UNIVERSITY OF COPENHAGEN
DEPARTMENT OF ECONOMICS

Topics in Economics of Higher Education: Choices and Returns

PhD Thesis

Anne Toft Hansen

Supervisors: Mette Ejrnæs, Mette Gørtz, and Paul Bingley

Submitted on: July 13, 2019

Contents

Acknowledgements

Writing this dissertation has for sure been a crazy ride. There has been many ups and downs but it has definitely made me stronger and wiser. Not only academically but also on a personal level. I owe everything to my family, friends, and colleagues that have acted as an incredible support system during the last three years.

First of all, I am grateful for having been given the possibility to write this dissertation. Without the financial support from University of Copenhagen, Kraks Fond - Institute for Urban Economic Reseach, and VIVE - The Danish Center for Social Science Research it would not have been possible.

Also a huge thanks to my supervisors Mette Ejrnæs, Mette Gørtz, and Paul Bingley. You have constantly been inspiring me and have challenged me to work harder and think outside the box. Mette Ejrnæs, you have also supported me personally through a very rough year and helped me to prioritize and keep my eyes on the price. Thank you so much for this.

Sally Sadoff and Michael Kuhn, thank you for all your comments and late-night Google Hangout discussions. Ulrik Hvidman, Hans Henrik Sievertsen, and Helene Willadsen, you have been amazing to work with. I have learned a lot from you guys.

Bo Honore, thank you for inviting me to Princeton University, for inspiring inputs, great beers, and for being an excellent host.

I have also been blessed with the best office buddies both at the Econ Department and at VIVE. Huge thanks to Jonas Hirani, Petra Cavalca, Sarah Sander, Bjørn Meyer, Mikkel Giver Kjer, Kirstine Karmsteen, Mette Bisgaard, and Malene Kessing. You have been great colleagues both when it comes to helping with technical issues or just going out for beers – even when I accidentally pushed a stranger's bike in the canals of Nyhavn - you were there to help me get the bike out of the water.

Also, my dearest friends, Sine, Sine, Tina, Thora, and Anne. Thank you for

thousands and thousands of pep talks and for making the time outside the office a great time.

Ulf, living with me during the final period of writing cannot have been easy. Still, you have given me an invaluable home base with your calm being and constant support. I cannot thank you enough.

Lastly, I would like to thank my mom and dad for their unconditional love and support. Mom, you are the strongest woman I know and I admire you immensely. I dedicate this dissertation to you.

Anne Toft Hansen
Copenhagen, June 2019

Dansk introduktion

I klassisk uddannelsesøkonomisk teori betragtes uddannelse som en investering. Et individ vælger at investere i en uddannelse, hvis det formodede fremtidige afkast er højere end omkostningerne ved at investere.

Den uddannelsesøkonomiske litteratur betoner ofte afkastet af uddannelse i monetær forstand, som fx fremtidig indkomst. Dog har litteraturen også rettet fokus på ikke-monetære afkast af uddannelse, som vedrører fx fertilitet [1], sundhed, kriminalitet og lykke. [2] En anden del af litteraturen, der omhandler valg af uddannelse, finder endvidere, at ikke-monetære afkast er vigtigere for valget af uddannelse end de monetære afkast. [3]

Denne Ph.d.-afhandling skriver sig ind i denne litteraturs forskellige forgreninger i forhold til både valg og afkast af uddannelse. Afhandlingen indeholder tre separate kapitler, der omhandler forskellige aspekter af videregående uddannelse. Kapitel et afdækker kønsforskelle i valg af videregående uddannelse, kapitel to undersøger betydningen af uddannelsessted for valg af bopæl og arbejdsplads. Kapitel tre undersøger, om der er en signalværdi [4] af karaktergennemsnit fra universitetet på arbejdsmarkedet i de første fem år efter afsluttet uddannelse. I det følgende går jeg i dybden med hver af de tre kapitler.

