VATT Working Papers 114

Lost Boys: Access to Secondary Education and Crime

Kristiina Huttunen
Tuomas Pekkarinen
Roope Uusitalo
Hanna Virtanen

VATT INSTITUTE FOR ECONOMIC RESEARCH

VATT WORKING PAPERS

114

Lost Boys: Access to Secondary Education and Crime

Kristiina Huttunen
Tuomas Pekkarinen
Roope Uusitalo
Hanna Virtanen

Valtion taloudellinen tutkimuskeskus
VATT Institute for Economic Research
Helsinki 2018

Kristiina Huttunen, VATT Institute for Economic Research, Aalto University, Helsinki GSE and IZA, kristiina.huttunen@aalto.fi

Tuomas Pekkarinen, VATT Institute for Economic Research, Aalto University, Helsinki GSE and IZA, tuomas.pekkarinen@vatt.fi

Roope Uusitalo, University of Helsinki, Helsinki GSE and IZA, roope.uusitalo@helsinki.fi

Hanna Virtanen, Research Institute of the Finnish Economy, hanna.virtanen@etla.fi

Acknowledgements: We thank Randi Hjalmarsson, Steve Machin, Olivier Marie, Chris Neilson, Emily Nix, Imran Rasul, and Uta Schönberg, as well as seminar participants at CEP, HECER, ISER Essex, UCL, University of Nottingham, Uppsala University, the Nordic Summer Institute for Labour Economics, and SOLE Meetings, Toronto. Financial support from the Strategic Research Council of Academy of Finland (grant number: 293445) is gratefully acknowledged.

Corresponding author: Kristiina Huttunen, VATT Institute for Economic Research, Arkadiankatu 7, 00101 Helsinki, Finland, kristiina.huttunen@aalto.fi

ISBN 978-952-274-228-5 (PDF)

ISSN 1798-0291 (PDF)

Valtion taloudellinen tutkimuskeskus
VATT Institute for Economic Research
Arkadiankatu 7, 00100 Helsinki, Finland

Helsinki, December 2018

Lost Boys: Access to Secondary Education and Crime

VATT Institute for Economic Research
VATT Working Papers 114/2018

Kristiina Huttunen – Tuomas Pekkarinen – Roope Uusitalo – Hanna Virtanen

Abstract

We study the effect of post-compulsory education on crime by exploiting a regression discontinuity design generated by admission cut-offs to upper secondary schools in Finland. We combine data on school applications with data on criminal convictions and follow individuals for 10 years. Our results show that successful applicants are less likely to commit crimes during the first five years after admission. Crime is reduced both during and outside the school year, indicating that the channel through which schooling affects crime cannot be explained by incapacitation alone. We find no effect on crime committed after 6 years from admission.

Key words: crime, education, school admission, incapacitation, human capital

JEL classes: K42, I2

1. Introduction

It is well established that individuals with low levels of educational attainment are vastly over-represented among criminal offenders.[1] Over the recent decades, the empirical literature has also progressed in addressing the causality of this relationship. Several papers have examined the effects of changes in the amount of compulsory schooling on crime and reported that the extension of compulsory schooling typically reduces crime. However, still not much is known about the mechanisms through which education reduces crime. For example, it is still unclear to what extent attending more years of school reduces crime by simply incapacitating individuals through keeping them off the streets or by increasing the opportunity cost of crime through higher human capital.[2] Furthermore, there is a clear lack of evidence on the effects of participation in post-compulsory schooling on crime—even though this is a margin of educational attainment that is crucial for learning skills that are directly relevant in the labour market.

In this paper, we exploit data on post-compulsory school admissions and longitudinal crime records to compare the crime trajectories of individuals who succeed in gaining admission to post-compulsory education to crime trajectories of those who are rejected. The Finnish education system bases admissions to post-compulsory education programmes on the end-of-compulsory school grade point average (GPA) of the applicants. In the case of programmes that are over-subscribed, this admission mechanism generates a regression discontinuity design (RDD) that allows us to compare the outcomes of the individuals that were just above the admission cut-off with the outcomes of those that just failed to be accepted. By linking these admission data to administrative data on criminal convictions we can estimate the causal effect of being admitted to post-compulsory schooling on crime. Furthermore, the ability to follow individuals over time allows us to

[1] According to Harlow (2003), 68% of State prison inmates in the United States did not have a high school diploma in late 1990's. In Finland, 48 % of all offenders, and 75 % of those sentenced to prison, had no post-compulsory degrees in year 2011, while the corresponding figure among the same age population was 15 percent. (Aaltonen et al, 2011).
[2] Lochner (2011) provides a survey on the theoretical and the empirical literature on the effect of education on crime.

disentangle the immediate incapacitation effects of admittance from long-term human capital effects.

Previous studies on education and crime have mainly relied on variation in compulsory schooling age across jurisdictions and cohorts to identify the causal effect of the length of schooling on crime. Lochner and Moretti (2004), Machin et al. (2011) and Hjalmarsson et al. (2015) exploit variation generated by changes in compulsory schooling laws in the US, the UK and Sweden, respectively. The results in these papers provide evidence that educational attainment, measured in years of schooling, reduces crime. In the previous literature, attempts at isolating the incapacitation effect of education have either focused on the effects of compulsory schooling laws on juvenile crime (Anderson 2014; Beatton et al. 2016) or have exploited events such as temporary school closures to examine the contemporaneous relationship between schooling hours and juvenile crime (Jacob and Lefgren 2003; Luallen 2006). Landersø et al. (2015) and Cook and Kang (2016), on the other hand, find evidence on the incapacitation effect by exploiting variation in the school starting age.

However, to the best of our knowledge, very few papers have attempted to disentangle incapacitation and human capital channels within the same empirical setting.[3] This paper aims to fill this gap in the literature. We use an RD design to study whether failure to be admitted into post-compulsory education has a causal effect on criminal activity over the 10-year follow-up period. The key to our analysis is the Finnish registry data that allow us to link individual records from the school application registry with records on education, crime and labour market outcomes, over a period of several years. These data make it possible to disentangle the different mechanisms through which schooling may affect criminal activity. For example, we can investigate whether criminal behaviour is affected during school terms during the years individuals are still enrolled at

[3] In a recent paper, Bell et al (2018) study the effect of changes in compulsory schooling age in the United States and Australia on crimes at different ages and finds evidence that most of the effect of the compulsory schooling derives from incapacitation.

school or, alternatively, after they obtain a post-compulsory degree that increases the returns to legitimate work. Similarly, we can examine whether the effect of education varies across types of crime.

In addition to disentangling mechanisms, our paper also provides new evidence on the effect of post-compulsory education on crime. We argue that is an important margin of education since post-compulsory education typically provides individuals with skills that are of direct relevance in the labour market. We are aware of only three earlier papers that focus on the effect of post-compulsory education on crime. Two papers explore variation in the supply of post-compulsory education across cohorts and regions (Machin et al.2012, Brugård and Falch 2013). Both find that expansion of post-compulsory schooling is associated with reduction in the aggregate crime (or imprisonment) rate in the location. Åslund et al. (2018) examine a reform that extended the length of vocational education from two to three years in Sweden, and find evidence that property crime decreased during the additional enrolment year. Our paper differs from these papers in that we study the effect of post-compulsory schooling at the extensive margin and can follow individuals' crime patterns during and after the post-compulsory school enrolment. Moreover, we are able to identify the effect of gaining entry to post-compulsory education without using variation generated by reforms that probably affected the content of the post-compulsory schooling as well.

Our results indicate that admission to post-compulsory schooling has a negative effect on the probability of committing any kind of crime within five years after the admission. However, we find that the differences between criminal activity of successful and unsuccessful candidates decrease in the long run, indicating that the crime reducing effect of schooling is driven by the incapacitation effect of schooling. Yet, the negative effect of gaining admission to post-compulsory schooling on crime occurs during both school terms and breaks, suggesting that the incapacitating effect of schooling is more extensive than simply the physical presence at school. When distinguishing between different types of crimes, we find that the admission to upper secondary school mainly affects property, traffic and other crimes (drug-related) while having no effect on violent crime. All these effects are driven by the effect of education on criminal behaviour among men. We find no effects on criminal activity among women.

The remainder of the paper is organized as follows. Section 2 describes the school application system in Finland. Section 3 presents the data and discusses the definition of the running variable and the threshold. Section 4 presents the empirical strategy and provides a graphical presentation of our research design and the regression results. This section also discusses the potential mechanism. Section 5 concludes.

2. Upper secondary school admissions in Finland

The context of our analysis is the Finnish post-compulsory school application system. The length of compulsory schooling in Finland is nine years. Compulsory school ends in the summer of the calendar year when the student turns sixteen. After finishing compulsory education, students can apply to secondary schools. Secondary schools are divided into two tracks: an academic track which is general in nature and provides basis for access to tertiary education and a vocational track that prepares students for specific occupations, such as hair dresser, car mechanic etc. Figure A1 describes the structure of the Finnish education system.

In order to gain access to secondary school, students have to apply through a centralized application system maintained by the Finnish National Board of Education (FNBE). The application process is depicted in Figure 1. The process starts during the February-March period in the final 9th year of compulsory school. Individuals can apply to up to five different educational programmes (either programmes in different schools or different programmes within schools). The allocation of places in each programme is based on a programme-specific admission score. For most schools, this score is solely based on average grades from the last (9^{th}) year of compulsory school. Some schools give extra points for experience and minority gender, or use aptitude tests in addition to grades. In addition, the weights given to different criteria vary across schools and across programs. As can be seen from Figure 1, the applicants only receive their compulsory school GPA's in May –some two months after submitting applications. This means that applicants do not know their own admission points or the admission thresholds when applying. All this should make strategic application behaviour very difficult.

Student selection to each programme follows a DA algorithm where each applicant is considered for her preferred choice in the first round. Each programme tentatively accepts applicants according to its selection criteria up to its capacity and rejects lower-ranking students. In the next rounds, the applicants rejected in the previous round are considered for their next preferred programme. Each programme compares these applicants to the admitted applicants from previous rounds, rejecting the lowest-ranking students in excess of its capacity. The algorithm terminates when every applicant is matched to a track or every unmatched candidate is rejected by every track she had listed in her application. At the end of this automated admission stage, in June of the final year of compulsory school, the applicants receive an offer according to the allocation result. Admitted applicants have two weeks to accept the offer based on the automated admission process while rejected applicants are placed on a waiting list in rank order based on their admission score. After these two weeks, the schools start to fill the remaining vacant slots by calling the applicants in their waiting list in rank order.

Each year around 5 percent of students fail to obtain a place in upper secondary schools in Finland, although there are more slots than the number of compulsory school leavers. The reason for this is that older cohorts can also apply for upper secondary school places.[4] Typically, these older applicants have been accepted in previous years but wish to switch to another programme. Older applicants also include applicants who were rejected by the programmes they applied to in previous years. The main educational options for applicants not accepted to secondary education are the optional 10th grade in comprehensive school and preparatory vocational training. However, as the failed applicants have finished compulsory schooling they are under no obligation to study and can also try to find employment.

