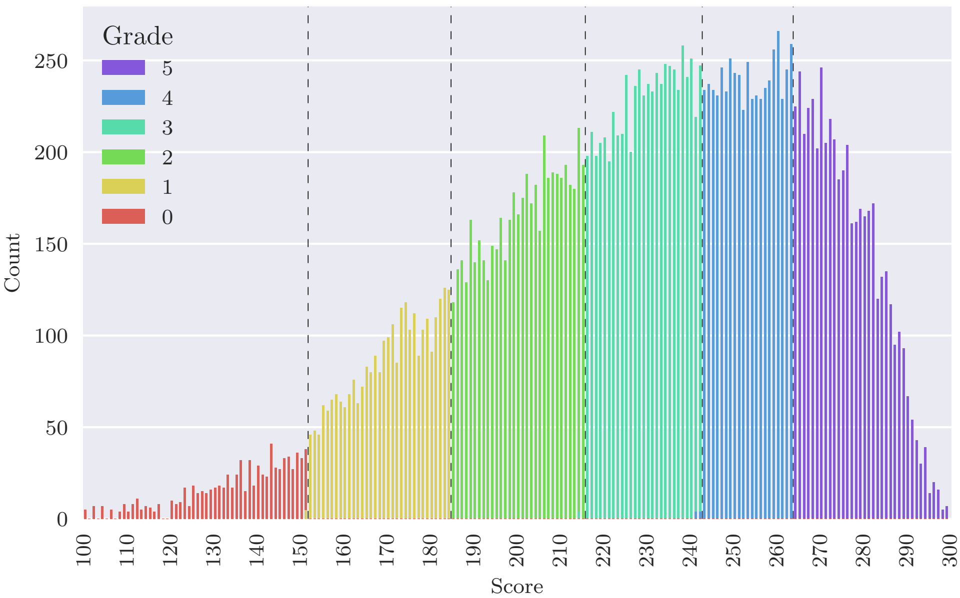

Figure 2: Scores and grades for the English exam in the spring of year 1990
Notes: The figure illustrates the distribution of scores and grades. The dashed vertical lines are the grade cutoffs.

students own teacher. The assistants are not allowed to score students whom they know, for example students from the same high school, relatives etc. The scores range between 0-60 in mathematics and science and between 0-300 in foreign languages. After the scores are set the Matriculation Exam Board sets the grade cutoffs. Before 1996 six grades were used. Since 1996 the top grade was split to two grades so that the total number of grades increased to seven.[1] Fig. 2 illustrates how the distribution of scores is discretized to grades in the English exam in an arbitrary year (1990).

In principle the scores are mapped to grades "on a curve": the best grade is awarded to the top 5%, the next to the following 15%, 20% 24%, 20%, 11% and 5% fail the exam. The share of grades is adjusted from this baseline to make the grades comparable over time and between subjects. For example the September exams have a lower grade distribution than March exams because they have relatively more repeaters.

After a student has received her grades she can apply for a rescoring of the exam. The rescoring can only increase the students score. There is a rescoring fee of €50, which is waived in the case the rescoring results in a higher score. Our data on the scores comes from after the rescoring so in principle there could be selection across the grade cutoffs. In 2015 there were only 470 rescoring application, which is 0.2% of all exams. In 2018 rescoring

[1]Known in Finland by their Latin abbreviations from highest to lowest L, E, M, C, B, A and I

resulted in a change in the score one times out of four. Based on this we argue that selection is not a problem in practice.

After passing the required exams the student receives a Matriculation Examination Certificate. The certificate lists the passed exams together with their grades.

2.3 Post-secondary education in Finland

Access to post-secondary education is based partly on the grades from high school exit exams. Our prior is that this is the main channel through which success in the exam affects future labor market outcomes. Next we describe how entry to post-secondary education works in Finland.

Higher education in Finland is publicly provided. It is divided to universities of applied science (*Ammattikorkeakoulu* in Finnish and from now on applied college) and universities. Applied colleges mostly grant bachelor's degrees in professions such as software engineering and nursing. Students enter from high schools or the vocational track of upper secondary education. Universities enter students only from high schools and grant bachelor's and master's degrees and doctorates. Most students exit with a master's degree.

Application to higher education is not centrally organized. Instead, programs (i.e. disciplines within a specific institution, say University of Helsinki Department of Law) choose who to admit. Most programs admit students in June and July each year.

Almost all programs hold entry exams. The material that is tested in the exam is typically unique to the program. Usually programs admit half of their students based jointly on the high school exit exam and the entry exam and half based only on the entry exam. Some of the less popular programs admit students based only on the high school exit exams. (Moisio and Vuorinen 2016) have estimated that the weight of the high school exit exam is 40% on average. There are private prep courses for the most competitive entry exams. 20% of the applicants to universities (9% for applied colleges) attend a preparatory course. (Moisio and Vuorinen 2016). In 2018 entry to post-secondary education was reformed to give much more weight to the high school exit exam grades. Our data comes from the time before the reform.

2.3.1 Example: Entry to law school at University of Helsinki

For a concrete example on how entry into higher education works, we'll detail the case for the law school at University of Helsinki. Applications have to be submitted in a two week period in mid March, when the high school exit exams are held as well. If the students wait until the last day to apply, they have taken all the exit exams, but do not have the initial results from them. The required reading for the entry exam is published two weeks before the application period starts. The applicants have two and a half months to study the reading before the entry exam. Half of all law school applicants attend a prep course (Moisio and Vuorinen 2016), which is probably a lower bound on the share for the law school at University of Helsinki, since entry to it is the most competitive.

The entry exam is scored from 0 to 50 and the best four exit exams award between 0 and 8 points for a maximum of 32 points from exit exams.[2] The higher English exam awards between 2 to 8 points for grades 2 to 7 and the lower Swedish exam awards between 1 to 5 points for grades 3 to 7. First hundred slots are awarded based on the joint score from the entry and exit exams. In 2016 the joint entry cutoff was 60 points. Last hundred slots are awarded based on the entry exam alone with a cutoff of 41 in 2016. So in 2016 if you applied to the law school at University of Helsinki whether you passed a grade cutoff in the high school exit exams affected your entry if your joint score was 60 (so you were the marginal applicant for the first 100 slots) and you scored less than 41 points in the entry exam (so you wouldn't have gotten in in the last 100 slots).

3 Data and Sample Selection

3.1 Data on exit exams

This study is built on administrative data set of the population of high school students who sign up for high school exit exams in Finland. The summary statistics of the exit exam data are detailed in Table 1. The examinations registry is administrated by the Matriculation Examination Board (MEB) and an electric registry contains all candidates starting from the March 1967 period. Due the old erroneous saving procedure, the exam information is missing from the 1969 March exams, foreign language exams prior to 1981 and the 1990 exams for mathematics and native Finnish. The MEB registry has information on all exams a students takes (including repeats), when and where the exams were administered, the scores and the final grades from the exams.

We concentrate on the foreign language exams because they have approximately 200 unique values for the score (compared to 60 for the other exams), which makes it possible to control for the running variable and specifically English and Swedish because the sample sizes are orders of magnitude larger than for other foreign language exams. The other exams with large sample sizes: Finnish, mathematics and humanities & science are limited by not having as many unique scores. Furthermore, the humanities & science exam was split to individual subject exams and the Finnish exam changed from

[2] Higher level math, Finnish, Science and higher level foreign languages award 8,6,5,4,3,2 points for the grades 7-2 and lower level math and languages award 5,4,3,2,1 for the grades 7-3.

Table 1: Summary Statistics

Grade	Fail	1	2	3	4	5	6	Pooled
Students (1000's)								1050
English (1000's)	48	139	247	284	247	207	38	1210
Swedish (1000's)	50	119	196	221	182	153	24	945
Mean Score	116	153	185	215	241	265	278	212
Std Score	(22)	(18)	(18)	(16)	(14)	(13)	(8)	(45)

Notes: The first three rows report the sample sizes for students, English exams and Swedish exams respectively. The final two rows show the mean and the standard deviation of the scores. The columns are the seven possible grades that a student can receive from the exam and the final column reports the summary statistics when the data is pooled over the grades.

two essay writing exams to having both reading comprehension and writing with a new scoring system during our sample. The unique institutional details of the Finnish exam make manipulation of the score possible.

3.2 Outcome variables

Our outcome variables are from different full-population registry data sets administrated by the Statistics Finland. We link variables to the exam data from the Register of Completed Education and Degrees from Finnish institutions (field, institution, level and date of the degree) and from the Longitudinal Employment Statistics Files from 1987 to 2015 (unemployment, employment, annual labor earnings, business income)

We want our earnings measure to capture the earnings that are generated from labor. Because of this we sum annual labor earnings and business income together. The labor earnings are earnings that are paid to salaried and hourly workers. Business income is income that small business owners (such as a plumber working as an independent contractor) earn on their labor. Because of tax incentives some labor earnings get shifted to capital income. We do not include capital income in our earnings variable because we don't know how much of capital income is actually labor earnings and because including capital income in our earnings measure makes it very noisy. We report means of outcome variables together with the results.

4 Identification

4.1 Identification Strategy

Our research question is what is the effect of getting a better grade in the high school exit exam on education and labor market outcomes? Our identification strategy is regression discontinuity design: we check whether the grade cutoffs are associated with discontinuous jumps in the outcome variable. If the unobserved characteristics change smoothly over the cutoff, they are captured by the running variable and a jump in the outcome variable at the cutoff is due to the cutoff.

We estimate a regression pooled over subjects $s \in \{\text{English, Swedish}\}$, cutoffs $c \in \{5, 16, 36, 60, 80, 95\text{th pctl}\}$ and dates $t \in \{\text{Mar 1982}, \text{Sep 1982}, \text{Mar 1983}, ..., \text{Sep 2013}\}$:

$$\tilde{y}_{s,c,t,a,i} = \beta_a D_i + f(distance_{s,c,i})\sigma^{below}_{s,c,a} + D_i f(distance_{s,c,i})\sigma^{above}_{s,c,a} + \epsilon_{a,i} \quad (1)$$

Where $\tilde{y}_{s,c,t,a,i}$ is the difference of the outcome of interest at age a with the mean of the outcome for the subject, cutoff, date and age bin $\tilde{y}_{s,c,t,a,i} = y_{s,c,t,a,i} - \bar{y}_{s,c,t,a,i}$, D is a dummy which takes the value 1 if the observation is above the cutoff and 0 otherwise and f is a function of the distance to the closest cutoff estimated separately for both sides of the cutoff and separately for each subject and cutoff pair, but pooling over exam years. σ's are the parameters for the function f estimated separately for each side of the cutoffs.[3] β_a is the effect of crossing the cutoff from a lower grade to a higher grade, which we identify from within subject, cutoff and year variation in outcomes across the cutoffs that is not explained by the function of the distance to the cutoff. We use a triangular kernel to give the observations closest to the cutoffs most weight. Because we have two observations for most individuals we cluster the standard errors at the student level. We do not cluster standard errors by the discrete values of the running variable (score) because Kolesár and Rothe (2018) shows with simulations that clustering is unreliable when there are only few unique values of the running variable on both sides of the cutoff.

Because our running variable (the distance to closest grade cutoff) takes only 20 unique values we are unable to use nonparametric rdd à la Cattaneo, Idrobo, and Titiunik (2017). We have to specify a functional form for f and choose a bandwidth. Our preferred f is linear, because the estimates are more precise, but we consider also a quadratic function in the robustness analysis in Appendix E, but no higher order polynomials because of arguments by Gelman and Imbens (2017). We evaluate the fit of f by running the same regression for dummy cutoffs. Dummy cutoffs are the mid-points between actual grade cutoffs. For example in the Swedish exam in year 1990 the second grade was awarded to the scores ranging from 152 to 183 so we place the placebo cutoff at 167.5. If we manage to control for the effect of the running variable with our f function, the parameter estimates for the effect of the placebo cutoffs should be zero. If they are non-zero this is evidence for the functional form assumption affecting our results. We report the results of the placebo regressions with our main results.

[3] As a robustness check we estimate the polynomials separately for each exam period. Our results are not affected.

Similarly to the functional form, we have to also specify the estimation bandwidth. Because of this we report our results with varying bandwidths in Appendix D. Our preferred bandwidth is ten, which is the widest possible bandwidth without any observations having to be left out because they are closer to other cutoffs. Our parameter estimates are stable around our preferred bandwidth and the estimates on the placebo cutoffs are not significantly different from zero in general.

4.2 Validity of the identification strategy

Our running variable of the exam score is ordinal data (we know that a score of 11 is more than a score of 10, but not that it is 10% greater), but we need to treat it as interval data to get identification. Since we don't have data arbitrarily close to the cutoffs we have to use data away from the cutoff and extrapolate from it to the cutoff. We do this by estimating a polynomial for the distance to the cutoff separately on both sides of the cutoff. Extrapolating to the cutoff then requires that the distance between points away from the cutoff is the same as the distance between points across the cutoff. This assumption would be violated for example in the case where the score could take any other value except the score just below the cutoff. Then the jump in the score to the adjacent one would be two points when jumping over the cutoff and one point otherwise.

Fig. 3 plots the distribution of the score around each cutoff pooled over subjects and years. We can see from the plots that the middle cutoffs have the most mass. The smoothness of the density of the score across cutoffs suggests that the distance between any adjacent scores is comparable over the cutoff and away from it.

We have a strong prior that manipulation of the score that could make unobservable covariates jump at the cutoff is unlikely because the exams are scored before the grade cutoffs are set and the grade cutoffs vary from year to year as we can see from Fig. 4, which plots the grade bands for English.