[1] Humlum, M. K., Kristoffersen, J. H., and Vejlin, R. (2017). College admission decisions, educational outcomes, and family formation. *Labour Economics*, 48:215-230.

[2] Oreopoulos, P. and Salvanes, K. G. (2011). Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives*, 25(1):159-184.

[3] See for example: Zafar, B. (2013) College major choice and the gender gap. *Journal of Human Resources*, 48(3):545-595 and Boneva, T. and Rauh, C. (2017). Socio-economic gaps in university enrollment: The role of perceived pecuniary and non-pecuniary returns. *CESifo Working Paper Series*

[4] Ifølge humankapitalteori er uddannelse relateret til et højere afkast, fordi uddannelse øger produktivitet. I modsætning hertil hævder signalteorien, at uddannelse medfører et afkast, fordi uddannelse afspejler medfødte evner (og ikke hvad den studerende har lært på en given uddannelse)

Kapitel 1

med Helene Willadsen

Kapitel 1 undersøger, hvordan kønsforskelle i valg af uddannelse relaterer sig til den kønsrelaterede lønforskel. Konkret indskriver kapitlet sig i litteraturen om "child penalty", som påviser, at lønforskelle mellem mænd og kvinder kan forklares med, at kvinder straffes lønmæssigt, når de får børn. I kapitlet benytter vi et unikt spørgeskema, som vi har distribueret til alle ansøgere på de videregående uddannelser i 2018. Med mere end 17.000 respondenter er spørgeskemaundersøgelsen det største datagrundlag om uddannelsesvalg endnu set i litteraturen om valg af uddannelse.

Vi finder tre centrale resultater. For det første, finder vi, at selvom kvinder og mænd placerer næsten lige stor vægt på økonomiske faktorer (forventet indkomst og beskæftigelse) i deres valg af uddannelse, prioriterer mænd i større omfang deres livskvalitet under studiet end kvinder. Kvinder prioriterer derimod fremtidige ikke-økonomiske faktorer som fx jobtilfredshed og forventet sandsynlighed for at få børn højere end mænd i deres uddannelsesvalg. For det andet forventer ansøgerne kønsrelaterede lønforskelle inden for alle fagområder, når de søger videregående uddannelse. Vi fokuserer på en gruppe ansøgere, der har både interesse og evner for matematik, og finder, at selv blandt denne mere homogene gruppe forventer kvinderne at tjene mindre end deres fremtidige mandlige kolleger, når de søger uddannelse. For det tredje indikerer vores resultater, at kvinder prioriterer hensynet til deres forventninger til at få børn, når de søger uddannelse. Kvinder, som rapporterer, at de med høj sandsynlighed har børn 10 år efter endt studie, er mere tilbøjelige til at vælge et ikke-matematisk fagområde sammenlignet med kvinder, der rapporterer en mindre sandsynlighed for at få børn.

Vores resultater indikerer, at den eksisterende litteratur om "child penalty" bør inkludere betydningen af uddannelsesvalg. Kapitlet dokumenterer, at mænd og kvinder prioriterer forskelligt i uddannelsessystemet, betinget af deres forventninger til at få børn. Mere opsigtvækkende viser kapitlet, at kvinder, der i høj grad forventer at få børn, er tilbøjelige til at vælge uddannelser med udsigt til et mere børnevenligt arbejdsmarked, selvom det måske også medfører en lavere løn: Vi finder, at kvinder muligvis forventer, at en karriere inden for et matematisk fagområde kan være svært at forene med at have børn. Implikationen af dette fund er, det kan være relevant at informere potentielle ansøgere om familievenlige karrieremuligheder frem for primært at informere om løn og arbejdsløshed, hvis politiske beslutningstagere ønsker at optimere antallet af (kvindelige) ansøgere til de matematiske uddannelser.