3. Data, Sample and Variables

3.1 Data

Our primary data set is the Finnish joint application registry which contains information on all the Finnish students that graduate from compulsory school and apply for secondary education. As nearly everyone applies, these data practically cover full cohorts of 16 year-olds in Finland. The data include information on grades in all subjects at the end of compulsory school (Mathematics, Finnish, English, History, etc.), grade point average, compulsory school ID, applications to secondary schools in preference ranking, programme codes for programmes applied to, admission scores (for each programme individual applies to), as well as information on whether the applicant was offered a slot in secondary school programmes, whether she was put on a waiting list and whether she accepted the slot she was offered. The data also contain a unique person identification code that allows us to link the data to other registers and follow individuals over time.

[4] Every year around 30-40 percent of all applicants to secondary schools had finished their compulsory schooling before the application year.

We focus on seven cohorts that graduated from compulsory school during 1996-2003 and applied to secondary education in the year they finished comprehensive school. We merge this information with data from several sources: from Employment Statistics that contain information on main activity (employment, unemployment, student etc.) and taxable earnings by source (wages, unemployment benefits, student grants, etc.); from the Register on Degrees and Examinations that contains the date and the type of completed degrees from all Finnish educational institutions; from the Student register that reports enrolment in education; and, finally, from data on prosecutions, sentences and punishments containing all convictions from criminal courts. We follow the target cohorts in these registers up to 2013.

The conviction data are based on all decisions by district courts. These data have information on the type of punishment (no conviction, fine, conditional imprisonment and unconditional imprisonment) and the type of crime (four-digit conviction classification). The data also contain information on both the timing of crime (date), and timing of conviction (year), and the principal crime for each conviction and the length of possible prison sentences. The main outcome measure for criminal activity, “Any conviction”, is an indicator for whether individual has been convicted for any crime in district court. In addition, we measure convictions by type of crime and sentence.

In Figure 2, we plot the crime-age profiles from the conviction data. The minimum age of criminal responsibility is 15 in Finland. The crime rates increase from age 16 and peak at ages 19 to 20 which is common to crime-age profiles reported in countless earlier studies. Property crimes are clearly the most common crime category. The share of traffic violations increases from age 18 which is the age when Finnish youth can obtain a driver’s license. At all ages, most convictions lead to fines or conditional imprisonment. Unconditional imprisonment is rare before age 19 but the severity of crimes and, therefore, the share of imprisonment also increases with age (Reported in Appendix Figure A2). As most Finnish students apply to upper secondary schooling the year that they turn 16 and most secondary school programmes last for three years, the students who enrol in secondary schools immediately after comprehensive school graduate at ages 19-20 when the crime rates are at their highest.

3.2 Definition of the threshold

As explained above, the admission system generates offers at several stages. The first offers are sent based on the automated DA algorithms and the applicants have two weeks to accept or reject these offers. The application register data reveal that sufficiently large numbers of offers are turned down at this stage so that there is no visible discontinuity based on the automated offers that we could exploit in our analysis. After the two weeks have passed from the initial offers, the schools start to fill vacant slots by calling those applicants who are on the waiting list (which is ranked according to admission points). In our data, the applicant is defined as being admitted after this round of calls and we define the admission cut-off score as the admission points of the last applicant admitted from the waiting list. Some admitted applicants, however, choose not to enroll in the programme and many schools have open slots/vacancies after the commencement of school year in August. Therefore, the enrolment does not correspond one-to-one with surpassing the admittance threshold in our data but we focus on enrolment as one outcome that is potentially affected by admittance.

3.3 Sample Construction

Our goal is to exploit the post-compulsory secondary education programme admission cut-offs to estimate the causal effect of school admission on crime. We have data on a total 476,476 first-time applicants to 17,047 programme-year combinations over seven years. In principle, each programme-year combination has an admission threshold that could be used in this analysis. However, in order to use the applications in our regression discontinuity design we need to restrict the number of programmes in four ways. First, in order to be able to measure the admission cut-offs using admission scores that are observable in our data, we only include programmes that rank their candidates on the basis of admission points in our analysis sample. This restriction excludes some special education programmes that use alternative criteria for admission. Second, we include only programmes that reject some applicants. This restriction excludes under-subscribed programmes from our analysis sample. Third, as we are interested in the effect of gaining entry to any secondary education, we concentrate on programmes that are critical at this margin in the sense that at least some of the applicants would remain without a slot in any secondary school if they fail to surpass the cut-off. In practice, we define these programmes as the ones that have the lowest threshold for at least some applicants (among the set of programmes that that they applied to). Finally, for our regression discontinuity strategy to work, we need a sufficient amount of applicants for whom the programme is critical and who are sufficiently close to the threshold. In what follows, we use two criteria for sufficient mass by looking at thresholds that have either one or five applicants within a unit of the admission scores at both sides of the entry threshold for whom that particular programme was critical in the sense that it was their best chance of gaining entry to post-compulsory secondary education.

The restrictions outlined above reduce our sample size significantly (as shown in table A1 in the Appendix). When focusing on programmes that are critical for at least one applicant within a unit of the admission score on both sides of the cut off, the number of programs drops from the total of 16,969 programme/year–combinations to 4,169. We call this set of programmes Sample 1. When we further restrict the sample to programmes with at least five candidates for whom the programme is critical within a unit of the admission score on both sides of cut off, the number of programmes drops to 669. These programmes had 62,158 applicants between 1996 and 2003. This sample is our preferred analysis sample (Sample 2). In this sample, 6,737 applicants were rejected from all programs they applied to.

In the Appendix Table A1, we provide a detailed comparison between the tracks and the applicants used in the analysis, with the full population of tracks and applicants. Panel A in Table A1 reveals that the tracks in the analysis sample are relatively large and concentrated in bigger cities. The analysis sample tracks are also more likely to be academic high school tracks (because these are on average larger) than all tracks in the full data. The minimum GPA for the admitted applicants in the analysis sample tracks is slightly lower than in the other tracks. In Panel B of Table A1, we compare the accepted and rejected candidates in these tracks. The rejected candidates have obviously lower GPAs, are more likely to be foreign born and more likely to live in large cities and have a less favourable family background.

Table A1 also reports the mean values of our main outcomes of interest. Admitted applicants are less likely to commit crimes both in the full population, 0.05 versus 0.12 within three years of the admission, as well as in our analysis samples, 0.04 versus 0.12 in sample 1 and 0.03 versus 0.10 in sample 2. The baseline differences in the probability of committing crimes are relatively similar across the two analysis samples and the full population. The difference in the crime rates remains and gets even large by the tenth year since admission.

3.4 Running variable

In our regression discontinuity framework, we use admission scores as the running variable. While the main determinant of the admission scores in all programmes is GPA from compulsory schooling, different schools apply different scales, give different weights to different grades and some use also other criteria in addition to GPA. To make the running variable comparable across different educational programmes, we rescale the admission scores to GPA units.[5] The programme-specific cut-off scores are defined on the basis of the lowest-scoring candidate that was observed being offered place to this programme. We drop the observations that are used to define the cut-off score from the analysis. The running variable, r_{ikt}, for applicant *i* to a track *k* in a year *t* is defined as her distance to the cut-off point τ_{kt} in GPA units.

$$r_{ikt} = (c_{ikt} - \tau_{kt})$$

These programme-specific running variables equal zero at the cut-off point for each programme in a given year. Figure 3 displays the distribution of this standardized running variable in our estimation sample. As shown in Figure 3, the cut-off scores are on the left tail of the distribution, but the distribution is smooth around the threshold. Table A2 shows that there is no discontinuity in the background characteristics of applicants at the cut-off.

[5] In practice we estimate programme-specific regression models where admission scores are explained with the GPA and then divide the score with the coefficient of GPA. This way, a one unit change in GPA has the same effect on the rescaled scores in each programme.

4. Empirical Strategy and Results

4.1. Graphical presentation of the reduced-form effects

Before turning to the regression results, we begin with a graphical presentation of the reduce-form impact of receiving an offer to post-compulsory schooling. The upper panel of Figure 4 plots offers and enrolments to post-compulsory education as a function of the standardized admission scores using sample 2. Each dot in the figure corresponds to the fraction admitted or enrolled within a bin of 0.1 GPA units.

In the upper left-hand side panel of Figure 4, we plot the fraction admitted applicants. As the admission threshold is defined based on the lowest-scoring admitted candidate, there are no admissions below the cut-off point. At the cut-off point, the share admitted jumps to 0.75. There is a clear discontinuity but the admission rate still remains well below 100%. Technically, this implies that some applicants on the waiting list did not receive an offer even though a person with a lower score in the waiting list was admitted. This could be due to some applicants not being reached at the point when schools fill their vacancies from the waiting list or simply due to measurement error in admission scores. For our analysis, this implies that the admission cut-offs do not work as a sharp regression discontinuity design for admissions but rather that passing the admission cut-off discontinuously increases the probability of being admitted to post-compulsory secondary school. In what follows, we estimate both the reduced-form estimates of passing the admission threshold on our outcomes of interest as well as the local average treatment effect (LATE) of being admitted where admissions are instrumented with being above the threshold.

The figure in the upper right-hand panel of Figure 4 plots the enrolment in post-compulsory education as a function of admission scores. The enrolment figure differs from admission mainly in the part below the cut-off scores. Many applicants that are not admitted in the admission process apply to schools directly during the months after the process has ended and are eventually enrolled in secondary education by the end of the year when they finish compulsory school. However, there is still a clear discontinuity, with the enrolment rates in post-compulsory education increasing from about 0.4 to about 0.8 at the cut-off score.

The figure in the lower left-hand panel plots the share of those admitted to any education, including both 10th grade and post-compulsory education. Most applicants that do not get into post-compulsory education are placed in voluntary 10th grade in compulsory school or in preparatory vocational education and are hence classified as being in education. The discontinuity at the cut-off point is still visible but the fraction in education only increases by about 15% at the cut-off. The lower right-hand panel of Figure 4 shows the share that has completed a post-compulsory degree within 10 years of finishing compulsory school. There is a clear jump in the probability of obtaining a degree in 10 years at the admission cut-off suggesting that successful admission also has long-term consequences for final educational attainment, even though the rejected individuals can apply again in the following years.

In Figure 5, we plot the propensity to commit any crime that leads to conviction in a district court within three and ten years after post-compulsory school admission as a function of the standardized admission score. The results show that there is a visible discontinuity in the probability of conviction at the admission cut-off. Those who exceed the post-compulsory school admission threshold are less likely to be convicted during the 3 or 10 year follow-up after the possible school admission than those who fail to exceed the admission cut-off.

Taken together, the descriptive evidence in Figures 4 and 5 suggests that receiving an offer to post-compulsory schooling may significantly increase enrolment in secondary education and reduce criminal behaviour. Overall, these descriptive figures shed light on the possibility that getting an offer to a post-compulsory education programme can affect criminal behaviour —both through incapacitation and keeping the youth off the street, and by increasing human capital.