We check the smoothness of the distributions in Fig. 3 statistically with the test of Frandsen (2017) (with k=0), which essentially checks that the number of observations is consistent with the binomial distribution with success probability of one third and number of trials the sum of the number of observations just above the cutoff and adjacent points (at distances -0.5,0.5 and 1.5 to the closest cutoff).

The results for the test are in Table 2. It confirms our visual inspection of the histograms that there is no manipulation of the score. We cannot reject the null hypothesis that the distributions are smooth at the 5% level over the individual cutoffs and when we pool all cutoffs together, except for the cutoff at the 80th percentile. We do not have an explanation as to why the distribution at the 80th percentile might not be continuous except by chance. The distributions are smooth over placebo cutoffs as indicated by

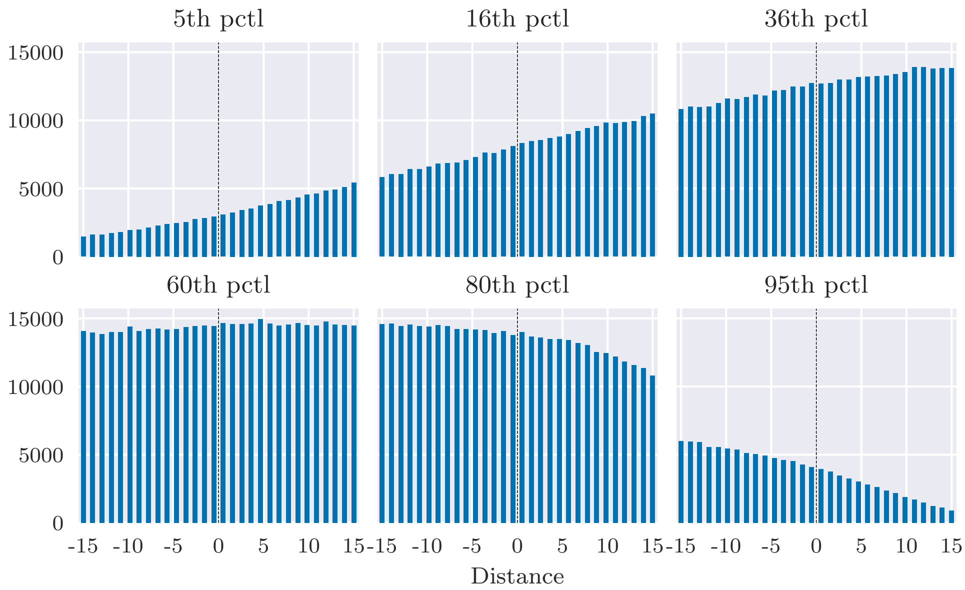

Figure 3: Distribution of the score

Notes: The figure plots the histogram of the distribution of our running variable, the score. To construct the figure we classify each observation (a result from an exam) to its closest grade cutoff and normalize the score with the value of the grade cutoff. For example in the English exam in year 1990 the first cutoff was at 151.5 and the second at 184.5. Hence if a student would score 160 we would classify her as being closest to cutoff 1 and her normalized distance would be 8.5. A student with the score 180 would be classified as being closest to cutoff 2 with a normalized score of -4.5. We pool the normalized scores over years and subjects to produce the histogram for each cutoff.

Table 2: No manipulation test

Cutoff pctl	5	16	36	60	80	95	Pooled
P-value	0.90	0.91	0.80	0.26	0.05	0.94	0.15
	(0.91)	(0.64)	(0.60)	(0.91)	(0.45)	(0.15)	(0.30)

Notes: Results of the test Frandsen (2017) (with k=0) with the null hypothesis of no manipulation of the normalized scores across the cutoffs (and placebos in parenthesis). The first columns report the P-values across the individual cutoffs pooled over years and subjects and the final column reports the P-value when we pool also across the cutoffs.

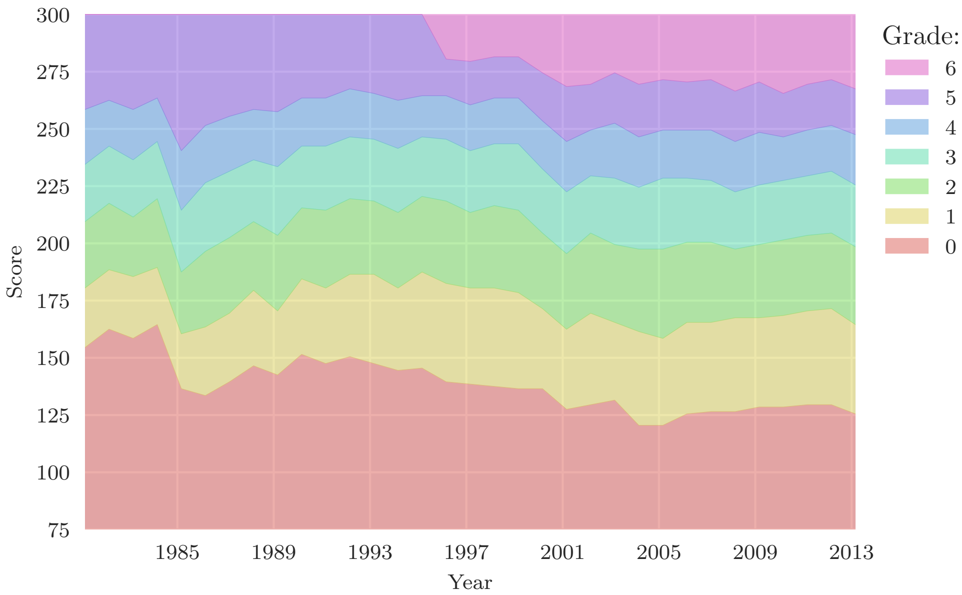

Figure 4: English exam grade cutoffs

Notes: The figure plots the grade cutoffs over time. We identify the cutoffs from the data by first classifying each score in an exam to a grade which is most frequent. For example in the English exam in year 1990 there were 36 observations with the score 151. 33 of the 36 got the grade 0 and 3 got the grade 1 so we classify the score 151 to grade 0. After completing this classification we set the cutoffs between the highest score that is classified as the smaller grade and the lowest score that is classified as the higher grade. Continuing with the English exam in 1990, the highest score classified to grade 0 is 151 and the lowest score classified to grade 1 is 152 so we set the cutoff between grades 0 and 1 to 151.5.

Table 3: Covariate balance

Cutoff %ile	5	16	36	60	80	95	Pooled
Age (Years)	18.98	18.89	18.82	18.78	18.77	18.77	18.81
Estimate	0.01	0.00	-0.00	0.00	-0.00	0.00	0.00
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Placebo	0.00	-0.01	-0.00	-0.00	0.00	-0.00	-0.00
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Woman (%p.)	53.53	55.25	57.55	60.39	62.66	60.55	59.18
Estimate	0.91	-0.61	-0.49	-0.26	0.60	-0.09	-0.08
	(0.91)	(0.55)	(0.44)	(0.41)	(0.41)	(0.76)	(0.21)
Placebo	-0.26	0.33	-0.23	-0.53	-0.30	-1.26	-0.25
	(0.66)	(0.48)	(0.42)	(0.41)	(0.44)	(1.19)	(0.20)
M. College (%p.)	6.98	8.94	11.76	15.67	20.65	31.05	15.60
Estimate	0.76	-0.05	-0.36	0.33	-0.31	0.56	0.00
	(0.47)	(0.31)	(0.28)	(0.30)	(0.34)	(0.75)	(0.15)
Placebo	-0.04	0.44	-0.17	0.18	-0.16	1.20	0.08
	(0.36)	(0.29)	(0.29)	(0.32)	(0.40)	(1.23)	(0.15)
F. College (%p.)	11.24	13.83	17.43	22.15	27.94	37.32	21.73
Estimate	1.11	0.78	0.29	0.47	0.22	0.37	0.44
	(0.59)	(0.38)	(0.34)	(0.34)	(0.38)	(0.78)	(0.17)
Placebo	0.12	-0.03	-0.15	0.31	0.23	0.08	0.09
	(0.44)	(0.35)	(0.34)	(0.36)	(0.43)	(1.26)	(0.17)

Notes: The five rows for each covariate report the mean, the effect of being above the cutoffs on the covariate (with its standard error in parenthesis) and the effect of placebo cutoffs. We run the regressions separately for each cutoff, where we pool over years and subjects. The results of these regressions are reported in the first six columns. In the last column we pool all cutoffs together.

the P-values in parenthesis.

A possible mechanism for manipulation are rescoring requests, which can only results in an increase in the score. The requests are rare, 0.2% of all exams and result in an increase in the score 25% of the time. We would expect to see the rescoring affecting the score most around the pass/fail cutoff, but this is not born out in the data: the distribution across the pass/fail cutoff is smooth.

We do not use covariates in our regressions, but here we check for jumps in them, because jumps would hint that there is possible manipulation at the cutoffs. We run regression Eq. (1) with each covariate as an outcome variable in turn. The covariates we consider are age of the student, her gender and having college educated parents. The results of the regressions are reported in Table 3.

The means of the covariates are reported in the row that starts with the underlined name of the covariate in Table 3. The mean age for taking an exam is 18 years 9 months and 23 days. 59% of the students are women. 16% of their mothers and 22% of the fathers are college educated.

When we pool the cutoffs together the share of college educated fathers jumps between 0 to 0.8% at the cutoff, but there are no other significant jumps in the covariates. The jump in college educated fathers is greatest at the first two cutoffs. There is also one statistically significant jump in the covariates at the placebo cutoffs. The age of the students decreases between zero and 6 days at the second cutoff.

5 Results on high school outcomes

The cutoffs affect mechanically the grades that the student receives from the exam. In the case where the student was on the margin of graduating from high school, failing to pass an exam can lead to postponing graduating or failing to graduate all together. In addition to the mechanical effect, the cutoffs can affect students choices to repeat the exams.

We consider the effect of crossing a cutoff on the grade the student gets from the exam, the probability of repeating the exam, the final grade the student gets after possibly repeating the exam and the probability of graduating from high school by age 30. Because students have a maximum of 2 years to repeat their exams, we drop the last two years of observations when considering the effect of cutoffs on final grades and repeating. Table 4 shows the effect of crossing a cutoff on the highs school outcomes. The row "Estimate" is the parameter estimate on the cutoff ("Placebo" on the placebo cutoff) dummy with its standard error below in parenthesis. The first row of each underlined outcome reports the mean of the outcome. The columns report the effect of crossing the individual cutoffs and the effect when we pool all cutoffs together.

The grade cutoffs determine the grade that the student receives from the exit exam almost always. The average difference in the grade across all grade cutoffs is 0.97. Crossing the pass / fail cutoff increases the grade by 0.91, which means that 10% of the students who score just below the cutoff still receive a passing grade. This could be due to extra points being awarded after the scoring for medical conditions and other adverse circumstances. Unsurprisingly the placebo cutoffs that are in the middle of the grades don't affect the grades.

The average effect of all grade cutoffs on the final grade of the student is 0.79. Not surprisingly the effect of the cutoff on the final grade is especially low (0.25) at the pass / fail cutoff. The effect on the final grade is smaller than the effect on the grade from the exam because students below the cutoff repeat the exam more often (especially when they fail the exam) and when

Table 4: High school outcomes

Cutoff Percentile	5	16	36	60	80	95	Pooled
Grade	0.60	1.55	2.53	3.51	4.48	5.40	3.21
Estimate	0.91	0.97	0.98	0.98	0.98	0.96	0.97
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Placebo	0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Final Grade	0.92	1.63	2.61	3.59	4.53	5.44	3.29
Estimate	0.25	0.82	0.81	0.81	0.86	0.86	0.79
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Placebo	0.00	-0.00	-0.00	-0.00	0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Repeated	0.33	0.11	0.12	0.12	0.09	0.08	0.12
Estimate	-0.51	-0.09	-0.12	-0.13	-0.15	-0.17	-0.16
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Placebo	0.00	-0.00	-0.00	-0.00	0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Graduated	0.83	0.97	0.98	0.99	0.99	0.99	0.97
Estimate	0.12	0.00	-0.00	0.00	-0.00	0.00	0.01
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Placebo	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Notes: The five rows for each outcome report the mean of the outcome, the effect of being above the cutoff on the outcome (with its standard error in parenthesis) and the effect of placebo cutoffs. We run the regressions separately for each cutoff, where we pool over years and subjects. The results of these regressions are reported in the first six columns. In the last column we report results were we pool all cutoffs together.

they do repeat, they improve their grade more often than students above the cutoff. 12% of all students repeat the exam. The probability to repeat is distributed unequally on different sides of the cutoff as crossing the cutoff reduces the probability to repeat the exam by 16% points. Even though clearing the higher cutoffs does not affect graduation from high school, students choose to repeat the exam to improve their final grade. The decrease in the probability of repeating as a share of the mean probability around the cutoff is increasing in the cutoffs. This means that the higher the cutoff, the more the repeating is concentrated to students who scored below the cutoff than above.

When deciding whether to repeat or not the students weight the opportunity cost of repeating an exam against the probability of succeeding in getting a higher grade and the benefit the higher grade gives. The cutoffs create discrete jumps in the probability of increasing the grade and possibly in the opportunity cost of repeating an exam as well. The probability of getting a higher grade by repeating an exam is higher just below the cutoff than just above because below even a small improvement in the score will result in the higher grade where as above the improvement has to equal at least one total grade. The opportunity cost of repeating the exam could be higher above the cutoff for example because a higher grade increases post-secondary education opportunities. The time spent preparing for repeating an exam could be spent more productively studying above the cutoff than below it. Because of the jump in the probability and possibly in the opportunity cost, at a minimum the cutoffs affect the composition of the repeaters and plausibly also the aggregate repeating when compared to the alternative of releasing the information at the level of the scores.