Kapitel 2

Universitetskandidater er en gruppe af arbejdere, som de fleste regioner ønsker at tiltrække. Et øget uddannelsesniveau i lokalbefolkningen øger ikke kun den samlede produktivitet men er også forbundet med sociale afkast, såsom reduceret kriminalitet og øget politisk deltagelse i lokalsamfundet. Derfor er det et centralt politisk mål for både statslige, men i særlig grad kommunale og regionale beslutningstagere at tiltrække og fastholde universitetsstuderende.

Kapitel 2 undersøger effekten af, hvor de studerende læser på universitet på, hvor de ender med at bo og arbejde efter afsluttet uddannelse. En særlig metodisk fordel er, at ansøgerne ikke kender de årlige adgangskvotienter til en given uddannelse, hvilket skaber en tilfældig variation omkring hvilken uddannelse, ansøgerne kommer ind på. Ved hjælp af et fuzzy regression discontinuity design, der kontrollerer for, at valg af fag og uddannelsessted er et ordinalt valg, undersøger jeg effekten af at studere i en storby (Århus eller København) eller uden for storbyerne på kandidaternes valg af bopæl og arbejdsplads 8 år efter, de har søgt ind på universitetet.

Jeg finder, at den geografiske placering af kandidaternes uddannelsessted har en effekt på, hvor kandidaterne efterfølgende vælger at bo og arbejde. Universitetsansøgere, der oprindeligt foretrak at studere i en storby, men ikke kom ind, og i stedet gennemfører en uddannelse på et universitet uden for en storby, er mere tilbøjelige til at bo og arbejde uden for storbyområderne 8 år efter, de har søgt uddannelse sammenlignet med kandidater, der blev færdiguddannede fra deres foretrukne universitetsvalg.

Resultaterne i kapitlet indikerer, at det kan være et nyttigt politisk værktøj at få flere til at studere uden for Århus og København, hvis målet er at øge andelen af højtuddannede i hele Danmark og ikke kun i storbyerne.

Kapitel 3

med Ulrik Hvidman and Hans Henrik Sievertsen

Et dårligt match mellem jobansøgere og arbejdsgivere kan have store konsekvenser for ulighed og produktivitet på samfundsniveau, ligesom mismatchet endvidere kan indebære potentielle omkostninger for de individuelle arbejdsgivere og jobansøgere. Det kan være udfordrende at skabe et godt match mellem arbejdsgivere og jobansøgere, da det ikke er muligt for arbejdsgivere at observere jobansøgernes sande produktivitet. Ifølge den økonomiske signalteori bruger arbejdsgivere jobansøgeres uddannelsesniveau, som et signal om deres arbejdsmarkedsproduktivitet. Men befolkningens uddannelsesniveau er steget over de seneste årtier, og derfor står arbejdsgiverne ofte over for et valg mellem ansøgere med tilsvarende uddannelsesniveau (fx en universitetsgrad) og har brug for andre signaler om produktivitet, som fx karakterer inden for et givent uddannelsesniveau.

Dette kapitel identificerer signalværdien af karakterer på indkomst og undersøger endvidere, hvor hurtigt arbejdsgivere lærer om dette signal, der ikke er relateret til arbejdsmarkedsproduktivitet. Vi genbesøger derfor et centralt spørgsmål i arbejdsmarkedsøkonomien: Er effekten af uddannelse på indkomst bedst forklaret af humankapital eller signalværdi af uddannelse?

På baggrund af eksogen variation i universitetsstuderendes karakterer, forårsaget af en karakterreform fra 2007, undersøger vi, om denne eksogene variation har en effekt på indkomst i de første fem år på arbejdsmarkedet. Vi finder, at de studerende, der oplever et positivt stød til deres eksamensgennemsnit, får en højere indkomst i de første to år efter afsluttet uddannelse, hvorefter effekten går mod nul. Vores resultater tyder derfor på, at arbejdsgiverne lærer deres medarbejders sande produktivitet at kende efter de første to år og justerer lønnen derefter.