4.2. Regression Analysis

To examine how admission to upper secondary school affects an individual's crime and labor market outcomes, we use the Sample 2 data in a RDD regression framework that exploits the admission cut-offs to identify the effect of post-compulsory education on our outcomes of interest. The reduced form of interest is:

$$y_{ikt} = \alpha_{kt} + \rho Z_{ikt} + (1 - Z_{ikt}) f_0(r_{ikt}) + Z_{ikt} f_0(r_{ikt}) + e_{ikt} \qquad (1)$$

where y_{ikt} is the outcome variable (e.g. admission, enrollment, completed degree, criminal conviction) for applicant i to track k in year t. Z_{ikt} is a dummy variable indicating whether the applicant is above the cut-off to track k in year t, and r_{ikt} is the running variable that is centered around the cut-off point (gets value 0 at cut-off). The effects of the running variable are controlled by a linear function, $f_0(r_{ikt})$. The equation includes an interaction term that allows the effect of the running variable to differ on different sides of the cut-off. We pool data on all programs and therefore have 569 indicator variables and their interactions with the running variable in our regressions. e_{ikt} summarizes the unobserved factors.

We estimate the equation (1) using nonparametric local linear regression with triangular kernel weights and optimal bandwidth for each programme-year combination derived using the selection procedure in Calonico et. al. (2014). In addition to our reduced-form specification, we also use our setting as a fuzzy RD design to estimate the local average treatment effect (LATE) of being admitted to post-compulsory education. We define the treatment variable for these regression, D_i, to indicate that an applicant is *observed* to receive an offer in the data. The first stage regression of this fuzzy RD strategy is the specification (1) where the outcome variable is admission.

4.3 Main Results

Table 1 reports the main results of our regression analysis. We begin by reporting the results for all school applicants in Panel A and then split the sample by gender in Panel B (Males) and Panel C (Females). We estimate the effect of admission on enrolment to secondary school in the fall after the admission and on the probability of at least one conviction by one, three, five and ten years after the admission.

The first row of Table 1 reports the reduced-form effect of exceeding the admission threshold on enrolment and crime. In the second row we report the first-stage estimate, i.e. the effect of exceeding the admission threshold on receiving an offer from at least one secondary program. In the third row, we report the local average treatment effects (LATE), measuring the effect of an offer on the outcome in each column. In the last row, we report the mean of each outcome variable below the admission threshold within the optimal bandwidth.

As the first stage results on the second row of Table 1 show, scoring above the cut-off score naturally has a large effect on the likelihood of receiving an offer and the first stage estimates are thus large and significant. Admission also has a strong effect on the enrolment in post-compulsory education in the next fall reported in column (1) of Table 1. It increases by 52 percentage points for male applicants and 66 percentage points for female applicants.

The effects on crime are reported in columns (2)-(5). The result for the whole sample in column (5) of Panel A suggests that those above the admission threshold are less likely to commit crimes over the whole follow-up period. Although we see no effect on crimes committed during the first school year immediately following admission, in column (2), the difference starts to show up within three to five years in columns (3) and (4). When splitting the sample by gender (Panels B and C), we find that the effect on crime is totally driven by male applicants. In general, the crime rates are much lower for the girls and admission to secondary education has no significant effect on any of our measures of crime for females. For male applicants, on the other hand, the admission clearly decreases the propensity to commit crime both in the medium and long-term. The IV (LATE) effect indicates that being admitted to upper secondary schooling decreases the probability of committing any type of crime by 7 percentage points within three years of admission, which corresponds to 53% reduction when compared to the mean below the cut-off (0.127). Within 5 years after admission, the reduction has increased to 12 percentage points (57 % when compared with the mean level below the cut-off, 20 %). The total crime during the whole follow-up period has decreased by 10.5 percentage points (39%).

4.4 The timing of crime effects

The results in Table 1 clearly indicate that being admitted to post-compulsory education has a negative effect on criminal activity among boys. However, the pattern of results in Table 1 is consistent with both the incapacitating effect of education as well as with the human capital enhancing effect. One way to disentangle these mechanisms from each other is to look at the timing of the effects of admission on crime more closely. If gaining admission reduces crime primarily during the years when the admitted applicants are most likely to be enrolled, we would argue that admissions reduce crime mainly through the incapacitating effects. However, if admissions lead to reductions in crime that only emerge over the long run, the results would be more in line with the human capital mechanism.

Figure 6 plots the estimates of the effect of admissions on crime by each school year (starting August 15th) from the pre-admission year until 10 years after the admission. In order to make sure that we follow same individuals throughout these ten years, we use a fixed bandwidth instead of track- and year-specific optimal bandwidths in these regressions. Reassuringly for our identification strategy, the results indicate no pre-admission difference in the criminal behaviour among boys. Neither do we see any clear impact on crime during the first post-admission school year. However, the gap between the annual crime rates of the admitted and non-admitted boys starts to increase during the second and third years after admission and remains below zero until year 5, although these annual estimates are quite imprecise.

Figure 7, on the other hand, plots the effect of admission on the probability of having at least one conviction as a function of the number of years from the post-compulsory school admission. The negative effect of admission on this cumulative crime outcome starts to emerge during the time when the admitted individuals are still likely to be enrolled and it continues to increase after this. However, in the long run the difference in the probability of at least one conviction between the admitted and non-admitted male applicants disappears.

Taken together, the results in Figures 6 and 7 are consistent with the incapacitating effect of schooling on crime. However, this incapacitation effect seems to be dynamic in nature and matches findings by Bell et al. (2018), who document similar effects using variation generated by a compulsory schooling extension in Australia. This kind of dynamic incapacitation keeps individuals enrolled during the time when they are most likely to start criminal careers and this effect persists also after finishing post-compulsory education since the lack of a criminal record increases relative returns to legitimate work. However, the results in Figure 6 and 7 do not give much support for the human capital channel as we fail to find enduring effects of schooling on crime.

4.5 The effects by types of crime

The effects of post-compulsory admission on different types of crime can also shed light on the mechanisms through which it reduces crime. Previous literature (e.g. Jacob and Lefgren 2003) has documented that the incapacitating effect of schooling on crime varies by types of crime. Adolescents are less likely to commit property crimes when in school, while violent crime may even increase during school hours. Thus, if we expect schooling to reduce crime through incapacitation, we would expect to see bigger reductions in property crime than in violent crime.

In Figure 8 and in Table 2 we examine separately the effects on crimes by crime type with fixed (Figure 8) and optimal bandwidths (Table 2). Figure 8 shows that the admission to post-compulsory education actually increases violent crime slightly during the first school year. This is in line with the results in Jacob and Lefgren (2003) who found that interactions at school may increase violent crime. At the same time, however, we see that admission decreases the probability of committing property, traffic and other crimes (including for example drug related offenses) from year 2 until year 5 after admission. After year 6, the difference in the cumulative crime rates between admitted and non-admitted remains quite stable. This pattern of the effects by crime types is again more consistent with the incapacitation than with the human capital channel. The results with optimal bandwidth (Table 2) are similar.

4.6 The effects within school year

To further examine the incapacitation channel, we investigate the timing of criminal activity more carefully. In Table 3 we report the effect of being above the admission threshold on the probability of committing a crime during the school year (from August 15th to May 30th) and during summer holiday breaks (from June 1st to August 14th). As previously, we find no clear impact on the propensity to commit crime during the first year after admission. By the third year, however, there is a clear reduction in crime committed during the school term as well as during the summer holidays. These results indicate that the contemporaneous effect of schooling on crime is more complex than just incapacitation channel. Those admitted to upper secondary school may be more active (in employment) during the summer holiday breaks as well, or simply interact with less crime-prone peers.

4.7 The effects on inactivity and educational attainment

Our data also allows us to examine the effect of admission on male inactivity and the likelihood of eventually completing a degree in Table 4. The results in Table 4 indicate that admission to post-compulsory education decreases the probability that a male applicant is not in education, nor in employment (NEET) in the following autumn by 14 percentage point. This is a big decrease (94 %) when compared with the baseline inactivity (15%). For females, there is a similar decrease in percentage points, although the baseline is higher (19%).

In Figure 9, we report the results on the effect of gaining admission on NEET in different years since possible admission. Gaining admission to post-compulsory education has a negative effect on the probability of being inactive during the first three years after the admission. However, after this the effect disappears. This pattern of results suggests that schooling may reduce crime among boys by keeping them active. Interestingly, we find no clear impact on crime rates during the first year after admission, although those not admitted are significantly more likely to be inactive. It may be that inactivity may not result in an immediate change in criminal behaviour and that the incapacitation effects of schooling may occur with a lag.

Although the timing of the crime effects, the effects by crime types and the effects on the inactivity are broadly consistent with the incapacitation mechanism, it is still possible that our failure to find evidence on the human capital channel is because gaining admission does not have any effect on the final educational attainment of the applicants. To rule out this possibility, we present the estimates of the effect of admission on the probability of graduation from post-compulsory education by three and ten years from admission, respectively, in columns (2) and (3) of Table 4. The results indicate that admission clearly increases the probability of obtaining a post-compulsory degree by the third year since admission for both boys (20 percentage points) and girls (29 percentage points), especially when compared to the baseline- which is only around 12 percentage. When comparing the probability of having obtained the degree by year 10 since admission, the difference diminishes, but remains significant. The results indicate that admission affects the timing of obtaining a degree for both genders, but also has an impact on completed education for males. Hence, admission should increase human capital even though we fail to find evidence for the human capital channel of crime reduction.

4.8 Robustness

Robustness with respect to control variables

To check that our results are not driven by some observable differences the admitted and non-admitted applicants around the entry threshold, Table A3 presents the results from the same set of regressions as in Table 1, with control variables for mother tongue of the applicant, age at graduation, gender, parental education and for living in one of the 15 largest cities in Finland. The results are robust to inclusion of controls.

Robustness to bandwidth choice and sample

One potentially disturbing pattern in the results above is the fact that the samples dictated by the outcome specific bandwidths vary a lot in size. In order to check how sensitive our results are to the choice of samples, we report with a fixed bandwidth in Table A4. The main results remain robust: being above the admission cut-off increases the probability of obtaining a degree and reduces crime. The clearest change is to the crime outcome in year 10: the sample is now much bigger and the effect on crime by year 10 does decrease. To further investigate this, we also report how sensitive our results are to choice of bandwidth. The Figure A3 reports the effect on cumulative crime by year 5 and 10, and cumulative property and traffic crimes by year 5, for males using different bandwidth choices. Our results indicate that our results are not sensitive to choice of bandwidth as long as the bandwidth does not exceed 0.5. The effects do get smaller when the bandwidth size gets very large.

Alternative sample design

In order to make sure we have enough observations around each cut-off point, we have used quite restrictive sample definitions and, as a result, have lost a big share of our data (as shown in table A1). In order to check how sensitive our results are to these restrictions, we use another sample where we restrict the data to applicants to programmes with at least 1 applicant for whom the program was his or her best chance to get into any program on both sides of the admission cut off. In these regressions we pool the data, and use same specification for all programs, since we are no longer able to use program specific running variables. Results are reported in table A5. The results are in line with our main analysis: Being above the admission threshold significantly reduced the propensity to commit crime, while increasing the propensity to finish a degree and being inactive. Again, the effect on cumulative crime is more robust.