In spite of almost everyone below the pass/fail cutoff repeating the exam, passing the cutoff still has a sizable effect on graduating from high school. The average effect of an cutoff on graduating high school is 1 percentage point. As expected the effect is driven completely by the lowest cutoff which increase the probability of graduating form high school by 10 to 14 percentage points.

6 Results on education outcomes

From the previous section on the effect of the cutoffs on high school outcomes we learned that the cutoffs affect the grades that students graduate with from high school. Because the grades are used in selecting students to post-secondary education, crossing a cutoff could improve a students education opportunities. Now we will find out how education outcomes are affected by the cutoffs.

The first five rows of Table 5 report the effect of the cutoffs on years of schooling of the highest attained degree at age 30. We transform the degrees to years of schooling in the following way: high school dropout = 9 years of schooling, high school degree = 12, vocational degree = 12, applied degree = 15, bachelor's degree = 15, master's degree = 17, licentiate degree = 19, PhD = 21. Years of schooling captures differences in the types of degrees students get on different sides of the cutoff. For example attending a university instead of an applied college. The mean years of schooling at age 30 varies significantly with the results of the exit exam. Students closest to the 5th percentile cutoff have on average 13 years of schooling where as students closest to the 95th percentile have 15 years of schooling. The average over all cutoffs is 14.1 years. Being above a cutoff has a small positive effect on years of schooling. Years of schooling increases by 0.01 to 0.05 years or from 4 to 18 days at the cutoff. To make the comparisons of the effect sizes easier we report the point estimate as percent of the mean in the rows titled "Relative effect". The effect of a cutoff on years of education is 0.2% of the mean. The final two rows for each outcome report the estimated effect of a placebo cutoff, which acts as a check on our functional form assumptions for the running variable. The effect of the placebo cutoff should not be significantly different from zero. This is true for the pooled sample, but not for the lowest 5th percentile cutoff.

Crossing a cutoff can cause changes in students education over and beyond what is reflected in years of schooling, which only captures changes in the type of degree. For example a student could attend a more selective institution or study a more desirable subject as a result of being above a cutoff, which would not be reflected in years of schooling. Because better results in high school exit exams afford more choices in post-secondary education, we proxy the quality of an education program by the exit exam results of the students who have graduated from the program. Since programs use also entry exams to select students, differences in the exit exam results do not capture the differences in the quality of the programs completely.

Because students can choose which exit exams they take, quantifying success in the exams is not straight forward. We convert the exit exam results to interval scale by anchoring them to mean earnings from ages 30 to 34. We regress the earnings on results from mathematics and Finnish as native language exit exams and use the fitted values as the anchored exit exam results for each student. We calculate the anchored exit exam value for each education program by averaging over anchored exit exam results of the students who have graduated from the program. Our outcome variable is the anchored exit exam value of the highest attained education program at age 30. A program is defined by institution and subject. For example if by age 30 a student would have a bachelor's degree in chemistry from Helsinki University of Applied Sciences and a master's in chemistry from University of Helsinki the value of her outcome would be the mean of the anchored exit exam results for students with a master's in chemistry from University of Helsinki. If a student hasn't graduated from any program after taking the high school exit exams, we code the observation as missing. Being above or below a cutoff doesn't change the exam results of other students, but can potentially affect which program the student enters and graduates from, thus affecting who her peers are. We explain the details of the anchoring in Appendix B.

The anchored exit exam results of the program that students graduate from is 28,000 euros on average. The average varies by cutoff from 27,000 euros at the 5th percentile cutoff to 30,000 euros at the 95th percentile. The effect of a cutoff is quantitatively small, but precisely estimated. A cutoff causes a student to attend a program with 10 to 70 euros higher anchored exit exam results on average, which is 0.15% of the mean. Hence a cutoff has a larger effect on the quantity of education than the average exit exam results of the program where the student graduated from. The estimates for the placebo cutoffs are not significantly different from zero.

The last three education outcomes we consider are the types of degree the student has attained by age 30. 20% of the students have a vocational degree, 38% have a degree from an applied college and 24% have a university degree. The share of university degree is strictly increasing in the closest cutoff. Only 4% of the students closest to the bottom cutoff have a university degree, where as half of the students who score closest to the top cutoff have one. Share of vocational degree is decreasing in cutoff, with 29% of students holding one closest to the bottom cutoff and 7% closest to the top cutoff. Applied degrees are in the middle. The share peaks at 44% closest to the third cutoff at the 36th percentile.

Crossing a cutoff causes a between 0.6 to 1.2 percentage point increase in the probability of having an university degree at age 30, which is 4% of the mean. The point estimate is positive for all cutoffs, but significant and largest for the cutoffs from the 36th percentile to the 80th percentile. The average effect of a cutoff on the probability of having an applied degrees is between -0.4 and 0.1 percentage points. The bottom cutoff has a large positive effect between 1.3 and 4.2 percentage points, the middle cutoffs have small and insignificant effects and the top cutoff has a large negative effect between -2.8 and -0.4 percentage points. This reflects the fact that applied degrees are the first choices for the lower scoring students and the second choices for the high scoring students. The over all effect of the cutoffs on the probability of having a vocational degree is negative, between -0.4 and -0.0 percentage points. The effect of the bottom cutoff is clearly positive between 3.4 and 6.1 percentage points, which is surprising, since crossing pass/fail cutoff has a positive effect on graduating from high school and vocational education is commonly thought of as a substitute for high school. The next three cutoffs on the other hand have a significant negative effect on the probability of having a vocational degree with the highest two having an insignificant negative effect. This means that for the lower scoring students vocational education is a real alternative to their preferred education. Higher scoring students attend vocational education rarely, but when they do it's more often their first choice.

Table 5 summarizes our results on education outcomes.

Based on these results on educational outcomes we would not expect to find large effects on labor market outcomes, which we will turn to next.

7 Results on labor market outcomes

We have yearly data on earnings and months employed from years 1987 to 2015. We estimate the regression in Eq. (1) separately for each age. The median effect of being above a cutoff on earnings is plotted in the first panel in Fig. 5. The solid line is the parameter estimate on the dummy of being above the cutoff. The shaded area is the 95% confidence interval of the estimate. The dashed line is the parameter estimate of the effect of a placebo cutoff. The placebo cutoffs are the midpoints of the grades and shouldn't affect the outcome if our chosen functional form is succeeding in capturing the effect of the running variable on the outcome. plots our findings for the effect of an cutoff on earnings pooled over cutoffs. The left panel plots the parameter estimates and the right panel the parameter estimates as percent of the mean earnings for the age. The effect of a cutoff on earnings is close to zero for age 20, but increases with age and peaks at age 48 between €600 and €1900. Earnings increase significantly with age. The mean earnings for at age 20 are €5800 and €39,300 at age 50. So it is not surprising that the effect of the cutoff in euros also increases with age. As we can see from the right panel of Fig. 5, the effect of the cutoff as a percentage of the mean earnings also increase in age. The effect is between -1 and 0% until age 28 and then starts increasing and reaches between 1.5 and 4.5% of the mean at age 48. Reassuringly the estimates for the placebo cutoffs are centered around zero and do not show a similar increase as the actual cutoffs with age. However none of the parameter estimates for the cutoffs are statistically significantly different from the placebos at the 5% level.

Part of the effect of the cutoff on earnings is explained by the effect of the cutoff on employment that we illustrate in Fig. 6. The cutoffs reduce employment by between 0.05 and 0 months, or 1.5 days in a year before age 28. The effect is centered around zero from age 28 to age 43. The cutoffs increase employment from age 43 onward between 0 to 0.1 months, or 3 days in a year. As a percentage of mean employment the effect of the cutoff matches the effect on earnings for ages 20 to 28 and is smaller after age 28. The estimates for the placebo cutoffs are similar in magnitude to the estimates of the actual cutoffs, which implies that we are not able to capture the effect of the running variable on employment around the placebo cutoffs. This casts serious doubt that we would be able to capture the effect of the running variable on employment around the actual cutoffs. Failing to do so

Table 5: Effect of cutoffs on education outcomes

Cutoff Percentile	5	16	36	60	80	95	Pooled
Years	13.00	13.41	13.82	14.26	14.70	15.01	14.10
Estimate	0.03	0.04	0.03	0.04	0.02	-0.00	0.03
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)	(0.01)
Relative effect (%)	0.19	0.27	0.23	0.32	0.17	-0.03	0.23
Placebo	0.06	0.03	-0.01	0.03	-0.04	0.02	0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.06)	(0.01)
Peers (€)	26760	27112	27583	28203	28968	29850	28103
Estimate	-22.96	-11.97	83.72	68.88	-0.06	69.14	42.52
	(43.82)	(33.07)	(31.27)	(33.31)	(38.49)	(76.51)	(16.23)
Relative effect (%)	-0.09	-0.04	0.30	0.24	-0.00	0.23	0.15
Placebo	26.13	13.09	10.19	27.80	-8.91	224.44	27.36
	(36.68)	(31.82)	(32.21)	(35.40)	(43.55)	(131.16)	(16.37)
University (pp)	4.23	8.64	15.79	26.03	38.34	50.00	24.21
Estimate	0.40	0.11	0.66	1.35	0.94	1.17	0.91
	(0.31)	(0.28)	(0.31)	(0.35)	(0.40)	(0.76)	(0.16)
Relative effect (%)	9.49	1.31	4.21	5.20	2.45	2.34	3.74
Placebo	0.34	0.38	0.16	0.62	-0.56	0.81	0.27
	(0.29)	(0.29)	(0.33)	(0.37)	(0.44)	(1.20)	(0.16)
Applied (pp)	35.38	43.80	44.21	39.88	31.78	20.26	37.88
Estimate	2.77	0.28	-0.38	-0.41	-0.49	-1.61	-0.26
	(0.73)	(0.50)	(0.42)	(0.39)	(0.38)	(0.60)	(0.19)
Relative effect (%)	7.83	0.63	-0.87	-1.02	-1.53	-7.94	-0.68
Placebo	0.99	0.13	-0.59	0.31	-0.09	-1.39	-0.10
	(0.58)	(0.45)	(0.40)	(0.38)	(0.38)	(0.80)	(0.19)
Vocational (pp)	29.38	30.64	24.73	17.74	11.54	6.61	19.80
Estimate	4.71	-1.21	-1.04	-0.74	-0.34	-0.11	-0.36
	(0.70)	(0.47)	(0.36)	(0.30)	(0.26)	(0.37)	(0.16)
Relative effect (%)	16.04	-3.95	-4.22	-4.15	-2.92	-1.61	-1.82
Placebo	-0.23	-0.32	-0.12	-0.34	0.02	0.33	-0.18
	(0.56)	(0.41)	(0.33)	(0.28)	(0.25)	(0.50)	(0.15)

Notes: The six rows for each outcome report the mean of the outcome, the effect of being above the cutoff on the outcome (with its standard error in parenthesis), the pseudo elasticity of the effect i.e. the effect divided by the mean and the effect of placebo cutoffs. The first six columns report the results for the separate regressions for each of the cutoffs and the last column reports the results where we pool all cutoffs together. The first outcome is years of schooling. The second outcome captures the exit exam success of the student's peers in post-secondary education. We anchor the mathematics and native Finnish exam results to earnings at age 35. We then calculate the mean of the anchored exam results for each education program (institution times subject). The final three outcomes are the probability (in percent) that the student has a particular degree. All outcomes are measured at age 30.

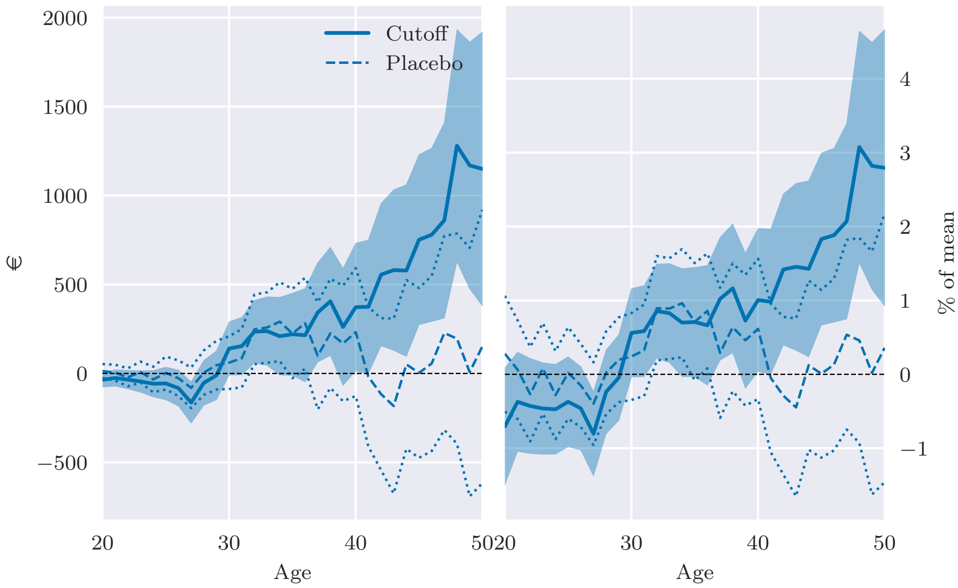

Figure 5: Earnings

Notes: The left panel of the figure plots the parameter estimates on the dummy of being above a cutoff of Eq. (1) for each age in the solid line and its 95% confidence interval in the shaded area. The dashed line are the parameter estimates on the dummy of being above a placebo cutoff for each age and its confidence interval. The left panel plots the same parameter estimates as a percentage of the mean earnings for the age. The sample sizes vary between 1.2 million for age 20 to 160,000 for age 50. Mean earnings at age 20 are €5,800 and €39,300 at age 50.