Disse resultater har direkte politiske implikationer, da karakterer er centrale for det bedst mulige match mellem arbejdsgivere og nyuddannede medarbejdere. Vores resultater indikerer, at hvis en studerende gives en tilfældig anden karakter, vil det (alt andet lige) få konsekvenser på arbejdsmarkedet for den studerende.

English introduction

In standard economic theory, education is an investment, and you choose to invest if the future returns are higher than the costs of investment. The future returns are often thought of in terms of monetary returns but education has also shown to have an effect on several nonpecuniary factors. In recent years, the literature has found an association between education and factors like fertility [5], crime, health, success in the marriage market, and happiness. [6] Additionally, the literature on the choice of education, shows that the prospects of different nonpecuniary factors are more important than pecuniary for prospective students. [7]

This dissertation adds to the literature with three self-contained chapters dealing with different aspects of choice and returns to higher education. Chapter 1 examines gender differences in the choice of higher education, chapter 2 investigates the role of university location on where graduates end up living and working, and chapter 3 examines the signaling value of the GPAs of university graduates in their first five years on the labor market.

Chapter 1

with Helene Willadsen

Chapter 1 presents the first attempt to relate the gender gap in the choice of education to the gender wage gap and the female child penalty in the labor market. For this purpose, we developed a unique survey and invited all higher educa-

[5] Humlum, M. K., Kristoffersen, J. H., and Vejlin, R. (2017). College admission decisions, educational outcomes, and family formation. *Labour Economics*, 48:215-230.

[6] Oreopoulos, P. and Salvanes, K. G. (2011). Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives*, 25(1):159-184.

[7] See for example: Zafar, B. (2013) College major choice and the gender gap. *Journal of Human Resources*, 48(3):545-595 and Boneva, T. and Rauh, C. (2017). Socio-economic gaps in university enrollment: The role of perceived pecuniary and non-pecuniary returns. *CESifo Working Paper Series*

tion applicants in 2018 to participate. With more than 17,000 respondents, this is the biggest survey seen in the literature on education choice. In the chapter, we focus on applicants who have shown an interest for math, by choosing to apply for at least one math field in their application for higher education. The chapter comprises of several smaller analyses that together draw a picture of the gender differences in choice of education among the group of applicants who have a preference for math.

Three important findings emerge. First, females and males attach nearly equal importance to pecuniary factors. Males appreciate the consumption value of studying more highly than females, who, instead, pay more attention to future non-pecuniary factors such as work satisfaction and expected probability of having children.

Second, prospective students expect a gender wage gap within all fields of study when applying for higher education. Even among our sample of math applicants, females expect to earn less than their male counterparts when applying for higher education.

Third, females seem to take future childbearing decisions into account when applying for education. Females who think that they will have children are more likely to choose a non-math field of study compared to females who report a lower probability of having children. This suggest that the child penalty in the labor market from the existing literature might fail to account that females already select themselves into different career paths based on expected childbearing. Specifically, some females might opt for education programs with child-friendly labor market prospects, although they may also earn less compared to math fields. In this sense, the child penalty strikes twice.

For policy purposes, our results indicate that females might believe that prospective careers in math fields are hard to reconcile with having children. As such, policies on promoting math fields should seek to inform prospective applicants about family-friendly careers rather than just informing about a good study environment or future labor market outcomes.

Chapter 2

University graduates are a group of workers that most regions wish to attract. Human capital spillovers to local regions do not only increase aggregate productivity but is also associated with social returns, such as reduced crime and increased political participation. Consequently, to attract and retain university graduates is a key goal for policy makers.