5. Conclusions

While a large literature has documented the relationship between crime and education, we still know little about the mechanism how education influences criminal behavior. Two channels have been emphasized in the previous literature: the incapacitation impact of schooling, and the human capital accumulation which makes the opportunity cost of crime and punishments more costly. In this study, we aim to understand these channels further by exploiting the admission cut-offs to post -compulsory schooling in Finland to estimate the long-term effects school entry on crime. We follow individuals near the admission cut-offs for several years after the admission date, to understand whether the effect of school entry on criminal behaviour occurs during the time when individuals are still in school, or after they have obtained formal degrees and entered the labour market.

Our results show that being successful in upper secondary school admission decreases the likelihood of committing crimes within five years of the admission by 12 percentage points (60 % decrease in crime propensity) among boys. The effect is sizeable when compared with previous estimates that have exploited variation in the length of schooling generated by changes in compulsory schooling.[6] The effects on cumulative crime are still visible ten years after finishing compulsory school but these effects are driven by effects that take place relatively soon after finishing comprehensive school. The effects of education on crime do partly coincide with the effects of education on non-participation rates, suggesting that the incapacitation channel is one possible mechanism. The effect is driven by reduction in property, traffic and other offences (drug-related), with no impact on violent crimes, which is in line with previous findings that the incapacitation effect of schooling mainly reduces property crime. However, when investigating the timing of crime in detail, we find that there is a similar reduction in crimes committed during holiday breaks as well. In addition, we find that difference in crime rates between admitted and non-admitted do not show up immediately, but within 2 to 3 years from admission. We find no effects for crime propensity 6-10 years after school admission, indicating that the school entry had no long-lasting effect on criminal behaviour.

[6] Jakob and Lechner (2001) find that juvenile property crime decreases by 14 percent on days when school is in session. Lochner and Moretti (2004) find that a one-year increase in average education levels in a state reduces state-level arrest rates by 11 percent or more. Hjalmarsson et al. 2015 find that one additional year of compulsory schooling leads to a 6.7.% reduction in convictions using Swedish compulsory school reform and data. A more recent paper by Bell et al. 2018, find a 6 % reduction in crime arrest rates in the US following different compulsory school reforms, and an 11 % reduction in Australia.

The results indicate that while the immediate incapacitation channel of schooling may be important, the effect of schooling on crime may work through other channels as well. There may be several mechanisms in play that explain why school entry may reduce crime with delay: Individuals who do not enrol in school may be more likely to interact with more crime-prone peers and it may take time before inactive juveniles start engaging in criminal activity.[7] The delayed effects are also in line with the dynamic incapacitation effects of schooling (Bell et al. 2018). Individuals who are enrolled at school at a critical age may have a lower propensity to commit crimes in the future as well. It is important to keep in mind that our focus has been on young individuals that are just at the beginning of the years when crime rates are highest; most are 16 years old when applying to upper secondary schooling. Being unsuccessful in school entry during this critical age may have severe consequences on these adolescents, not only through directly keeping them active and off the street, but also by offering different peer groups and future prospects, that may all have an impact on criminal behaviour.

[7] Several papers have documented peer effects in criminal behaviour (see e.g. Kling et al. 2005, Bayer et al. 2009, Billings et al. 2014, 2018).

Literature

Aaltonen, M, Kivivuori, J., and Martikainen, P., (2011): "Social determinants of crime in a welfare state: Do they still matter?" *Acta Sociologica*, 54 (2): 161-181

Anderson, D. M. (2014): "In school and out of trouble? The minimum dropout age and juvenile crime", *The Review of Economics and Statistics*, 96 (2), 318-331.

Åslund, O, H Grönqvist, C Hall and J Vlachos (2018): "Education and Criminal Behavior: Insights from an Expansion of Upper Secondary School", Labour Economics, ISSN: 0927-5371, Vol: 52, Page: 178-192

Bayer, P., R. Hjalmarsson, and D. Pozen, (2009): "Building Criminal Capital behind Bars: "Peer Effects in Juvenile Corrections," *The Quarterly Journal of Economics*, 124 (2009), 105–147.

Beatton, T., Kidd, M. P., Machin, S., and Sarkar, D. (2016): "Larrikin youth: New evidence on crime and schooling", CEP Discussion Paper 1456.

Bell, B., Costa, R., and S. Machin, (2018): "Why does education reduce crime?", IZA Discussion Paper No 11805.

Billings, S. B., Deming, D. J., & Rockoff, J. (2014): "School Segregation, Educational Attainment and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg". *The Quarterly Journal of Economics*, 129(1), 435-476.

Billings S, Deming DJ, Ross S. (2018): "Partners in Crime". *American Economic Journal: Applied Economics*. Forthcoming

Brugård, K. H. and T. Falch, (2013): "Post-compulsory education and imprisonment", *Labour Economics*, 23, 97-106.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014): "Robust nonparametric confidence intervals for regression discontinuity designs", *Econometrica*, 82(6), 2295-2326.

Cook, P. J. and S. Kang. (2016): "Birthdays, Schooling, and Crime: Regression-Discontinuity Analysis of School Performance, Delinquency, Dropout, and Crime Initiation." *American Economic Journal: Applied Economics*, 8(1): 33-57.

Jacob, B., and L. Lefgren, (2003): ‘‘Are Idle Hands the Devil’s Work- shop? Incapacitation, Concentration, and Juvenile Crime,’’ *American Economic Review* 93, 1560–1577.

Hjalmarsson, R., Holmlund, H., and Linquist, M. J. (2014): “The effect of education on criminal convictions and incarceration: Causal evidence from micro-data”, *The Economic Journal*, 125, 1290-1326.

Kling, J., J. Ludwig and L. Katz. (2005): “Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment.” *The Quarterly Journal of Economics* 120 (1): 87-130.

Landersø, R., Nielsen, H. S. and Simonsen, M. (2016): “School Starting Age and the Crime-age Profile”. *Economic Journal*. doi:10.1111/ecoj.12325

Lochner, L. and Moretti, E. (2004): “The effect of education on crime: evidence from prison inmates, arrests and self-reports”, *American Economic Review*, vol. 94(1), pp. 155–89.

Lochner, L. (2011): “Non-Production Benefits of Education: Crime, Health and Good Citizenship”, in: E. Hanushek, S. Machin and L. Woessmann (eds.), *Handbook of the Economics of Education,* Volume 4, North Holland, Amsterdam, pp. 183–282.

Luallen, J. (2006): “School’s Out . . . Forever: A Study of Juvenile Crime, At-Risk Youths and Teacher Strikes”, ‘*Journal of Urban Economics* 59 (2006), 75–103.

Machin, S., Marie, O., and Vujic, S., (2011): “The Crime Reducing Effect of Education.” *Economic Journal* 121 (552), 463–484.

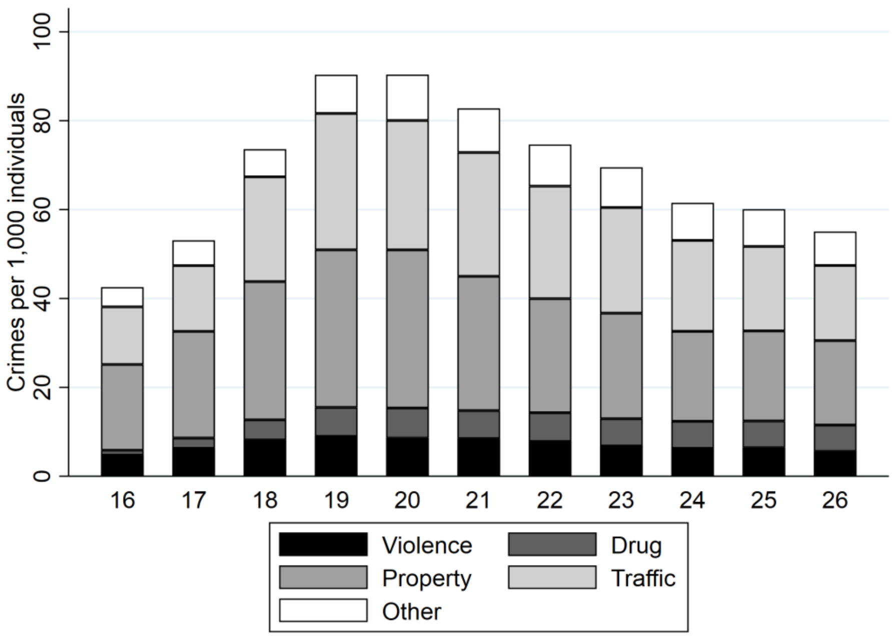
Machin, S., Marie, O., and Vujic, S., (2012): “Youth Crime and Education Expansion.” *German Economic Review*, vol. 13(4), pages 366-384, November.

Figures

Figure 1 Application process timeline

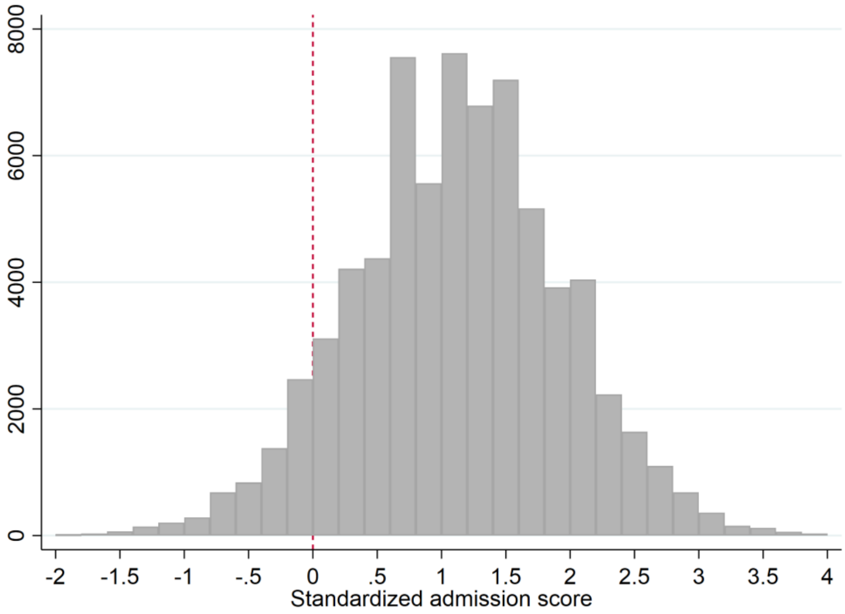
Timeline	
February-March	Students submit applications to post-compulsory education
March 12	Application deadline
May 30	Students receive their compulsory school GPA's
June 13	Students receive the offers based on the automated mechanism
June 27	Deadline for accepting the offers based on the automated mechanism
July - August	Empty slots are filled from the waiting list
Second week of August	Final enrolment decisions are made