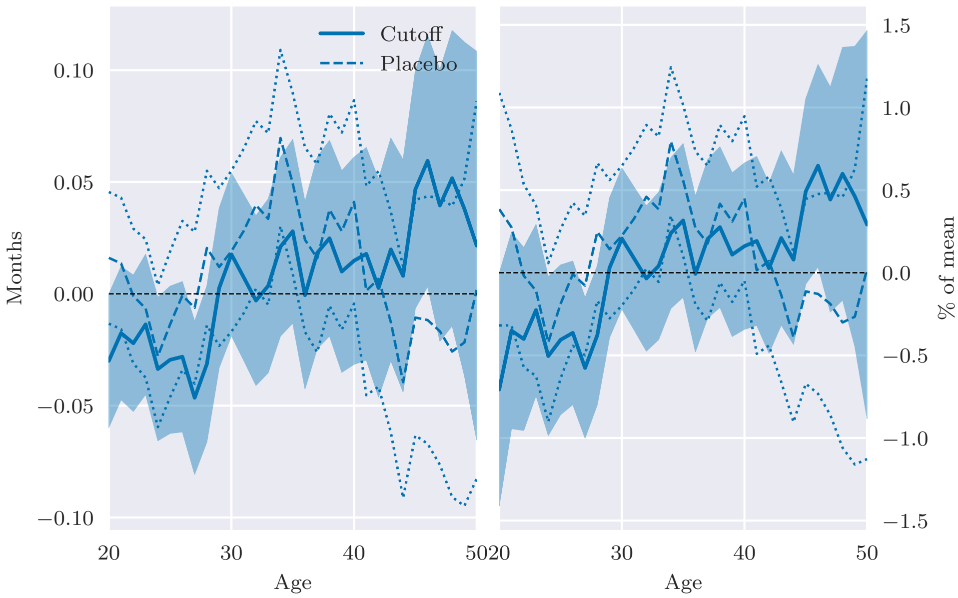

Figure 6: Employment

Notes: The left panel of the figure plots the parameter estimates on the dummy of being above a cutoff of Eq. (1) for each age in the solid line and its 95% confidence interval in the shaded area. The dashed line are the parameter estimates on the dummy of being above a placebo cutoff for each age and its confidence interval. The left panel plots the same parameter estimates as a percentage of the mean employed months for the age. The sample sizes vary between 1.2 million for age 20 to 160,000 for age 50. Mean employment varies between 4 months at age 20 to 10 months at age 44.

would result in spurious effects of the cutoffs.

The increasing age profile of earnings is not consistent with the grades having information value to the employers about the potential productivity of the students. If this were the case the effect should be positive at younger ages and decrease as the employers get more information on the actual productivity of their employees. The increasing age profile could be consistent with the grades containing information about the ability of the students to learn to be productive, which would not increase their earnings early in career, but cause employers to hire them to careers where productivity increases faster with experience and hence with a steeper age profile of earnings.

Another explanation for the increasing age profile of the effect of the cutoff on earnings would be that passing a cutoff increases education opportunities, which increases investment in human capital. The opportunity cost of getting more education is not working as much which would translate to lower earnings and employment early-in-career. The benefit is higher earnings once students start in full-time employment. Since the exit exam grades are used in determining access to post-secondary education, crossing a cutoff therefore increases the education opportunities students have on the margin. The effect of the cutoffs on earnings and employment is consistent with this explanation as the effect is negative for both earnings and employment during the years when students are typically in post-secondary education and the effect on earnings becomes positive when students start in full-time employment. Based on this reasoning we expect that increased education opportunities explain at least some of the observed effect on earnings. But how much is due to education?

Previous studies on high school exit exams by Canaan and Mouganie (2018) and Ebenstein, Lavy, and Roth (2016) assume that the exams effect on earnings is solely due to increased access to education. Next we show evidence that in the Finnish context at most two thirds of the effect on earnings is due to education opportunities and at least one third due to other mechanisms.

We use median earnings of the students who have graduated from the same education program as an outcome variable. Crossing a cutoff doesn't affect the median earnings of the actual programs per se, but as we have seen, crossing a cutoff has an effect on where the student ends up graduating from and thus also the median earnings of her program.

The effect on program's median earnings is the upper bound for the effect on student's earnings that is caused by increased access to education when the differences in programs' median earnings is greater than the differences in programs' earnings for the marginal student in the program. By marginal students we mean the students who graduate from a different program depending on which side of the cutoff they score in the exit exams i.e. the compliers. The exit exam grades are used in selecting students into post-secondary education, but are irrelevant after getting access. It is plausible that the grades affect graduating from an education program only through affecting the access to the program and not for example increasing the likelihood of graduating conditional on having access. If this is the case, when the marginal students score above the cutoff they are among the last students to enter their programs.

The marginal students earnings could differ less across the cutoffs than the median students earnings because of positive selection into the programs that explains part of the difference in median earnings. We already found evidence of this selection in Section 6. The median exit exam scores of the programs above the cutoff are 0.13% higher.

It is also possible that the difference in earnings for the marginal student is greater than the difference in median earnings between the programs. Kirkeboen, Leuven, and Mogstad (2016) shows that students sort into programs based on their comparative advantage i.e. the students who prefer a program over another gain more from the program than students who have the opposite preferences. We know that marginal students prefer the program they graduate from when they score above the cutoff, but we don't

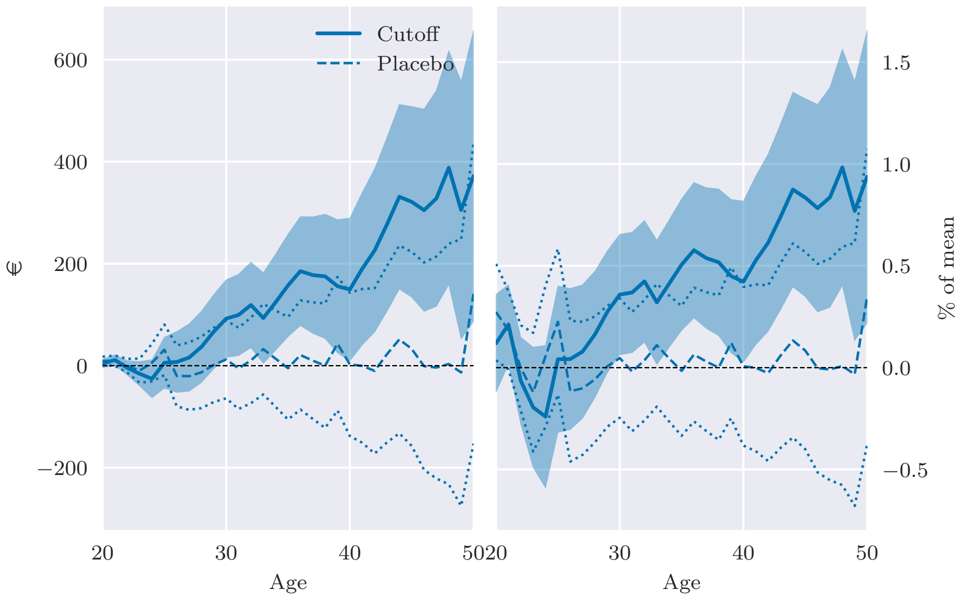

Figure 7: Degree median earnings

Notes: The left panel of the figure plots the parameter estimates on the dummy of being above a cutoff of Eq. (1) for each age in the solid line and its 95% confidence interval in the shaded area. The dashed line are the parameter estimates on the dummy of being above a placebo cutoff for each age and its confidence interval. The left panel plots the same parameter estimates as a percentage of the mean program median earnings for the age. The sample sizes vary between 1.4 million for age 20 to 150,000 for age 50. Mean of the education programs' median earnings at age 20 are €3,500 and €39,700 at age 50.

know if this holds also for the median student in the program. However since the marginal students above the grade cutoffs are among the last students to enter their programs they could benefit less from the instruction than the median student.

Since different education programs can have different age profiles for earnings, we use the median earnings of the program at an age corresponding to the students age. For example for a student who graduates from high school at age 19 and gets a Master's Degree in Chemistry from the University of Helsinki at 25, the outcome variable at ages between 19 and 24 would be the median earnings of high school graduates for the corresponding age (median earnings at age 19 for the outcome at age 19, median earnings at age 20 for the outcome at age 20 etc.) and from age 25 onward the median earnings of graduates from University of Helsinki with a Master's degree in Chemistry for the corresponding age.

The effect of the cutoff on education programs median earnings is illustrated in Fig. 7. The age profile of the effect is increasing in age similarly to the cutoffs effect on earnings. The effect on programs median earnings is zero at age 20, when practically all students have graduated from high school, but not yet from post-secondary education. The effect starts to increase steadily after age 25, when student's start graduating from post-secondary education. The effect is between €100 and €300 at age 35 and between €150 and €500 at age 45. The estimate on the placebo cutoffs is centered around zero for all ages. Similarly to the estimates on individuals earnings, the standard errors are so large that the placebo effects are significantly different from the actual estimates at the 5% level. In the right panel we show the parameter estimate as a share of the mean for the corresponding age. Just like for earnings, the relative size of the effect increases consistently with age and peaks at between 0.5 and 1.5% of the mean at age 47. To relate the size of the effect on programs median earnings to the effect on earnings we divide the former point estimates with the latter. From age 30 to age 45 the effect on programs median earnings is between 40 and 90% of the effect on earnings with a mean of 55%. This result on the ratio of the effects of programs median earnings and individuals earnings is close to Canaan and Mouganie (2018), who estimate a jump at the cutoff in the average earnings of Master's level graduates from the same major and institution that is 60% of the estimated effect on the individuals earnings.

8 Robustness

The estimated effects for placebo cutoffs are in our view the most compelling evidence for the robustness or lack thereof of our results. Because the discreteness of our data forces us to choose a functional form for how we control for the running variable and the bandwidth, we show here that our results are not much affected by these choices.

Let us start first by visually inspecting the rdd-plots, which we have relegated to Appendix C. The high school outcomes are directly affected by the cutoffs and unsurprisingly reveal clear jumps at the cutoffs. For the education and labor market outcomes we see that the jumps across cutoffs are usually not clearly visible except for the outcomes related to earnings. There is also quite a bit of variance in the mean of the outcomes at each distance to the cutoff indicating that our results for low bandwidths should be very sensitive to the choice of bandwidth.

We examine the sensitivity of our results to the bandwidth by repeating our analysis for bandwidths from only including observations that are at most two points removed from the cutoffs to including observations that are at most 15 points removed. We plot the parameter estimate on the dummy of being above the cutoff with its standard error and the same for the placebo cutoffs in Appendix D. As expected from the rdd-plots, the high school results are unaffected by the choice of bandwidth. For the education and labor market results we see that the estimates are stable between bandwidths 4 to 10. The standard errors shrink as the bandwidth increases. The estimates on the placebo cutoffs are also sensitive to the choice of bandwidth and generally are consistently close to zero between bandwidths 4 to 10. As the bandwidth increases the estimate on the effect of a cutoff and the effect of the placebo diverge away from zero.

Next we consider the sensitivity of our results to the choice of controlling for the running variable linearly. We repeat our analysis with estimating a second degree polynomial for the score separately on both sides of the cutoffs. The figures in Appendix E plot the effects of the cutoffs with standard errors when we control for the running variable with first or second degree polynomials for our preferred bandwidth of 10. The effects of the cutoffs are less precisely estimated when we use a second order polynomial, but regardless the point estimates are very similar.

One concern is that since we don't keep the cohorts constant when estimating age effects, the age profiles that we report are actually a result of cohort effects, i.e. some cohorts benefiting more from high grades than others. To alleviate this worry we repeat our analysis restricting our sample to students who took their exit exams from 1982 to 1985 and who we can follow from age 20 to age 50. The results are basically the same as when we use all available data. The results can be found in Appendix F.

9 Conclusion

We look at how success in a high stakes exam affects students future education and labor market outcomes. We use discontinuities in grading in Finnish high school exit exam to control for the skill of the student. This regression discontinuity setting allows us to compare student who just marginally got a higher grade to those that got a lower grade. We find that the effect of a higher grade in a language exam causes higher earnings and that the earnings gain is increasing in age. The effect peaks between €600 and €1900 at age 48. Summed over 30 years from age 20 to age 50 the effect is between €1,000 and €19,000 or 1% of the summed mean earnings. The effect of the higher grade is roughly half of the effect of increasing the score by one standard deviation without an increase in the grade. Estimates for placebo cutoffs are not statistically different from zero indicating that we succeed in identifying the effect of the higher grade independent of the skill of the student.

The increase in age is inconsistent with students above a cutoff being perceived as more productive than students below. If this was the case then the earnings effect should decrease in age as the true productivity of the students below and above the cutoff are revealed to be the same over time. The increase in age is consistent with crossing the cutoff increasing access to higher education and resulting in more investment in human capital. However by comparing average earnings of student's peers in post-secondary education across a cutoff, we find that differences in education caused by a higher grade explain at most 60% of the effect on earnngs. The remaining 40% could be explained by the students above the cutoff getting jobs that build more valuable experience and human capital or an increase in the confidence of the students which would make them benefit more from post high school education.