Chapter 2 examines the effect of university location on the choice of where to live and work of Danish university graduates. The fact that the annual admission cut-offs into higher education are unpredictable makes me able to account for self-selection into a given university location. Using a fuzzy RD-design that takes into account that the choice of field and location of education is an unordered choice, I estimate the causal effect of studying either in a metropolitan or a non-metropolitan region on where graduates live and work eight years after applying for university. My findings suggest that the location of study does have an effect on where graduates choose to live and work. Prospective students who initially preferred to study in a metropolitan region but did not get accepted and instead received an offer in a non-metropolitan university are more likely to live and work outside the metropolitan regions eight years after applying for university compared to graduates who graduated from their first choice university.

Another finding of this paper is that universities are able to retain some of its graduates. Some of the graduates who did not get accepted for their first choice university location, but instead graduate from a non-metropolitan second choice university are retained in their university location. Despite the fact that most of these graduates preferred to study in a metropolitan region.

This chapter provides evidence that pushing applicants for university to study in a non-metropolitan region will result in some of these graduates staying and entering the labor market in these non-metropolitan regions. Therefore, it could be a useful policy tool to achieve the goal of increasing the share of highly educated in local non-metropolitan regions.

Chapter 3

with Ulrik Hvidman and Hans Henrik Sievertsen

How workers are allocated across jobs has implications for inequality and efficiency at an aggregate level and involves large potential costs for employers and employees at the micro level. A major challenge in the matching process is that labor market productivity is imperfectly observed. According to job-market signaling theory, employers use completed schooling as a signal of labor market productivity to screen workers. However, as degrees have become increasingly common, these credentials constitute very crude signals and mask valuable information about the applicants' ability. Consequently, employers are often faced with a choice between applicants with similar levels of educational attainment (e.g., a university degree) and may therefore look for other signals of productivity, such as information on educational achievement.

Chapter 3 examines the signaling value of GPA on earnings and how fast employers learn about the worker's true productivity. We exploit a Danish grading reform causing as-good-as random variation in university graduates' GPA. Students who were enrolled in university during the implementation had their existing grades recoded to a new grading scale. Using this reform-induced variation we identify the effects of GPA on labor market outcomes. We find that a higher GPA causes higher earnings in the first two years after graduation, whereafter the signaling effect goes to zero. This suggests that employers learn about the workers true productivity after two years and adjust earnings accordingly.

These results have direct policy implications. Grades are relevant in the labor market matching process for university graduates. We find that if a student is given an arbitrary grade, all else equal (including exam performance), the student will have a different labor market outcome in the short run. This finding suggests that the grading system affects matching efficiency.

Chapter 1

Choosing to Do Math: Gender Differences in the Choice of Higher Education in Denmark

Choosing to Do Math: Gender Differences in the Choice of Higher Education *

Anne Toft Hansen †

Helene Willadsen ‡

Abstract

This paper examines gender differences in education choice by exploring how males and females value pecuniary and non-pecuniary factors in their choice of education. We relate the gender gap in the choice of education to the gender wage gap and the expected female child penalty in the labor market. For the purpose of this paper, we conducted a large-scale survey in 2018 for the full cohort of higher education applicants in Denmark. Focusing on applicants who had applied for at least one math-related field of education, we find that females expect to earn significantly less than males within all fields of study, thereby anticipating the gender wage gap. Additionally, we find gender differences in the choice of education based on the expected probability of having children. Furthermore, females are heterogeneous in their expectations for having children. Females who are almost certain that they will have children are more likely to apply for non-math fields and do not put much weight on pecuniary factors in their choice of education. The females who expect a lower probability of having children weight both expected earnings and probability of employment when choosing education and are also more likely to choose math-related fields of education compared to females who report a high probability of having children.

*We thank conference participants at the 2019 CEN meeting at The Rockwool Foundation, as well as seminar participants at University of Copenhagen, Center for Economic Behavior and Inequality, and VIVE-The Danish Center for Social Science Research. Furthermore, we thank Sally Sadoff, Michael Kuhn, Claus Thustrup Kreiner, Mette Gørtz, Mette Ejrnæs, and Marco Piovesan for valuable inputs

†The Danish Centre for Social Science Research & the Department of Economics, University of Copenhagen. E-mail: ath@vive.dk.