Figure 2 Crime-age profile by crime types



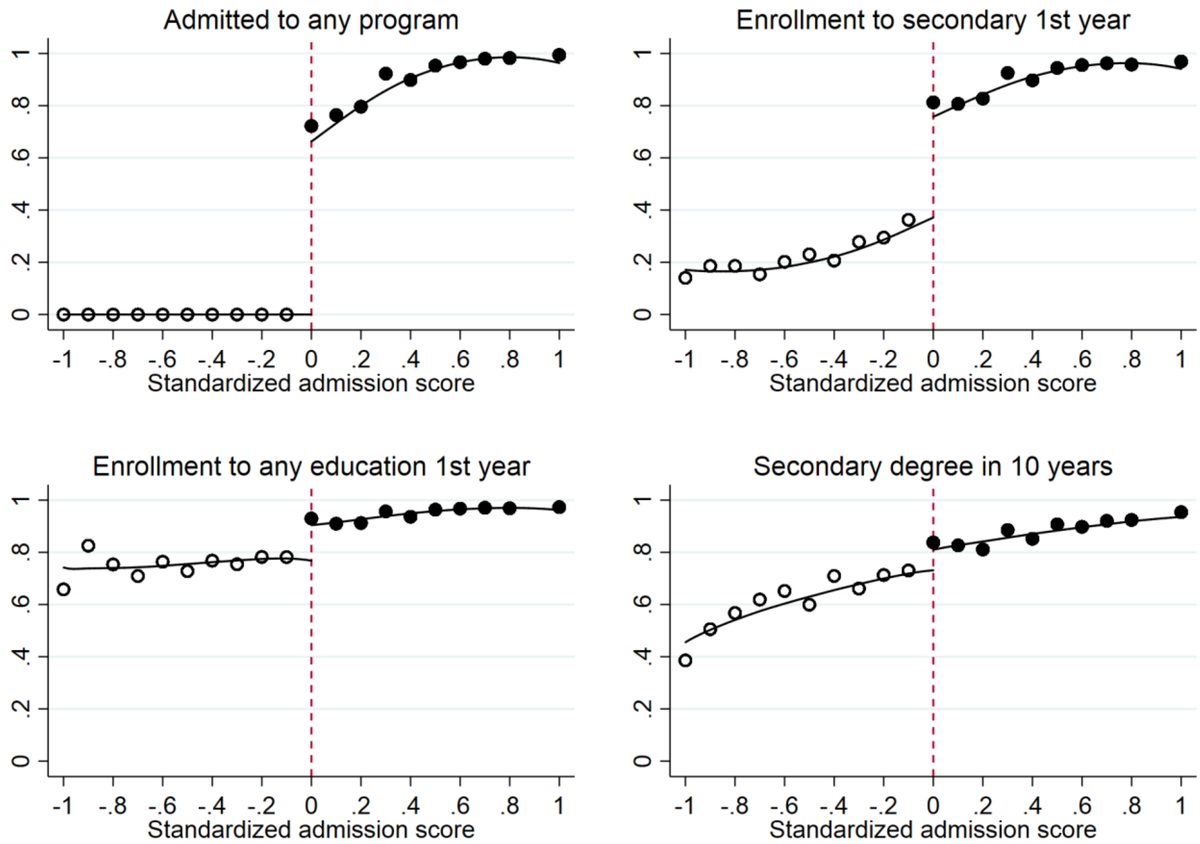
Notes: Crimes convicted at court by age and crime type. On the spot fines by police are excluded from the figure.

Figure 3 Distribution of observations with respect to cut-off grade



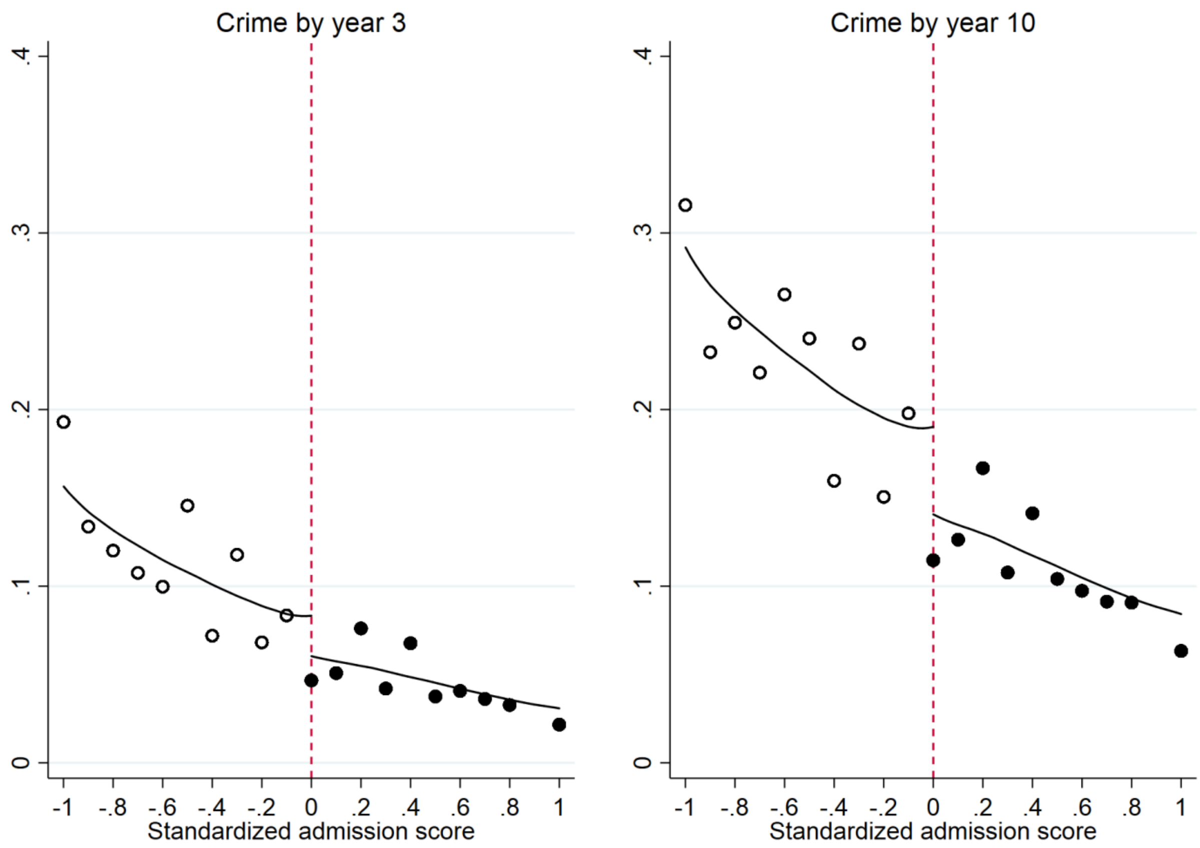
Notes: Distance to cut-off score in the program with lowest admission threshold among the applicant's applications. Admission scores are re-scaled so that one unit change of the GPA has the same effect on the standardized score in all programmes.

Figure 4 The effect of exceeding admission threshold on admissions, enrollment and education



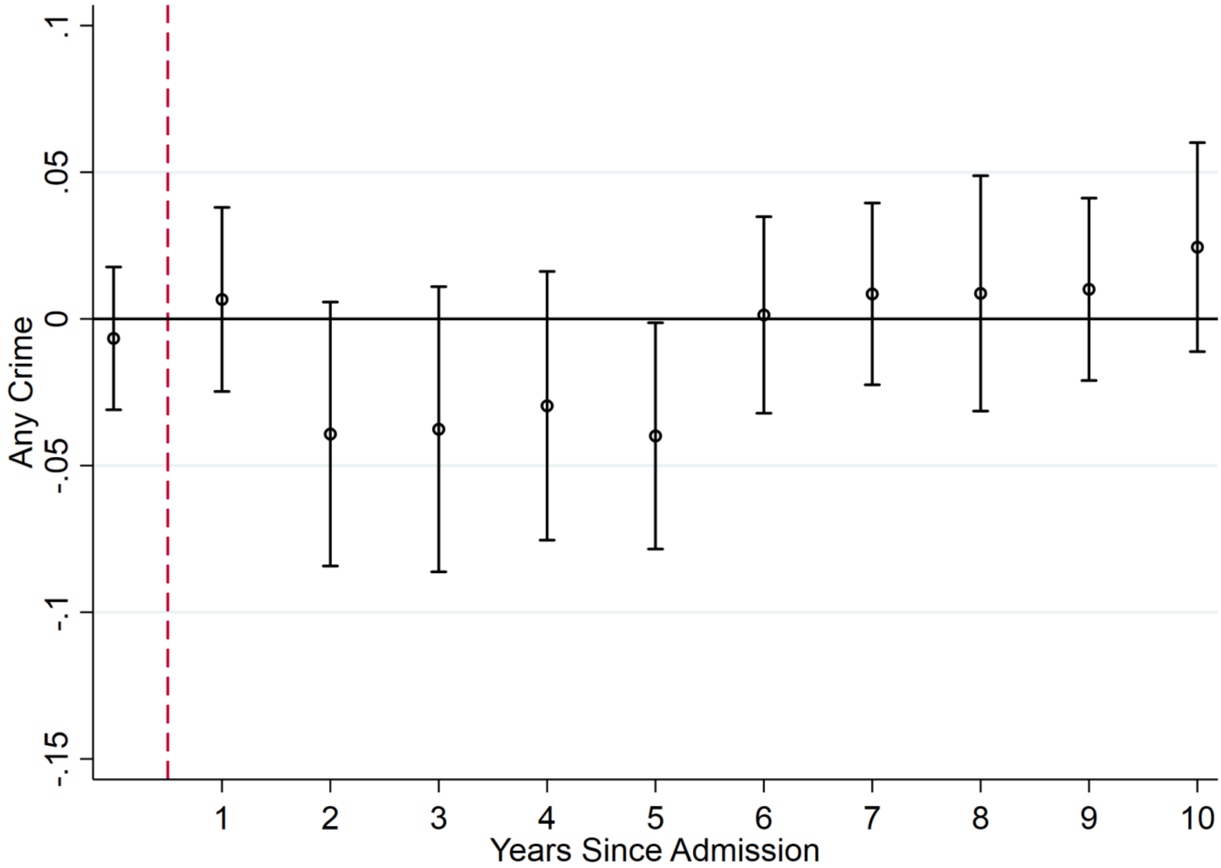
Notes: The figure shows the share of those admitted to any post-compulsory programme (left) and share of those enrolled to any post-compulsory program by next fall (left) plotted against track-specific standardized running variable. The sample is all programs with at least 5 applicants on both side of the cut off for which the program was the best chance of gaining admission (sample 2). The lines represent local second order polynomial estimates using the edge kernel and the optimal IK bandwidths. The dots correspond to the sample means by 0.1 GPA point bins.