We contribute to the understanding on how high stakes exams affect students. Our results complement the previous findings that success in high stakes exams have large effects on earnings in the first ten years after the exam. We find that in Finland there is no effect on earnings ten years after the exam. Ebenstein, Lavy, and Roth (2016) finds that in Israel a one standard deviation exogenous increase in the score increases earnings by 13% on average and 30% for students with high course grades. In contrast in Finland, a one standard deviation increase in the score[4] increases earnings by just one percent on average ten years after the exam. The effect is similar across the skill distribution except for at the highest grade cutoff (95th percentile). There increasing the score by one standard deviation increases earnings by 5%. Canaan and Mouganie (2018) use the cutoff that separates those who fail the exam from those who pass it. They find that passing the exam on the first try increase earnings by 12% ten years after the exam. In Finland crossing the pass/fail cutoff on the first try doesn't affect earnings significantly (-0.2% with a standard error of 1%). Both Canaan and Mouganie (2018) and Ebenstein, Lavy, and Roth (2016) assume that the earnings effect of the exit exam is caused by changes in education. In our context only up to 60% of the earnings effect is caused by changes in education.

Since we find evidence of effects from crossing a cutoff on earnings we conclude that two students who are essentially equal in skill end up in different circumstances because of the cutoffs. Releasing the information on a finer scale would reduce the inequality that the cutoffs introduce. A possible side effect could be to induce a smoother distribution of exam repeating where the decision to repeat would be due to real costs and benefits of repeating rather than the arbitrary distance to a grade cutoff.

[4] Students with a higher grade have on average half a standard deviation higher score than students with a lower grade. So we multiply the effect of an higher grade by two to arrive at the effect on increasing the score by one standard deviation.

References

Canaan, Serena and Pierre Mouganie (2018). "Returns to Education Quality for Low-Skilled Students: Evidence from a Discontinuity". In: *Journal of Labor Economics* 36.2, pp. 000–000. DOI: `10.1086/694468`. eprint: `https://doi.org/10.1086/694468`. URL: `https://doi.org/10.1086/694468`.

Cattaneo, Matias D, Nicolás Idrobo, and Rocıo Titiunik (2017). "A Practical Introduction to Regression Discontinuity Designs". In:

Ebenstein, Avraham, Victor Lavy, and Sefi Roth (2016). "The Long-Run Economic Consequences of High-Stakes Examinations: Evidence from Transitory Variation in Pollution". In: *American Economic Journal: Applied Economics* 8.4, pp. 36–65.

Frandsen, Brigham R (2017). "Party bias in union representation elections: Testing for manipulation in the regression discontinuity design when the running variable is discrete". In: *Regression Discontinuity Designs: Theory and Applications*. Emerald Publishing Limited, pp. 281–315.

Gelman, Andrew and Guido Imbens (2017). "Why high-order polynomials should not be used in regression discontinuity designs". In: *Journal of Business & Economic Statistics* just-accepted.

Kirkeboen, Lars J, Edwin Leuven, and Magne Mogstad (2016). "Field of study, earnings, and self-selection". In: *The Quarterly Journal of Economics* 131.3, pp. 1057–1111.

Kolesár, Michal and Christoph Rothe (2018). "Inference in Regression Discontinuity Designs with a Discrete Running Variable". In: *American Economic Review* 108.8, pp. 2277–2304. URL: `https://ideas.repec.org/a/aea/aecrev/v108y2018i8p2277-2304.html`.

Loukkola, Anna (2017). *Progress of studies 2015*. Tech. rep. Statistics Finland. URL: `http://tilastokeskus.fi/til/opku/2015/opku_2015_2017-03-17_tie_001_en.html?`.

Moisio, Johanna and Birgitta Vuorinen (2016). *Valmiina valintoihin: Ylioppilastutkinnon parempi hyödyntäminen korkeakoulujen opiskelijavalinnoissa*. Opetus-ja kulttuuriministeriö.

Tyler, John H, Richard J Murnane, and John B Willett (2000). "Estimating the labor market signaling value of the GED". In: *The Quarterly Journal of Economics* 115.2, pp. 431–468.

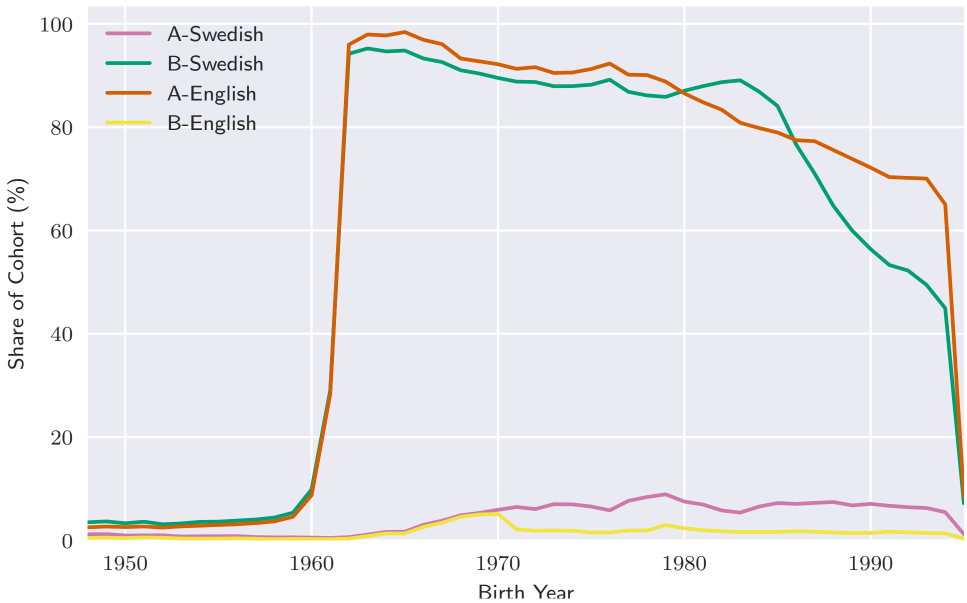

Figure 8: Share of A- and B-level Swedish and English exam takers in high school student cohort by birth year

A Additional institutional details

Fig. 8 plots the share of high school students who take final exams in the two English and Swedish exams. We see that the share of students who take the exit exam in A-level (higher) English has declined constantly from birth cohort 1962 onward. This decline is not due to student switching to B-level English as the share of B-level exit exams has remained low and constant. The same is not true for the Swedish exam. First the share of students who take the Swedish B-level exam is close to the share of English A-level exam takers and declines similarly until students born in 1980, when the decline stops and the share stays constant until birth cohort 1984 after which the share starts to decline quickly because the Swedish exam becomes optional for native Finnish speakers after year 2005.

B Anchoring of exit exam results

We are seeking to quantify the success of a student in the exit exams with a single interval measure. We first quantify the success in a single exam by the percentile of the score to make the exams comparable across exam seasons. Finnish as a native language and mathematics exams are the most popular exams, so we use them to quantify success. We do this by anchoring the exam results to mean earnings at ages 30 to 35. Specifically we regress the mean earnings on the percentiles of the mathematics and Finnish exams:

$$earnings_i = \sum_s \delta_s D_{s,i} + \gamma_s D_{s,i} f(p_{s,i}) + \beta X_i + \epsilon_i \tag{2}$$

where $s \in$ {Basic math, advanced math, Finnish} are the subjects we include in the regression D is a dummy for having taken the exam and p is the percentile of the score in the exam and X_i are predetermined control dummies for gender, college educated mother and father, native language and cohort. We use the resulting parameter estimates on the high school exit exam results δ_s, γ_s to calculate the fitted values for earnings given students score in the exit exams, but keeping the control variables the same for all, i.e. male, no college educated parents, Finnish as native language and having taken the exit exam in year 1974. These fitted values are the anchored exit exam results.

C RDD-plots

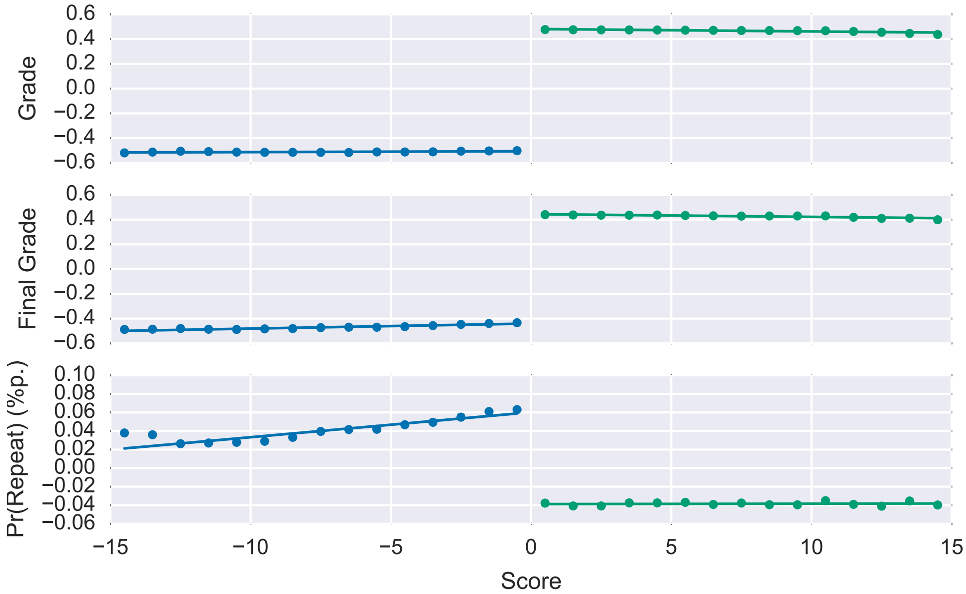

Figure 9: The relationship between high school outcomes and distance from the cutoff

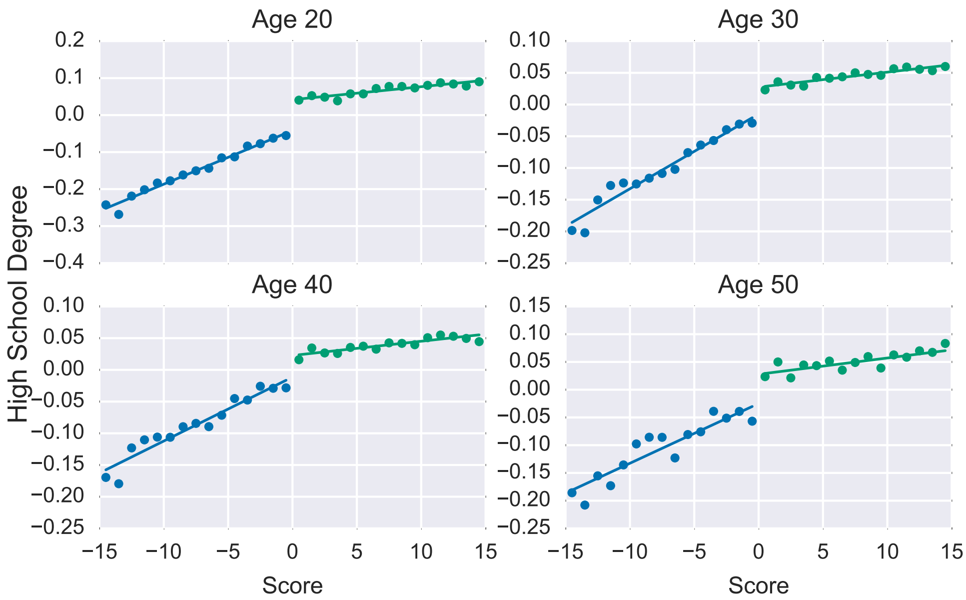

Figure 10: The relationship between high school degrees and distance from the cutoff

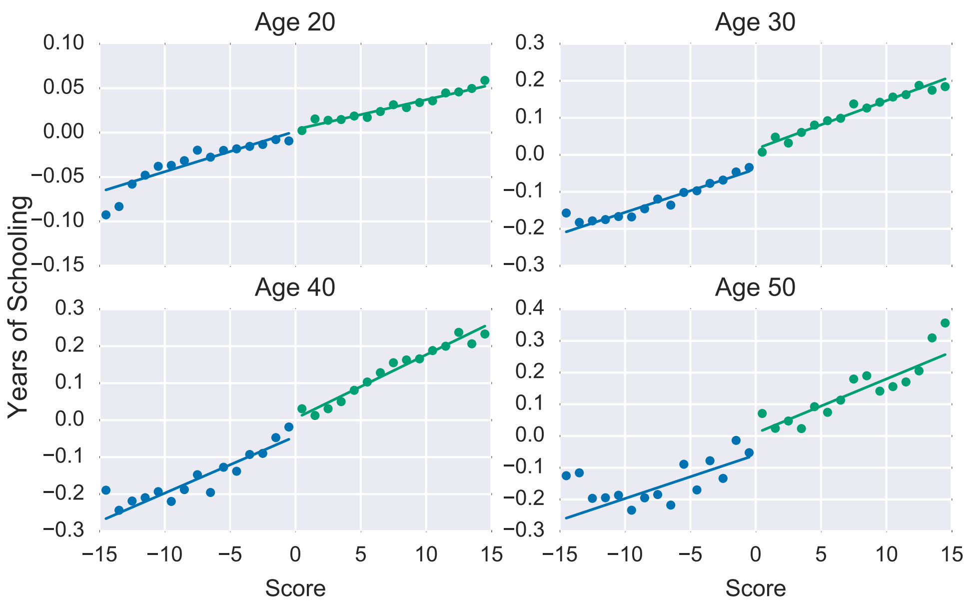

Figure 11: The relationship between years of schooling and distance from the cutoff