‡The Department of Economics, University of Copenhagen. E-mail: hw@econ.ku.dk.

1 Introduction

Differences in the field of higher education is an important factor for explaining the gender gap in earnings between highly educated males and females (Black et al., 2008). Major differences in the choice of education of males and females prevail worldwide. Males tend to apply for math-related education (e.g., science, technology, engineering, and mathematics (STEM)), and females tend to apply for non-math education (Zafar, 2013; Mostafa, 2019). In 2012 in the OECD countries, only 14 percent of females enrolled in STEM fields compared to 39 percent of males (OECD, 2015). Additionally, there are large gender differences in preferences concerning math-related fields. Females enrolling in math-related fields are more likely to choose fields within health, while males are more likely to choose fields such as engineering and economics. These male-dominated fields are associated with substantial labor market returns, even after controlling for ability sorting into different fields (Arcidiacono, 2004). If policy-makers wish to increase the proportion of graduates from math-related fields and/or improve gender equality, it is useful to understand why females opt out of math-related fields, which generally have higher returns in terms of future earnings and labor market opportunities.

It is a well-established finding that males earn more than females when controlling for hours worked, education, work experience, and industry (Blau and Kahn, 2000). This gender gap has decreased over the years, although it persists. Females in Denmark, as well as in other Scandinavian countries, are active on the labor market, with a labor market participation rate of only five percentage points lower than males (Kleven et al., 2018; OECD, 2018b). However, females earn around 12–20 percent less than males per hour worked, and this difference has been constant since the 1970s (Gupta et al., 2006). Several studies attribute this gender gap in earnings to a child penalty (lower earnings for females with children compared to females without children) (See for example: Blau and Kahn, 2000;

Kleven et al., 2018; Nielsen et al., 2004). Kleven et al. (2018) have identified a significant and large earnings gap of 20 percent between males and females following the birth of the first child, which tends to be non-existent for childless males and females. Besides, following the birth of the first child, females move into more family-friendly firms and the public sector. This claim is supported by Nielsen et al. (2004), who also find that females move into the public sector, which is associated with a lower child penalty.

This paper is the first attempt to relate the gender gap in the choice of education to the gender wage gap and the female child penalty on the labor market: If higher education applicants anticipate the child penalty, then expectations for future earnings, and family formation could play an important role in explaining gender differences in the choice of education. To relate the choice of education to future outcomes, such as earnings and fertility decisions, we need to gain insights into higher education applicants' beliefs about these outcomes and how they relate to their choice of education. Therefore, to elicit these beliefs, we designed a survey in 2018 for higher education applicants in Denmark. We reached out to the full cohort of higher education applicants, making it, to the best of our knowledge, the largest survey ever conducted on higher education choices. We elicited the beliefs on actual choices of education as we reached out to the applicants immediately after they had applied for education and before they had received an acceptance letter. For each of the applicants' top two educational choices, the survey inquired about both pecuniary and non-pecuniary factors, such as expected earnings and the probability of having children. We subsequently use the survey to estimate a choice model on fields of education choice.

For the purpose of this paper, we use a sample of applicants who have applied for a math field as their first or second choice of education. We define math fields as fields of higher education that require the highest level of math from high school for admission. Therefore, gender differences in preferences for math, cannot solely explain the differences in the choice of education.

Our choice model estimation suggests that pecuniary factors explain about 26 percent of the variation in the choice of education for females and 22 percent for males. The non-pecuniary factors 10 years after graduation are more important for the choice of education for females than males. These factors explain 51 percent of choice for females and only 34 percent for males. The consumption value of studying explains the largest share of the variation in males' choice (i.e., 44 percent) compared to 23 percent of that in females' choice. Furthermore, we find that expectations for income and the probability of having children are statistically significant drivers of the gender differences in the choice of education.