Figure 5 The effect of exceeding admission threshold on crime



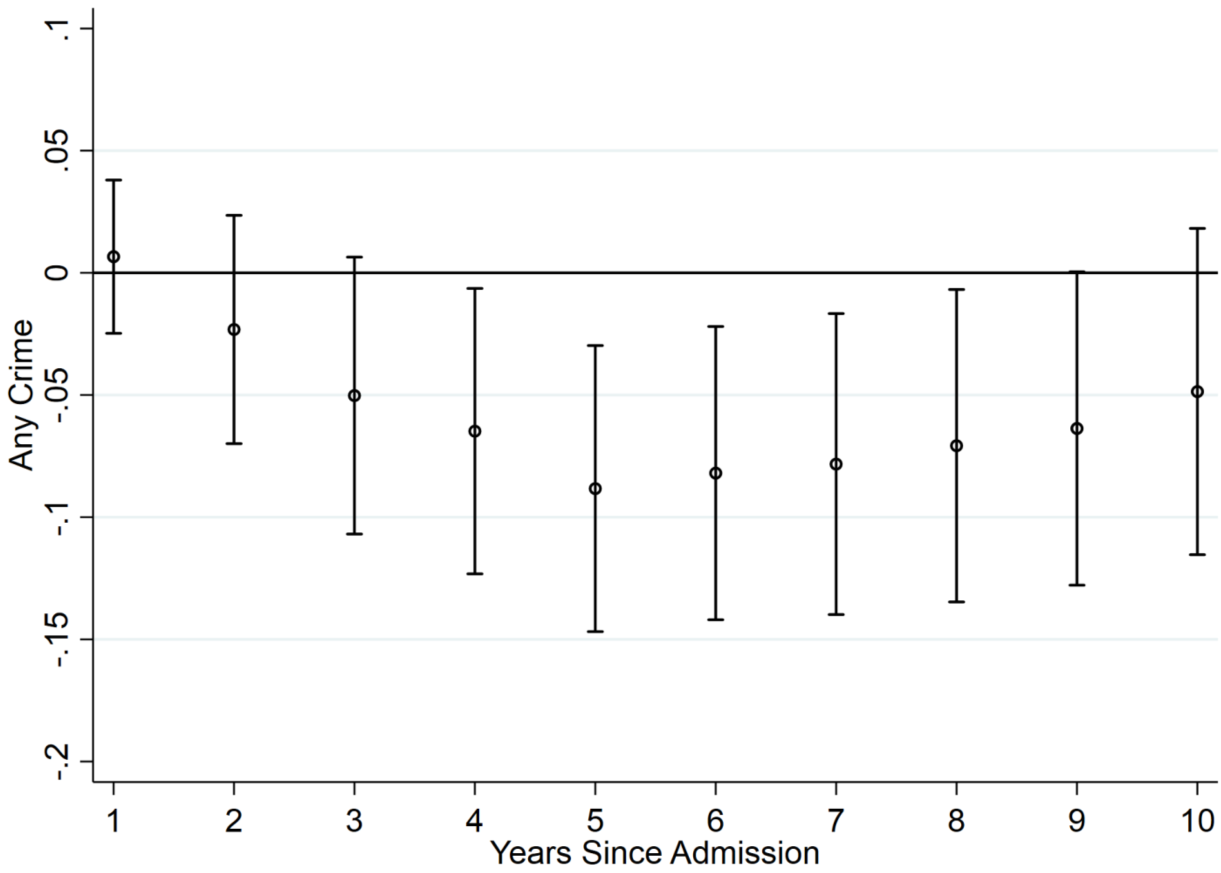
Notes: The figure shows the share of those who had committed any crime within two years (left) within ten years (right). The sample is all programmess with at least 5 applicants on both side of the cut off for which the programme was the best chance of gaining admission (sample 2). The lines represent local second order polynomial estimates using the edge kernel and the optimal IK bandwidths. The dots correspond to the sample means by 0.1 GPA point bins.

Figure 6 The effect of admission on crime: males



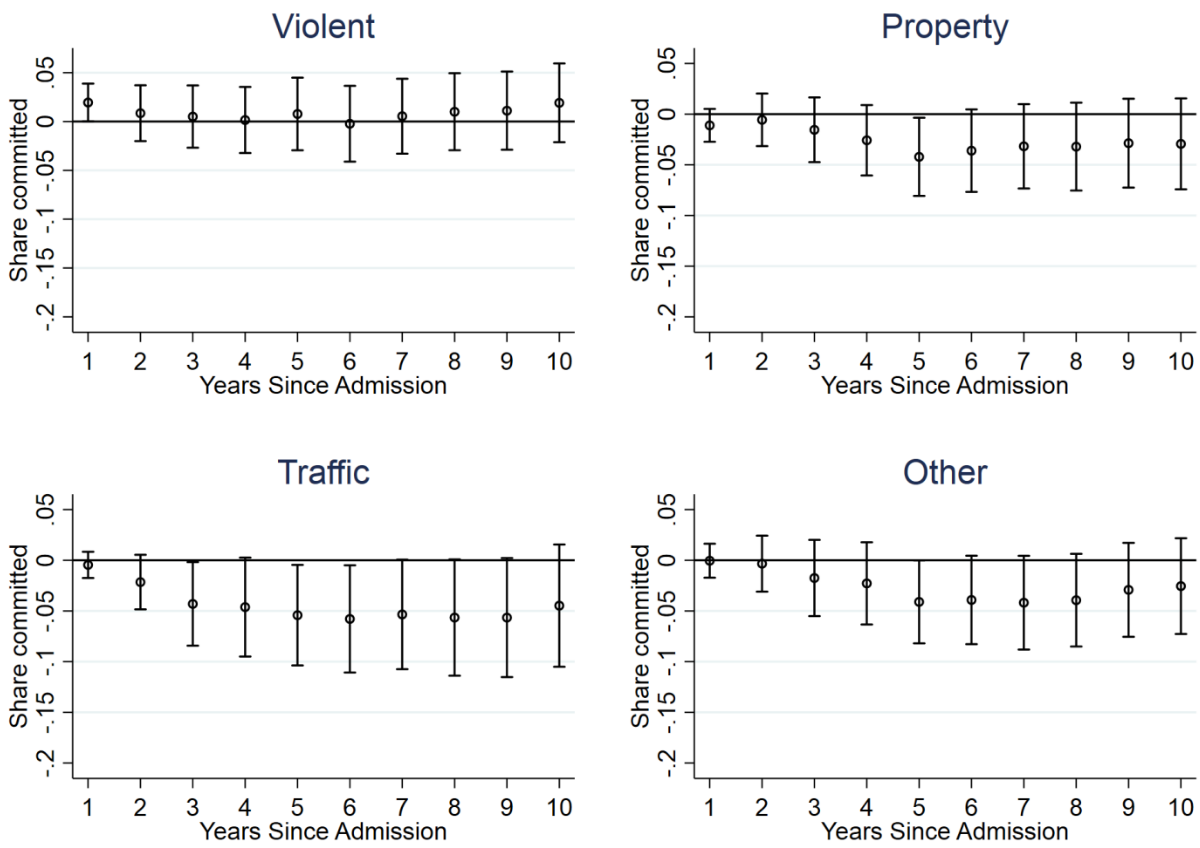
Notes: Each panel plots the effect of admission on annual crime from years 0 to 10 obtained from separate regressions. Year 0 is defined as August 15th the fall when starting ninth grade until August 14th the following year, year 1 is the admission year August 15th-August 14th following year etc. We use bandwidth 0.5 for all outcomes and for all programs.

Figure 7 The effect of admission on cumulative crime: males



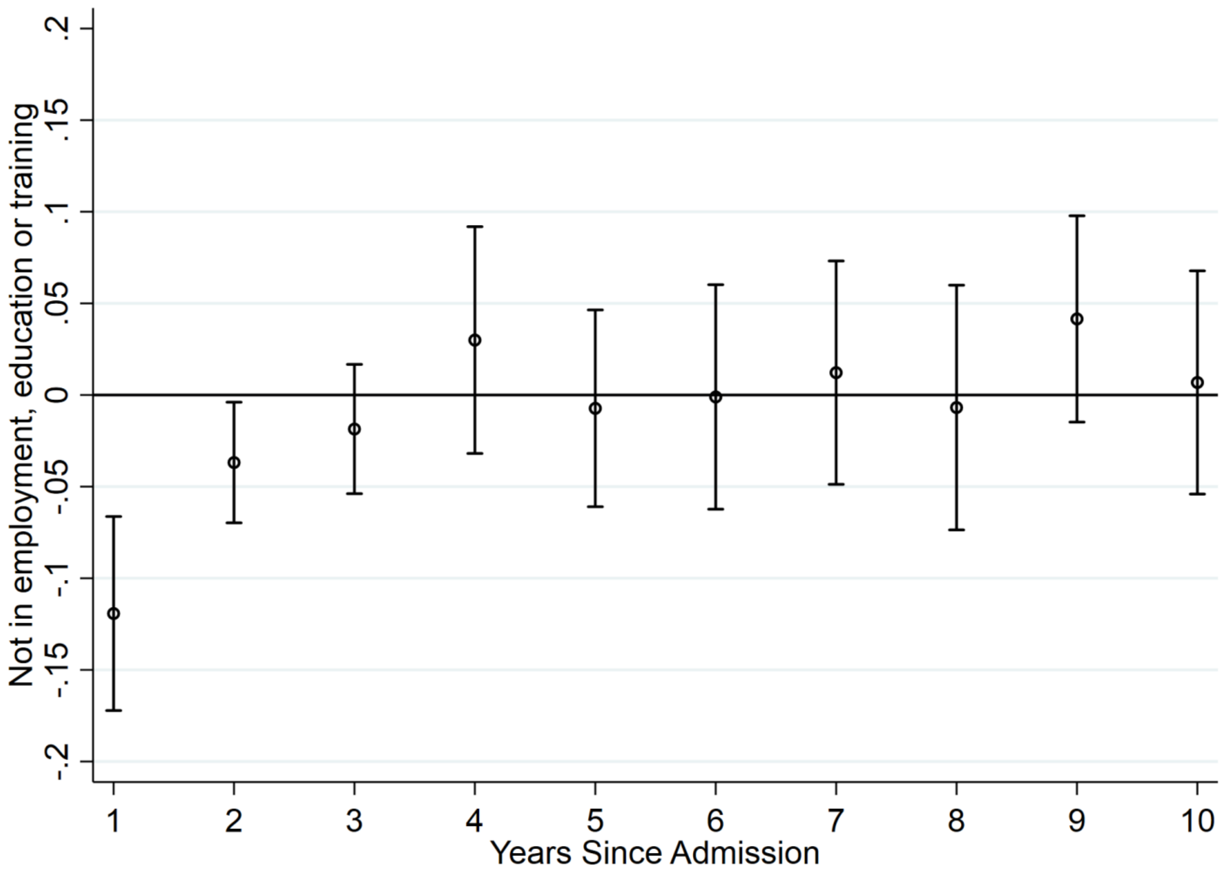
Notes: Each panel plots the effect of admission on cumulative crime since admission (year 1) from years 1 to 10 obtained from separate regressions. Year 0 is defined as August 15th the fall when starting ninth grade until August 14th the following year, year 1 is the admission year August 15th-August 14th following year etc. We use bandwidth 0.5 for all outcomes and for all programmes.

Figure 8 The effect of admission on cumulative crime, by crime type: males



Notes: Each panel plots the effect of admission on cumulative crime from years 0 to 10 obtained from separate regressions. Year 0 is defined as August 15th the fall when starting ninth grade until August 14th the following year, year 1 is the admission year August 15th-August 14th following year etc. We use bandwidth 0.5 for all outcomes and for all programmes.