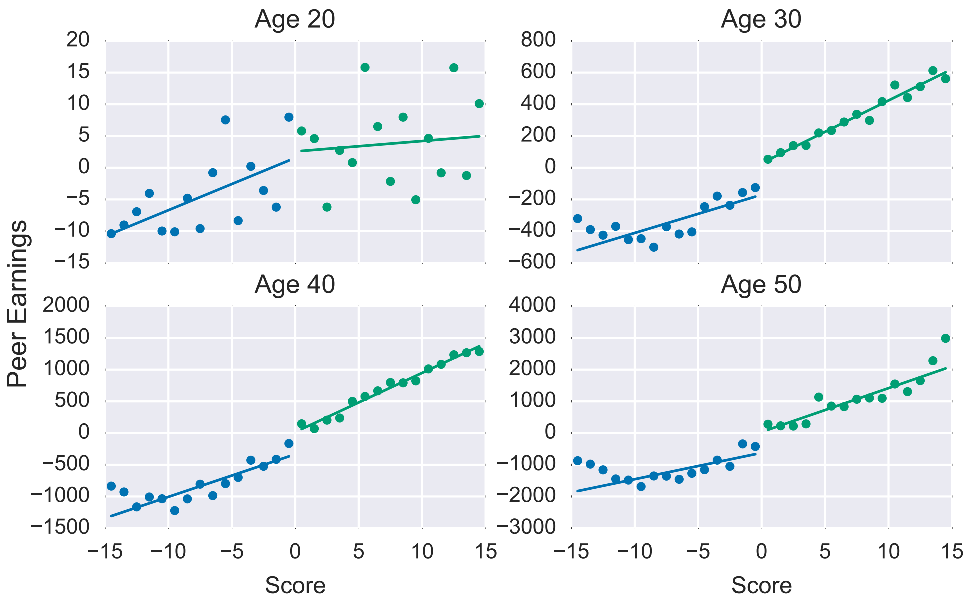

Figure 12: The relationship between peer earnings and distance from the cutoff

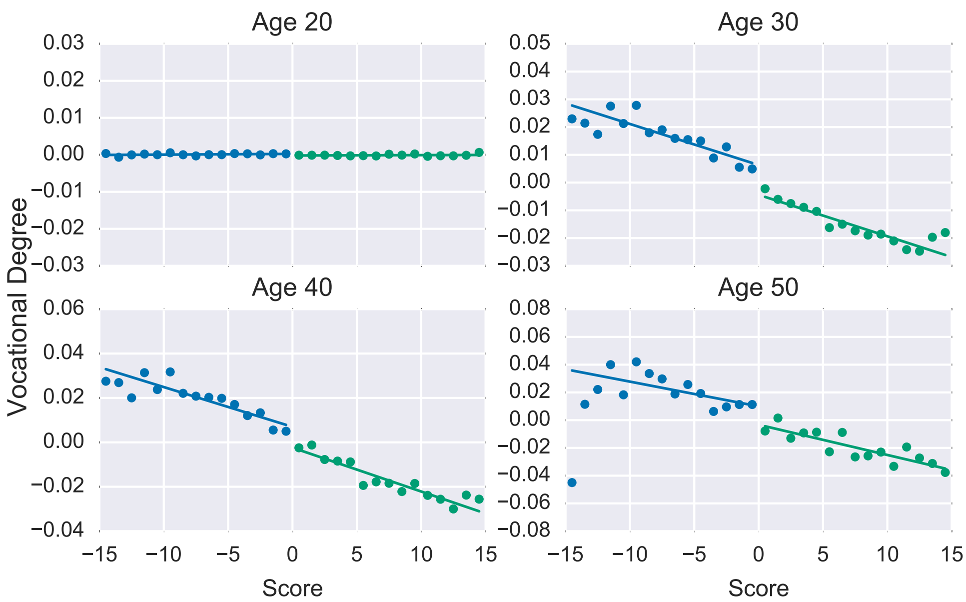

Figure 13: The relationship between vocational degrees and distance from the cutoff

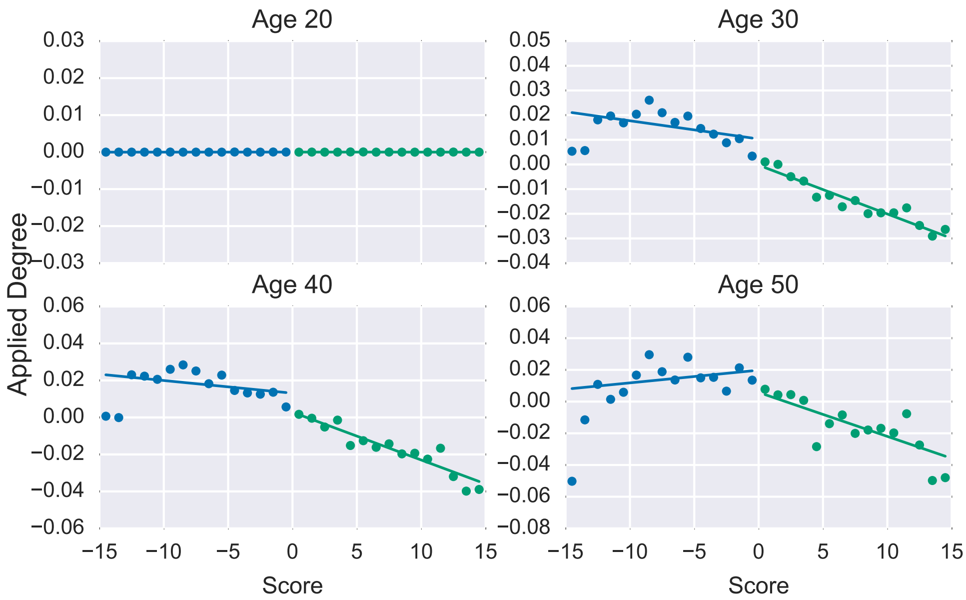

Figure 14: The relationship between applied degrees and distance from the cutoff

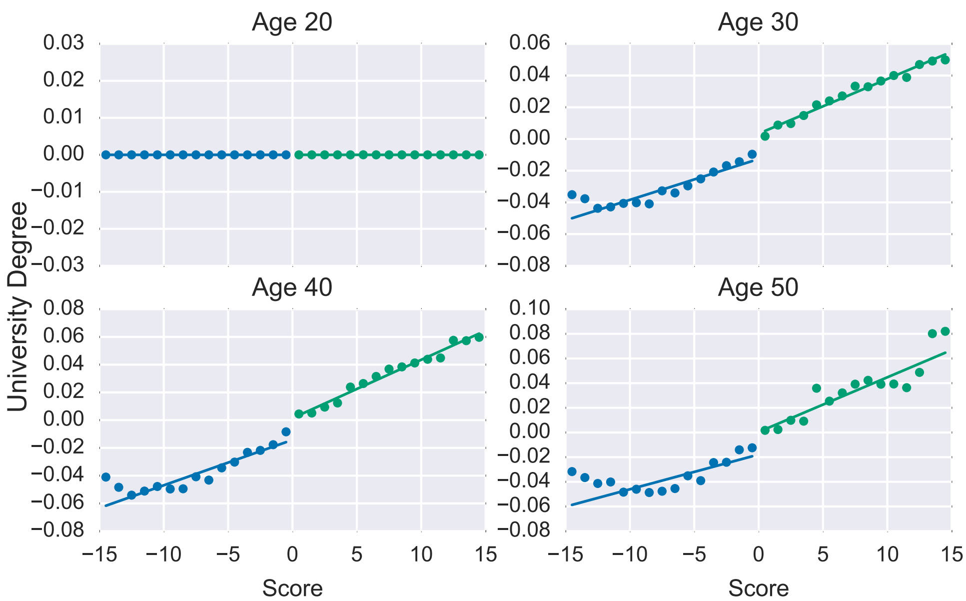

Figure 15: The relationship between university degrees and distance from the cutoff