Figure 9 The effect of admission on inactivity: males



Notes: Each panel plots the effect of admission on NEET (inactivity) status since admission (year 1) from years 1 to 10 obtained from separate regressions. Year 0 is defined as August 15th the fall when starting ninth grade until August 14th the following year, year 1 is the admission year August 15th-August 14th following year etc. We use bandwidth 0.5 for all outcomes and for all programmes.

Table 1 Effect of being above admission threshold/admission on crime

	(1)	(2)	(3)	(4)	(5)
	Enrolled to secondary	Crime by year 1	Crime by year 3	Crime by 5	Crime by 10
			Panel A All		
Reduced form:	0.326***	0.003	-0.011	-0.036**	-0.043*
	(0.035)	(0.007)	(0.014)	(0.017)	(0.025)
First stage:	0.614***	0.701***	0.710***	0.688***	0.678***
	(0.036)	(0.025)	(0.026)	(0.031)	(0.039)
LATE: Admitted	0.531***	0.004	-0.015	-0.052**	-0.064*
	(0.043)	(0.009)	(0.020)	(0.024)	(0.037)
Mean below*	0.307	0.027	0.080	0.127	0.187
Observations	9,689	14,817	12,469	11,376	7,955
Optimal bandwith	0.364	0.531	0.442	0.407	0.300
			Panel B Males		
Reduced form:	0.396***	0.006	-0.051*	-0.082***	-0.077*
	(0.048)	(0.015)	(0.030)	(0.029)	(0.041)
First Stage:	0.757***	0.686***	0.751***	0.706***	0.725***
	(0.037)	(0.033)	(0.030)	(0.033)	(0.037)
LATE: Admitted	0.523***	0.009	-0.068*	-0.116***	-0.105*
	(0.058)	(0.021)	(0.040)	(0.042)	(0.056)
Mean below*	0.335	0.046	0.127	0.201	0.266
Observations	4384	5653	5654	5637	4247
Optimal bandwith	0.320	0.386	0.383	0.406	0.295
			Panel C Females		
Reduced form:	0.325***	-0.001	0.010	-0.006	-0.008
	(0.059)	(0.006)	(0.008)	(0.014)	(0.018)
First Stage:	0.489***	0.663***	0.636***	0.619***	0.582***
	(0.060)	(0.045)	(0.030)	(0.036)	(0.039)
LATE: Admitted	0.664***	-0.002	0.017	-0.009	-0.013
	(0.080)	(0.009)	(0.013)	(0.022)	(0.032)
Mean below*	0.284	0.008	0.033	0.054	0.087
Observations	4146	4054	9167	6735	5666
Optimal bandwith	0.298	0.321	0.669	0.521	0.441

Notes: Each entry is an estimated effect from a local linear regression , triangular kernel weights and a program-specific optimal bandwidth (using CCT2014 bandwidth selection rule). Robust standard errors clustered at program-level are displayed in parentheses. *Mean below admission cutoff within optimal bandwidth. Optimal bandwidth is the mean bandwidth across all observations measured as GPA units.

Table 2 Effects by type of crime by year 5 and by year 10, males.

Panel A)	Cumulative crime by time 5:			
	Violence	Property	Traffic	Other
Reduced form:	0.010	-0.022	-0.036	-0.047**
	(0.019)	(0.019)	(0.028)	(0.019)
First Stage:	0.709***	0.730***	0.754***	0.710***
	(0.034)	(0.032)	(0.032)	(0.035)
LATE: Admitted	0.014	-0.030	-0.047	-0.067**
	(0.027)	(0.026)	(0.038)	(0.027)
Mean below*	0.045	0.082	0.095	0.064
Observations	5330	5192	4784	5259
Optimal bandwith	0.371	0.375	0.318	0.369

Panel B)	Cumulative crime by time 10:			
	Violence	Property	Traffic	Other
Reduced form:	0.015	-0.018	-0.051	-0.068**
	(0.023)	(0.024)	(0.039)	(0.027)
First Stage:	0.720***	0.732***	0.737***	0.730***
	(0.034)	(0.035)	(0.038)	(0.036)
LATE: Admitted	0.021	-0.025	-0.070	-0.093**
	(0.032)	(0.033)	(0.053)	(0.036)
Mean below*	0.073	0.108	0.146	0.101
Observations	5089	4701	3966	4503
Optimal bandwith	0.357	0.338	0.282	0.305

Notes: Outcomes measure whether individual has committed at least one crime in the category by the given time (year 5 or year 10 after admission). . *Mean below admission cutoff within optimal bandwidth.

Table 3 Effects during school years and summer breaks, males.

	Year 1 School year	Summer	Year 2 School year	Summer	Year 3 School year	Summer
	Panel A All					
Reduced form:	0.003	-0.002	-0.011	-0.001	-0.030	-0.028
	(0.016)	(0.009)	(0.019)	(0.012)	(0.026)	(0.017)
First Stage:	0.678***	0.770***	0.735***	0.778***	0.720***	0.696***
	(0.035)	(0.035)	(0.032)	(0.036)	(0.031)	(0.036)
LATE: Admitted	0.004	-0.002	-0.015	-0.002	-0.041	-0.040
	(0.023)	(0.012)	(0.026)	(0.015)	(0.036)	(0.025)
Mean below*	0.035	0.017	0.067	0.028	0.106	0.042
Observations	5229	5862	6082	5528	5742	4728
Optimal bandwith	0.362	0.416	0.404	0.395	0.395	0.326

Notes: See text under table 1.

Table 4 Effect of being above admission threshold/admission on activity and education

	NEET year 1	Degree by year 3	Degree by year 10
	Panel A All		
Reduced form:	-0.075***	0.148***	0.042**
	(0.018)	(0.021)	(0.020)
First stage:	0.704***	0.646***	0.725***
	(0.032)	(0.031)	(0.031)
LATE: Admitted	-0.107***	0.229***	0.058**
	(0.025)	(0.032)	(0.029)
Mean below*	0.167	0.123	0.718
Observations	10,329	12,950	12,128
Optimal bandwidth	0.364	0.470	0.454
	Panel B Males		
Reduced form:	-0.099***	0.129***	0.072**
	(0.030)	(0.026)	(0.030)
First Stage:	0.689***	0.644***	0.759***
	(0.038)	(0.029)	(0.034)
LATE: Admitted	-0.143***	0.200***	0.095**
	(0.044)	(0.041)	(0.040)
Mean below*	0.151	0.124	0.680
Observations	4494	7931	6395
Optimal bandwidth	0.320	0.541	0.499
	Panel C Females		
Reduced form:	-0.077**	0.190***	0.015
	(0.032)	(0.033)	(0.033)
First Stage:	0.541***	0.647***	0.630***
	(0.052)	(0.040)	(0.053)
LATE: Admitted	-0.142**	0.294***	0.024
	(0.058)	(0.047)	(0.053)
Mean below*	0.197	0.122	0.756
Observations	3571	6664	5036
Optimal bandwith	0.291	0.477	0.372

Notes: Each entry is an estimated effect from a local linear regression , triangular kernel weights and a program-specific optimal bandwidth (using CCT2014 bandwidth selection rule). Robust standard errors clustered at program-level are displayed in parentheses. NEET 1st (2nd) is an indicator whether individual is not in education nor in employment at end of the admission year, or end of the next calendar year. Degree in 3 (10) years, is an indicator whether individual has any post-compulsory degree within 3 (10) and a half years from the admission date. *Mean below admission cut-off within optimal bandwidth.

Appendix

Figure A1 Finnish educational system

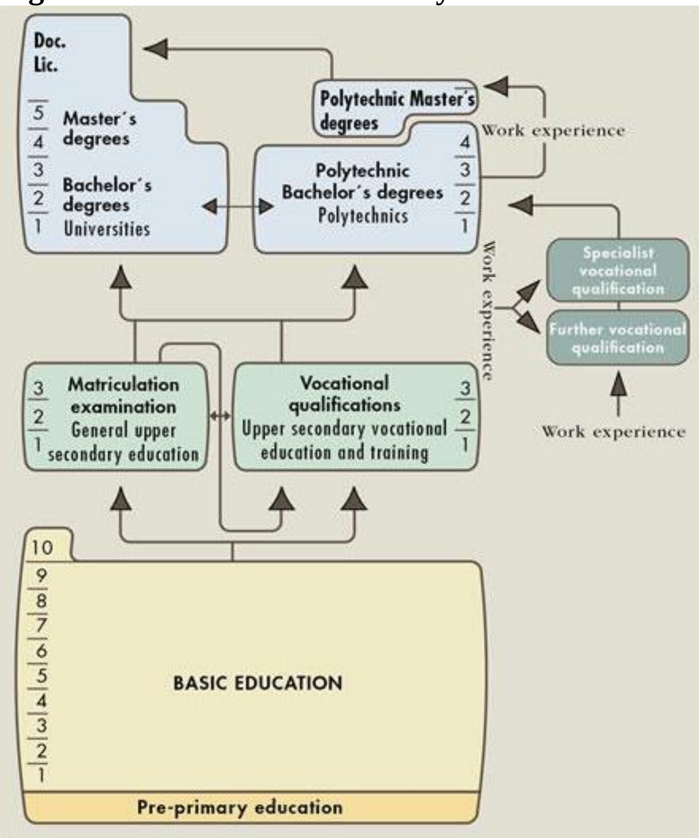
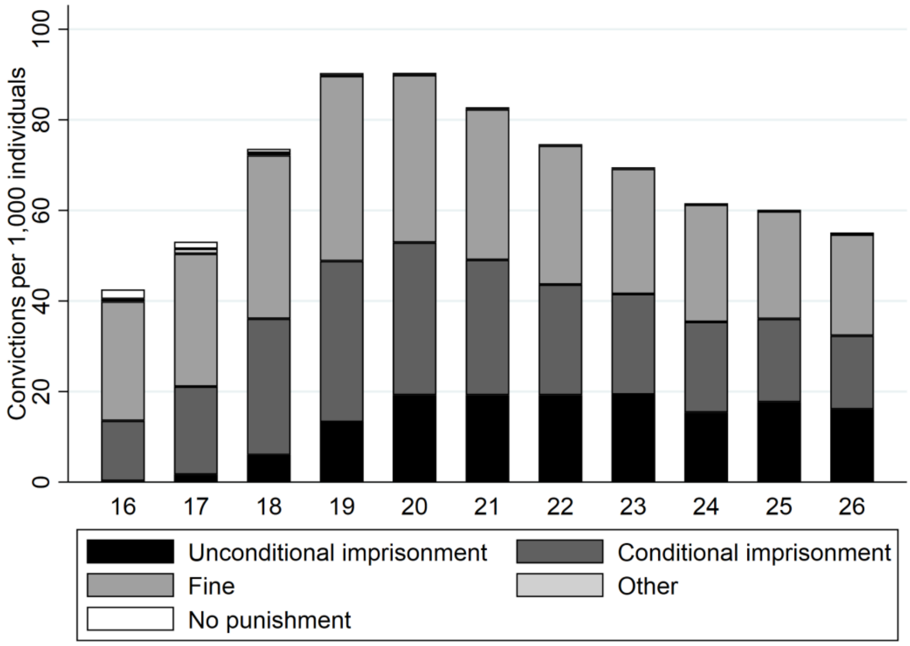
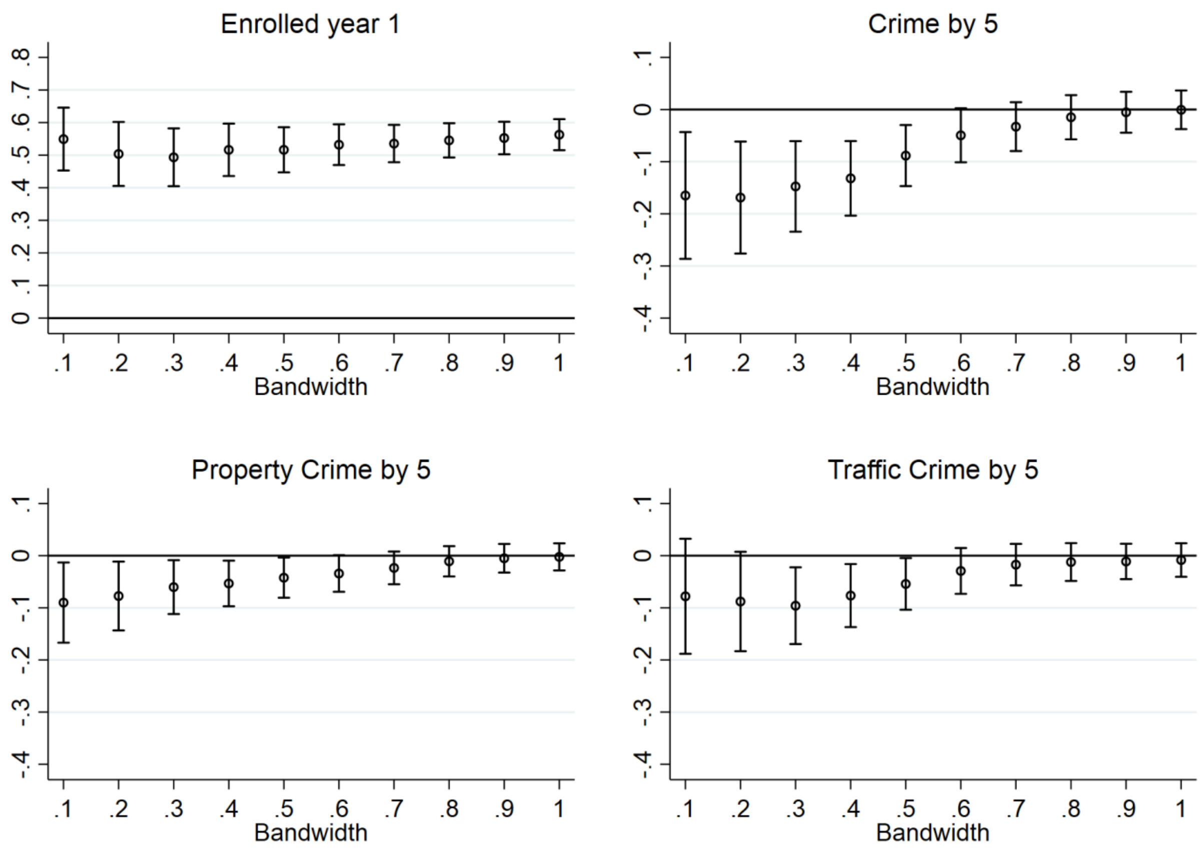