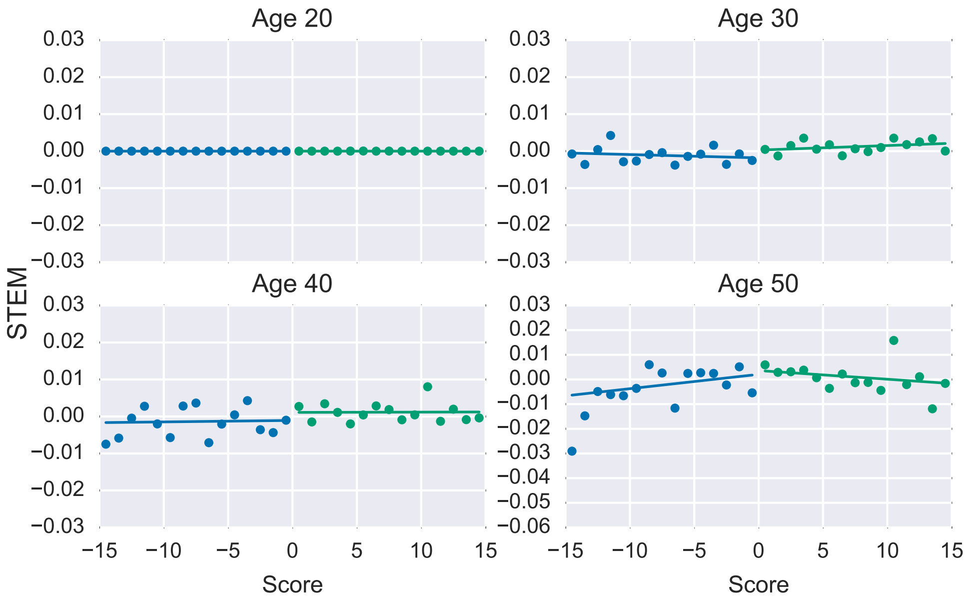

Figure 16: The relationship between stem degrees and distance from the cutoff

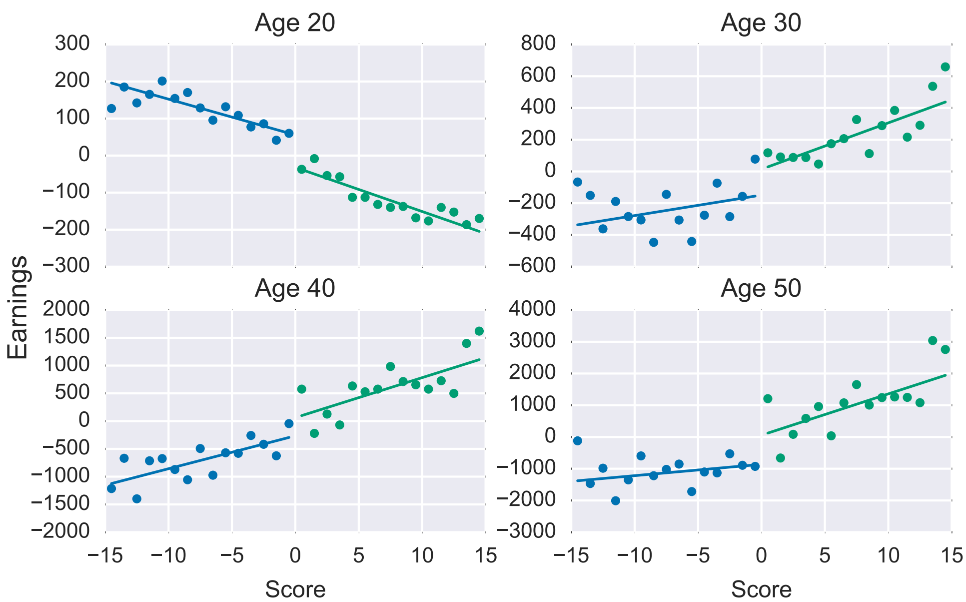

Figure 17: The relationship between earnings and distance from the cutoff

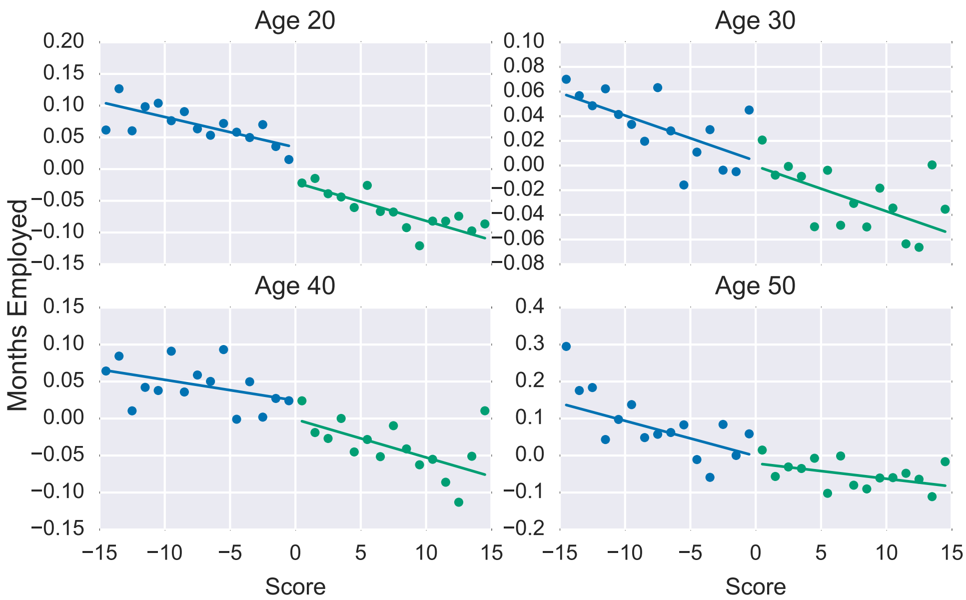

Figure 18: The relationship between employment and distance from the cutoff

D Results with alternative bandwidths

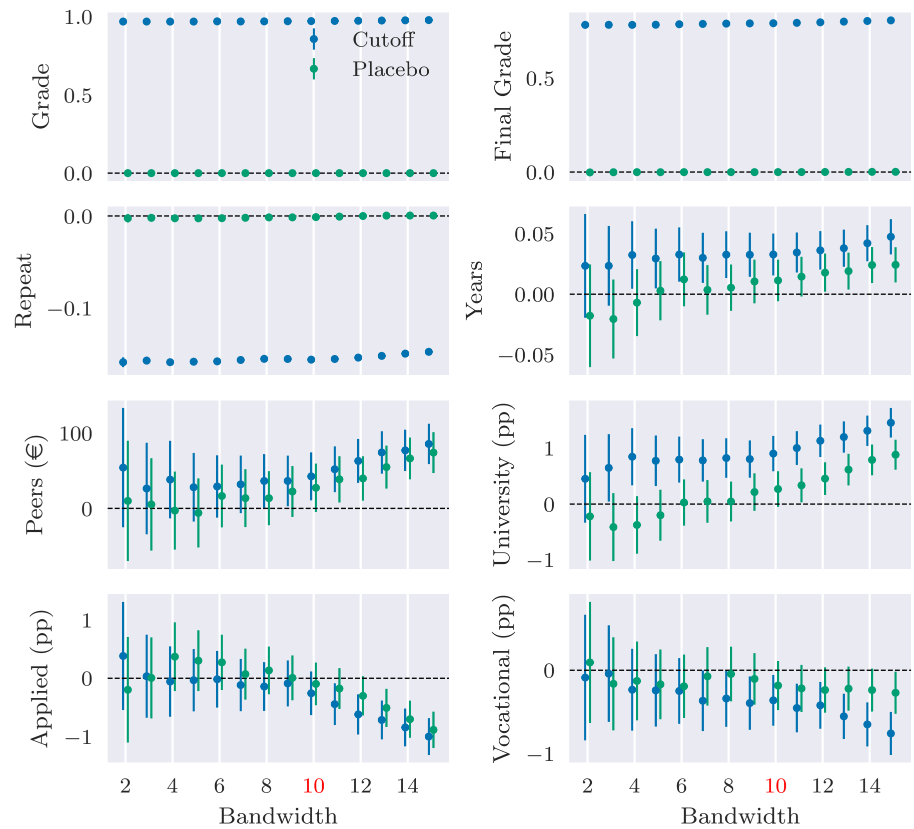

Figure 19: The effect of cutoff on high school and post secondary education outcomes with varying bandwidths

Notes: Each panel plots the parameter estimates and their standard errors for different bandwidths for the age that is at the top of the panel. When the bandwidth is higher than 10, observations start to overlap different cutoffs. I.e. an observation can be within the bandwidth of one cutoff from above and simultaneously be within the bandwidth of another cutoff from below.

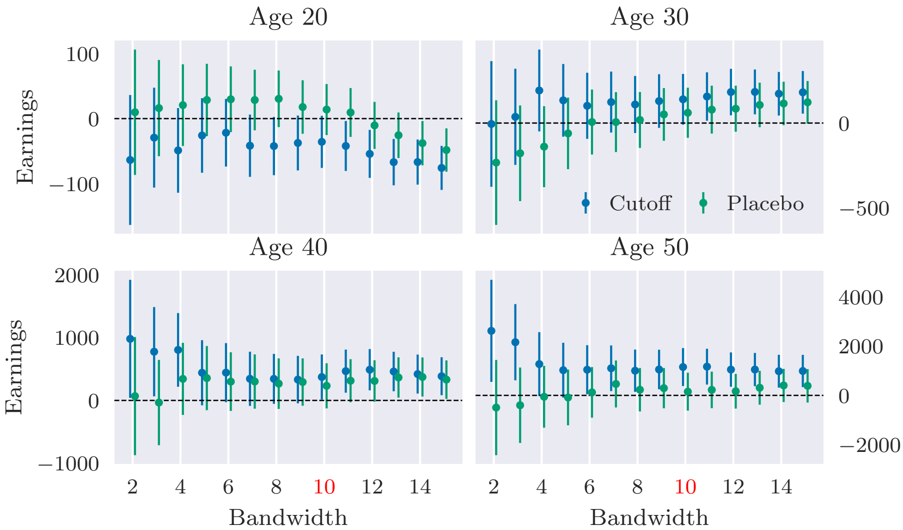

Figure 20: Earnings

Notes: Each panel plots the parameter estimates and their standard errors for different bandwidths for the age that is at the top of the panel. When the bandwidth is higher than 10, observations start to overlap different cutoffs. I.e. an observation can be within the bandwidth of one cutoff from above and simultaneously be within the bandwidth of another cutoff from below.

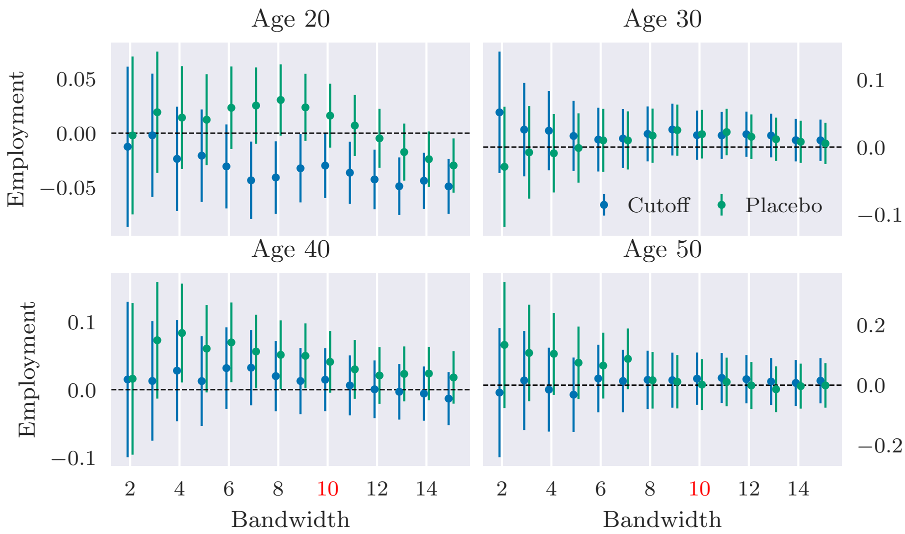

Figure 21: Employment

Notes: Each panel plots the parameter estimates and their standard errors for different bandwidths for the age that is at the top of the panel. When the bandwidth is higher than 10, observations start to overlap different cutoffs. I.e. an observation can be within the bandwidth of one cutoff from above and simultaneously be within the bandwidth of another cutoff from below.

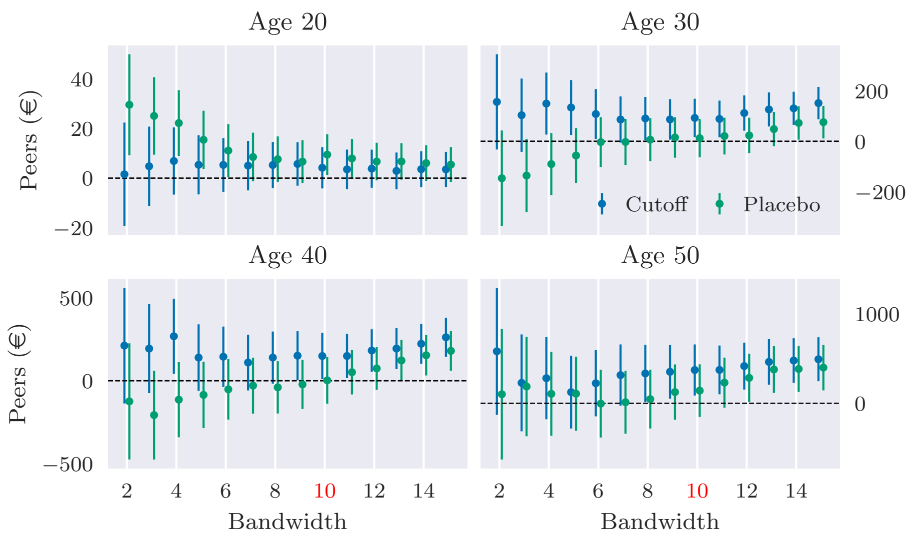

Figure 22: Degree median earnings

Notes: Each panel plots the parameter estimates and their standard errors for different bandwidths for the age that is at the top of the panel. When the bandwidth is higher than 10, observations start to overlap different cutoffs. I.e. an observation can be within the bandwidth of one cutoff from above and simultaneously be within the bandwidth of another cutoff from below.

E Results controlling the exam score with a second order polynomial

Table 6: High school outcomes controlling the exam score with a second order polynomial

Cutoff Percentile	5	16	36	60	80	95	Pooled
Grade	0.60	1.55	2.53	3.51	4.48	5.40	3.21
Estimate	0.90	0.96	0.97	0.98	0.98	0.97	0.97
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Placebo	0.00	0.00	-0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Final Grade	0.92	1.63	2.61	3.59	4.53	5.44	3.29
Estimate	0.23	0.81	0.80	0.80	0.85	0.86	0.78
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Placebo	-0.00	-0.00	-0.00	-0.00	0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Repeat	0.33	0.11	0.12	0.12	0.09	0.08	0.12
Estimate	-0.51	-0.09	-0.12	-0.14	-0.15	-0.17	-0.16
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Placebo	-0.00	-0.00	-0.00	-0.00	0.00	-0.01	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
High School	0.83	0.97	0.98	0.99	0.99	0.99	0.97
Estimate	0.10	0.00	-0.00	0.00	0.00	0.00	0.01
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Placebo	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Notes: The five rows for each outcome report the mean of the outcome, the effect of being above the cutoff on the outcome (with its standard error in parenthesis) and the effect of placebo cutoffs. We run the regressions separately for each cutoff, where we pool over years and subjects. Here we control for the running variable with a second order polynomial. The results of these regressions are reported in the first six columns. In the last column we report results were we pool all cutoffs together.

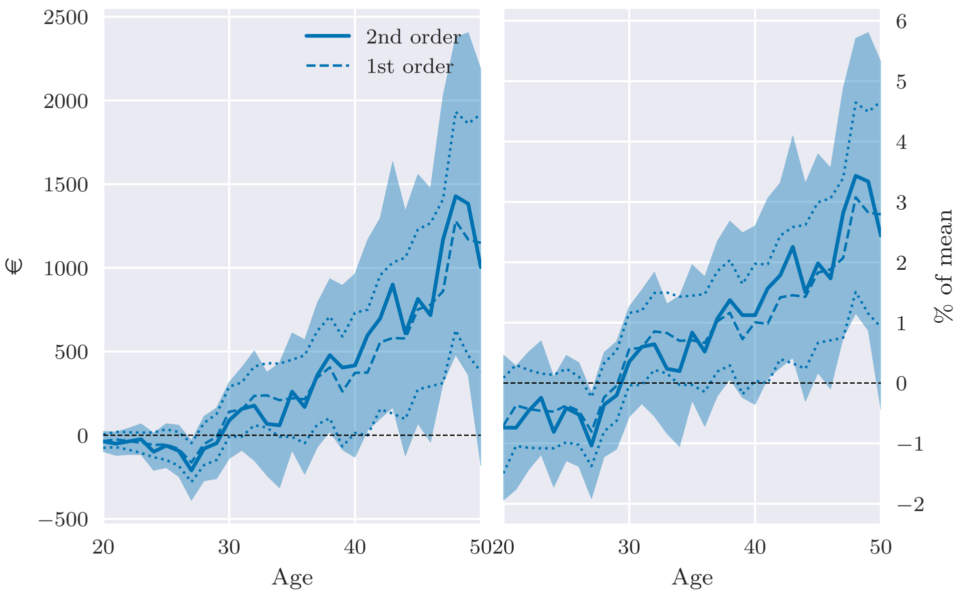

Figure 23: Earnings controlling the exam score with a second order polynomial

Notes: The solid line is the parameter estimates for exam years 1982-1987. The dashed line are the parameter estimates for the full sample.

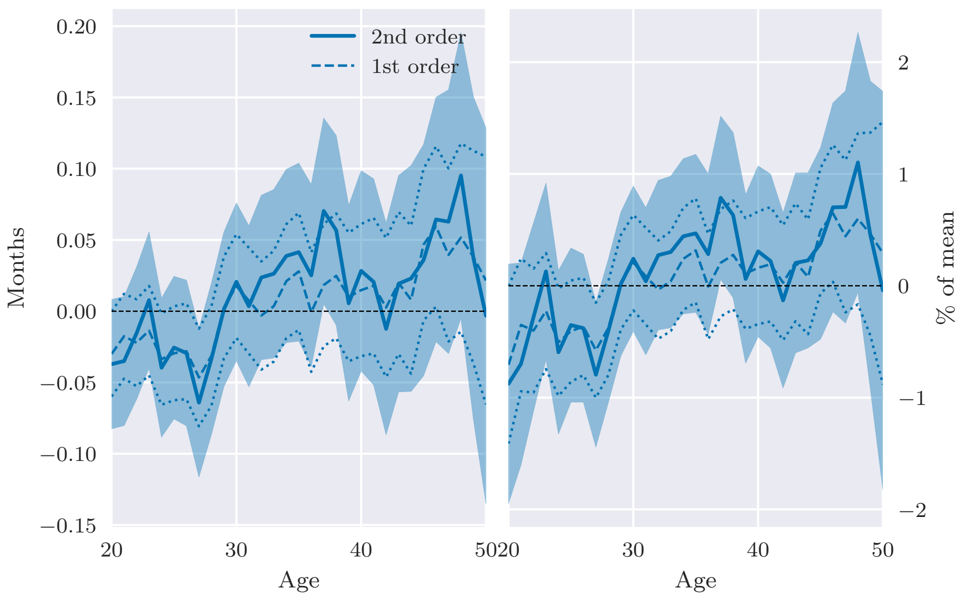

Figure 24: Employment controlling the exam score with a second order polynomial

Notes: The solid line is the parameter estimates for exam years 1982-1987. The dashed line are the parameter estimates for the full sample.