Figure A2 Crime-age profile by conviction



Notes: The sample consists of all school applicants in Finland in years 1996-2003. Each cohort is followed from age 16 until age 26.

Figure A3 Robustness to bandwidth choice to different outcomes, males: Enrolled year 1, Any crime by year 5, Property crime by year 5, Traffic crime by year 5



Notes: Each panel plots the effect of admission on outcome from separate regressions that use different bandwidth.

Table A1 Characteristics of total data vs. estimations samples: all applicants

	Total data		Sample 1		Sample 2	
PANEL A: Programme level information						
Programme characteristics						
Mean number of applicants	84.11		158.08		300.17	
High schools	0.24		0.34		0.59	
Lowest GPA among admitted	6.09		6.15		6.61	
No of programmes	16,969		4,169		669	
PANEL B: Individual level information						
	Total data		**Sample 1**		**Sample 2**	
	Admitted	Rejected	Admitted	Rejected	Admitted	Rejected
Individual characteristics						
Male	0.51	0.52	0.51	0.57	0.46	0.55
GPA	7.72	6.36	7.90	6.35	8.15	6.52
Non-Finnish or Swedish speaker	0.01	0.03	0.01	0.03	0.01	0.03
Lives in the 15 largest city	0.23	0.34	0.32	0.37	0.39	0.39
Crime year 0	0.01	0.04	0.01	0.04	0.01	0.03
Parental background						
Father's income	33,531	30,727	36,266	30,475	40,434	32,571
Father has secondary degree	0.71	0.63	0.74	0.63	0.77	0.65
Father has HE	0.17	0.11	0.21	0.11	0.27	0.14
Mother's income	22,840	21,318	23,971	21,451	25,919	22,801
Mother has secondary degree	0.77	0.67	0.78	0.67	0.80	0.68
Mother has HE	0.13	0.08	0.16	0.08	0.20	0.10
Outcomes						
Crime within 3 yrs	0.05	0.12	0.04	0.12	0.03	0.10
Crime within 10 yrs	0.11	0.24	0.10	0.25	0.08	0.21
Secondary degree within 10 yrs	0.90	0.62	0.92	0.63	0.94	0.66
No of individuals	411,351	18,895	16,0331	12,724	55,421	6,737

Table A2 Discontinuity of covariates at cut-off

Variables	All		Males		Females	
Individual characteristics						
Male	-0.007	(0.029)				
Crime before admission	-0.003	(0.005)	-0.003	(0.011)	-0.003	(0.003)
GPA	0.008	(0.009)	0.012	(0.015)	0.034**	(0.014)
Native language Finnish	-0.010	(0.008)	-0.012	(0.012)	-0.013	(0.011)
Native other than Finnish or Swedish	0.011	(0.007)	0.013	(0.010)	0.009	(0.010)
Age at the time of graduation	0.028**	(0.012)	0.021	(0.017)	0.043***	(0.014)
Lives in the 15 largest city	-0.012	(0.010)	-0.005	(0.014)	-0.019	(0.019)
Parents						
Information on mother	0.001	(0.006)	0.010	(0.008)	-0.004	(0.008)
Information on father	-0.007	(0.011)	0.028	(0.018)	-0.032*	(0.017)
Information on both parents	0.005	(0.013)	0.032	(0.020)	-0.029	(0.019)
Father's income	1914	(5137)	6493	(9793)	-2844	(2266)
Father in NEET	-0.005	(0.027)	-0.053	(0.044)	0.012	(0.045)
Father has post-compulsory degree	-0.018	(0.029)	-0.052	(0.051)	0.018	(0.052)
Father has HE	0.009	(0.020)	-0.009	(0.035)	0.017	(0.028)
Mother's income	797	(688)	1639	(1228)	1023	(1444)
Mother in NEET	0.006	(0.022)	0.018	(0.027)	0.021	(0.041)
Mother has post-compulsory degree	-0.024	(0.026)	-0.019	(0.042)	-0.063	(0.043)
Mother has HE	-0.005	(0.015)	0.027	(0.025)	-0.014	(0.021)

Notes: Each cell corresponds to coefficient from reduced-form regression (being above threshold). See text under table 1 for estimation details.

Table A3 Effects when controlling for covariates

	Enrolled to secondary	Crime by 1	Crime by 3	Crime by 5	Crime by 10
	Panel A All				
LATE: Admitted	0.534***	0.004	-0.020	-0.038	-0.056
	(0.045)	(0.008)	(0.020)	(0.025)	(0.039)
Observations	8592	13245	11159	10196	7078
	Panel B Males				
LATE: Admitted	0.550***	0.010	-0.079**	-0.083*	-0.116*
	(0.060)	(0.024)	(0.040)	(0.043)	(0.061)
Observations	3,896	5027	5062	5055	3797
	Panel C Females				
LATE: Admitted	0.610***	-0.004	0.013	-0.012	-0.034
	(0.089)	(0.011)	(0.013)	(0.024)	(0.034)
Observations	3689	3589	8182	5996	5023

Notes: The covariates include Indicator for mother tongue, age at graduation (in years), indicator for living in one of the 15th largest municipalities in Finland, gender, mother/father found in the register, mother's level of schooling. See text under table 1.

Table A4 Effects with fixed bandwidth

	Enrolled to secondary	Crime by 1	Crime by 3	Crime by 5	Crime by 10
	Panel A All				
LATE: Admitted	0.544***	0.010	-0.016	-0.059***	-0.050**
	(0.034)	(0.010)	(0.019)	(0.021)	(0.025)
Observations	13049	13049	13049	13049	13049
	Panel B Males				
LATE: Admitted	0.515***	0.007	-0.050	-0.088**	-0.049
	(0.044)	(0.019)	(0.034)	(0.036)	(0.041)
Observations	6905	6905	6905	6905	6905
	Panel C Females				
LATE: Admitted	0.599***	0.008	0.010	-0.028	-0.036
	(0.054)	(0.007)	(0.017)	(0.024)	(0.029)
Observations	6144	6144	6144	6144	6144

Notes: Each entry is an estimated effect from a local linear regression. We use bandwidth 0.5 for all outcomes and for all programmes.

Table A5 Robustness to sampling: Sample 1

	Enrolled to secondary	Crime by 1	Crime by 3	Crime by 5	Crime by 10
	Panel A All				
LATE: Admitted	0.538***	-0.027***	-0.029***	-0.039***	-0.036***
	(0.020)	(0.007)	(0.010)	(0.011)	(0.012)
Observations	24100	145500	108324	109696	109758
Mean	0.345	0.043	0.113	0.171	0.238
Optimal bandwidth	.456	1.904	1.446	1.454	1.453
	Panel B Males				
LATE: Admitted	0.534***	-0.042***	-0.039**	-0.051***	-0.040**
	(0.025)	(0.011)	(0.016)	(0.017)	(0.018)
Observations	13468	57757	48903	49676	51156
Mean	0.365	0.068	0.170	0.254	0.337
Optimal bandwidth	0.453	1.460	1.254	1.267	1.298
	Panel C Females				
LATE: Admitted	0.548***	-0.002	-0.004	-0.018**	-0.032***
	(0.032)	(0.004)	(0.008)	(0.009)	(0.012)
Observations	12044	92421	76683	79249	69213
Mean	0.314	0.010	0.035	0.117	0.105
Optimal bandwidth	0.509	2.954	2.051	2.103	1.837

Notes: Each entry is an estimated effect from a local linear regression, triangular kernel weights and a programme-specific optimal bandwidth (using CCT2014 bandwidth selection rule). Robust standard errors clustered at programme-level are displayed in parentheses. The sample include all applicants to tracks with at least 1 applicant on both side of the cut-off .