Figure 25: Degree median earnings controlling the exam score with a second order polynomial

Notes: The solid line is the parameter estimates for exam years 1982-1987. The dashed line are the parameter estimates for the full sample.

Table 7: Education outcomes controlling the exam score with a second order polynomial

Cutoff Percentile	5	16	36	60	80	95	Pooled
Years	13.00	13.41	13.82	14.26	14.70	15.01	14.10
Estimate	0.03	0.04	0.03	0.04	0.02	-0.00	0.03
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)	(0.01)
Relative effect (%)	0.19	0.27	0.23	0.32	0.17	-0.03	0.23
Placebo	0.06	0.03	-0.01	0.03	-0.04	0.02	0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.06)	(0.01)
Peers (€)	26759.88	27112.29	27583.11	28202.59	28968.33	29849.56	28102.56
Estimate	-22.96	-11.97	83.72	68.88	-0.06	69.14	42.52
	(43.82)	(33.07)	(31.27)	(33.31)	(38.49)	(76.51)	(16.23)
Relative effect (%)	-0.09	-0.04	0.30	0.24	-0.00	0.23	0.15
Placebo	26.13	13.09	10.19	27.80	-8.91	224.44	27.36
	(36.68)	(31.82)	(32.21)	(35.40)	(43.55)	(131.16)	(16.37)
University (pp)	4.23	8.64	15.79	26.03	38.34	50.00	24.21
Estimate	0.40	0.11	0.66	1.35	0.94	1.17	0.91
	(0.31)	(0.28)	(0.31)	(0.35)	(0.40)	(0.76)	(0.16)
Relative effect (%)	9.49	1.31	4.21	5.20	2.45	2.34	3.74
Placebo	0.34	0.38	0.16	0.62	-0.56	0.81	0.27
	(0.29)	(0.29)	(0.33)	(0.37)	(0.44)	(1.20)	(0.16)
Applied (pp)	35.38	43.80	44.21	39.88	31.78	20.26	37.88
Estimate	2.77	0.28	-0.38	-0.41	-0.49	-1.61	-0.26
	(0.73)	(0.50)	(0.42)	(0.39)	(0.38)	(0.60)	(0.19)
Relative effect (%)	7.83	0.63	-0.87	-1.02	-1.53	-7.94	-0.68
Placebo	0.99	0.13	-0.59	0.31	-0.09	-1.39	-0.10
	(0.58)	(0.45)	(0.40)	(0.38)	(0.38)	(0.80)	(0.19)
Vocational (pp)	29.38	30.64	24.73	17.74	11.54	6.61	19.80
Estimate	4.71	-1.21	-1.04	-0.74	-0.34	-0.11	-0.36
	(0.70)	(0.47)	(0.36)	(0.30)	(0.26)	(0.37)	(0.16)
Relative effect (%)	16.04	-3.95	-4.22	-4.15	-2.92	-1.61	-1.82
Placebo	-0.23	-0.32	-0.12	-0.34	0.02	0.33	-0.18
	(0.56)	(0.41)	(0.33)	(0.28)	(0.25)	(0.50)	(0.15)
STEM (pp)	12.98	15.51	16.04	16.25	16.63	17.85	16.07
Estimate	1.19	-0.27	0.80	0.11	-0.38	-0.19	0.17
	(0.52)	(0.37)	(0.31)	(0.29)	(0.31)	(0.59)	(0.15)
Relative effect (%)	9.15	-1.73	4.97	0.68	-2.26	-1.08	1.05
Placebo	0.28	0.11	0.17	0.29	0.26	1.59	0.26
	(0.42)	(0.33)	(0.30)	(0.29)	(0.33)	(0.99)	(0.15)

Notes: The five rows for each outcome report the mean of the outcome, the effect of being above the cutoff on the outcome (with its standard error in parenthesis) and the effect of placebo cutoffs. We run the regressions separately for each cutoff, where we pool over years and subjects. Here we control for the running variable with a second order polynomial. The results of these regressions are reported in the first six columns. In the last column we report results were we pool all cutoffs together.

F Results for exam years 1982-1987

Table 8: High school outcomes for exam years 1982-1987

Cutoff Percentile	5	16	36	60	80	Pooled
Grade	0.59	1.55	2.52	3.50	4.48	2.85
Estimate	0.96	0.99	0.99	1.00	1.00	0.99
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Placebo	0.00	0.00	-0.00	0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Final Grade	1.02	1.64	2.59	3.54	4.50	2.94
Estimate	0.07	0.85	0.85	0.89	0.95	0.81
	(0.02)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Placebo	0.00	-0.00	-0.00	-0.01	-0.00	-0.00
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Repeat	0.37	0.10	0.08	0.07	0.04	0.10
Estimate	-0.59	-0.08	-0.10	-0.10	-0.10	-0.14
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Placebo	0.00	0.00	-0.00	-0.00	0.00	-0.00
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
High School	0.84	0.97	0.99	0.99	1.00	0.97
Estimate	0.11	0.00	0.00	0.00	-0.00	0.01
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Placebo	0.01	0.00	-0.00	-0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Notes: The five rows for each outcome report the mean of the outcome, the effect of being above the cutoff on the outcome (with its standard error in parenthesis) and the effect of placebo cutoffs. We run the regressions separately for each cutoff, where we pool over years and subjects. The results of these regressions are reported in the first six columns. In the last column we report results were we pool all cutoffs together.

Figure 26: Earnings for exam years 1982-1987
Notes: The solid line is the parameter estimates for exam years 1982-1987. The dashed line are the parameter estimates for the full sample.

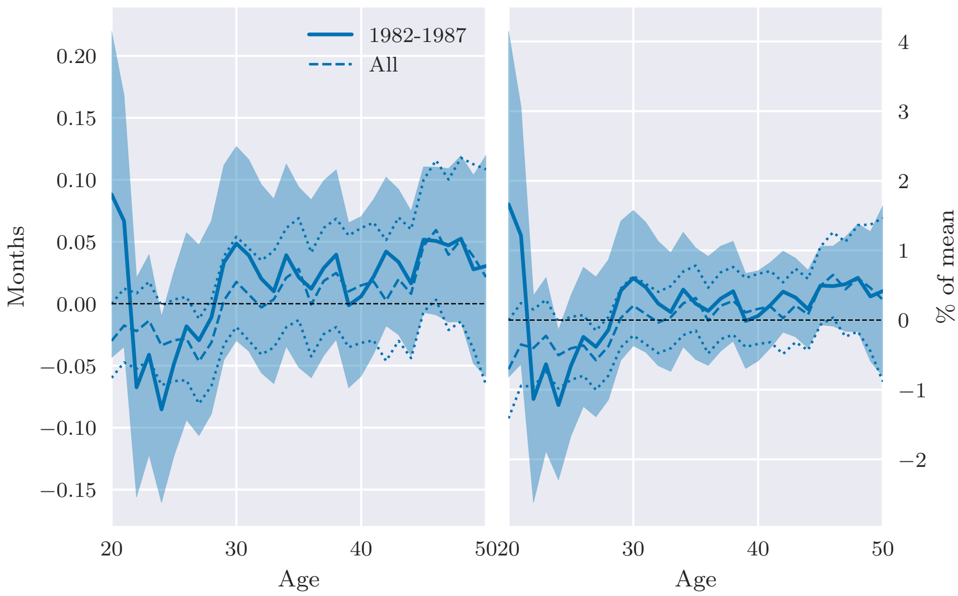

Figure 27: Employment for exam years 1982-1987
Notes: The solid line is the parameter estimates for exam years 1982-1987. The dashed line are the parameter estimates for the full sample.

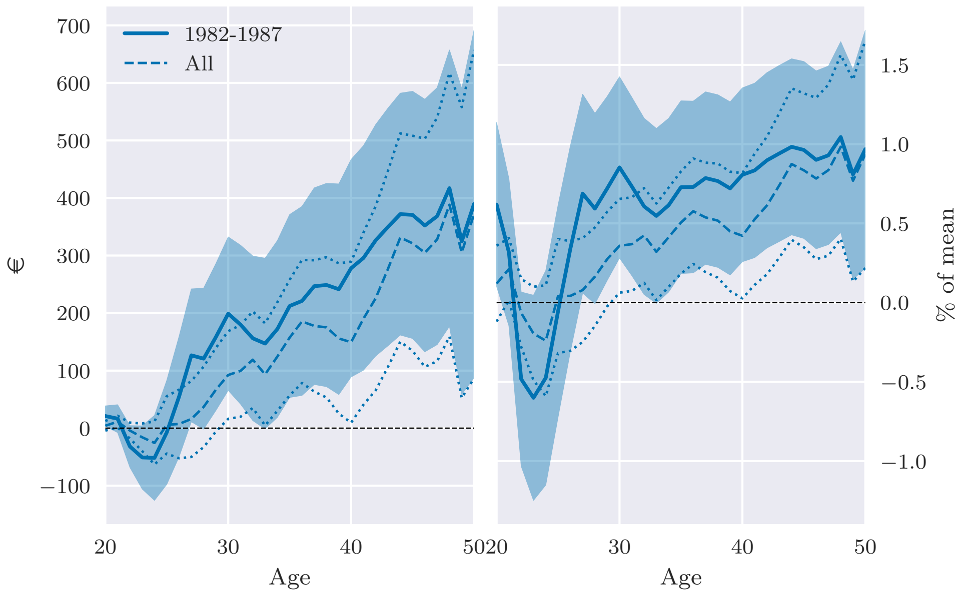

Figure 28: Degree median earnings for exam years 1982-1987

Notes: The solid line is the parameter estimates for exam years 1982-1987. The dashed line are the parameter estimates for the full sample.

Table 9: Education outcomes for exam years 1982-1987

Cutoff Percentile	5	16	36	60	80	Pooled
Years	12.94	13.24	13.63	14.13	14.67	13.86
Estimate	0.06	0.06	0.03	0.05	0.06	0.05
	(0.06)	(0.04)	(0.03)	(0.03)	(0.04)	(0.02)
Relative effect (%)	0.47	0.42	0.22	0.33	0.43	0.37
Placebo	0.07	0.04	-0.00	0.07	-0.01	0.03
	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.02)
Peers (€)	26487.09	26733.68	27122.65	27746.61	28616.07	27499.80
Estimate	-50.31	72.95	150.08	14.12	30.43	59.44
	(74.12)	(56.79)	(55.92)	(64.27)	(79.02)	(30.60)
Relative effect (%)	-0.19	0.27	0.55	0.05	0.11	0.22
Placebo	49.92	55.83	16.48	92.87	-105.97	29.16
	(62.79)	(55.47)	(59.76)	(70.70)	(90.64)	(30.96)
University (pp)	3.23	6.37	12.59	22.87	36.70	18.72
Estimate	0.47	0.58	1.45	1.47	1.07	1.16
	(0.49)	(0.44)	(0.53)	(0.65)	(0.81)	(0.29)
Relative effect (%)	14.43	9.05	11.55	6.43	2.91	6.19
Placebo	0.33	0.74	0.76	1.69	-0.46	0.75
	(0.44)	(0.48)	(0.59)	(0.73)	(0.90)	(0.30)
Applied (pp)	35.03	42.43	44.00	41.54	33.73	40.02
Estimate	3.07	0.19	-1.04	-0.94	0.15	-0.17
	(1.30)	(0.91)	(0.79)	(0.76)	(0.79)	(0.39)
Relative effect (%)	8.76	0.45	-2.37	-2.27	0.44	-0.42
Placebo	0.87	0.19	-1.19	-0.52	-0.07	-0.36
	(1.04)	(0.83)	(0.77)	(0.78)	(0.80)	(0.37)
Vocational (pp)	33.18	36.36	30.92	22.52	14.35	26.22
Estimate	5.63	-1.48	-1.68	-0.76	-0.91	-0.60
	(1.27)	(0.89)	(0.73)	(0.64)	(0.59)	(0.34)
Relative effect (%)	16.96	-4.06	-5.45	-3.37	-6.38	-2.27
Placebo	-0.20	-1.02	-0.42	-1.51	-0.74	-0.80
	(1.03)	(0.79)	(0.68)	(0.61)	(0.55)	(0.32)
STEM (pp)	11.94	14.31	14.84	15.13	15.16	14.65
Estimate	1.03	0.53	0.84	0.01	-0.58	0.28
	(0.87)	(0.65)	(0.56)	(0.56)	(0.60)	(0.28)
Relative effect (%)	8.64	3.72	5.69	0.04	-3.83	1.90
Placebo	-0.08	-0.17	0.61	0.51	-0.22	0.17
	(0.73)	(0.59)	(0.56)	(0.58)	(0.65)	(0.27)

Notes: The five rows for each outcome report the mean of the outcome, the effect of being above the cutoff on the outcome (with its standard error in parenthesis) and the effect of placebo cutoffs. We run the regressions separately for each cutoff, where we pool over years and subjects. The results of these regressions are reported in the first six columns. In the last column we report results were we pool all cutoffs together.