The finding that pecuniary factors are important but that non-pecuniary factors are even more important is in line with the existing literature on the choice of education. One strand of this literature focuses mainly on the pecuniary reasons for enrolling into higher education (Arcidiacono et al., 2012; Baker et al., 2018; Hastings et al., 2015; Jensen, 2010; Wiswall and Zafar, 2014) but recognizes that factors other than earnings are important for the choice of education. Another strand of the literature specifically investigates the significance of several non-pecuniary factors for enrolling in higher education. Boneva and Rauh (2017) find that social factors, enjoyment at work, and parental approval are more important than pecuniary returns when explaining why students with high socioeconomic status (SES) are more likely to apply for university than those with a low SES. Delavande and Zafar (2019) find that attending a university and following course work that is consistent with one's beliefs are important determinants when choosing an education. Zafar (2013) finds that males attach more importance to pecuniary workplace factors than females, and these expectations explain a significant part of the different patterns in the choices of education between males and females. However, our results show that females tend to attach equal importance to pecuniary factors as males.

Next, we investigate gender differences in expectations for income and the probability of having children. Descriptively, we find large differences in expected income and probability of having children. Males expect to earn significantly more than females in all fields, and females expect a higher probability of having children in all fields except social science math. We compare the expected income of males and females to actual income of previous cohorts within each field. We find that the expected earnings are significantly higher than the actual earnings. Additionally, we find that the applicants expect the gender gap in earnings fairly accurate within each field. We also find that correcting the expected income with the actual income in the choice model estimation does not account for gender differences in the education choice.

In a similar exercise, we estimate the gender gap if females on average had the same expectations for earnings and the probability of having children as males. Most interestingly, changing females' expectations about having children would move more females into male-dominated educations, although the effects are small. To explore this further, we base our choice model estimation on subsamples of males and females that indicate a high and low probability of having children. This analysis shows that females who expect a high probability of having children put less weight on pecuniary factors in their choice of education and are more likely to choose non-math fields. Females who expect a low probability of having children put more weight on pecuniary factors and are more likely to choose a math education within social science math. We do not find the same results for males, who only differ slightly in their choice of education depending on their expectations about having children. Taken together, this suggests that some education programs might be unattractive for females who expect to have children.

Related work by Wiswall and Zafar (2017) shows that a quarter of the gender gap in earnings early in one's career can be explained by different preferences for workplace characteristics. Females are willing to pay 7.3 percent of their wage for workplace flexibility, whereas males are only willing to pay 1 percent. Wiswall and Zafar (2017) interpret this as women purchasing positive workplace attributes. Our results, however, suggest a different interpretation in that females might anticipate a child penalty associated with certain fields. If this holds, then females and males might have the same willingness to pay for job attributes but face different constraints on the labor market related to having children.

From a policy perspective, if the goal is to increase the share of graduates from math related fields, our results points to specializing initiatives aimed at attracting more students. We find gender differences in the determinants of choice, where the consumption value of studying is more important for the male choice of education compared to the female choice. Females place future non-pecuniary factors much higher, suggesting that initiatives to attract more students should be designed differently for males and females. Furthermore, the wide spread policies of informing applicants about earnings might fail to attract more students compared to informing them about non-pecuniary factors, especially females with high expectations for having children.

The remainder of this paper is organized as follows. Section 2 describes the higher education system in Denmark. Section 3 presents our survey and the descriptive statistics, section 4 describes the choice model estimation, section 5 presents the results followed by sensitivity checks, and section 6 concludes.

2 Institutional Setting

In Denmark, education, including higher education, is free. Students enrolled in higher education receive a monthly grant of US$1000, and they also have the option of applying for governmental student loans in combination with the student grant [1].

[1] Denmark is one of the countries that has the highest public spending on education in the world; it allocated 6.3 percent of its GDP to it in 2014 (OECD, 2018a)

High school graduates apply for college and university education through a two-tiered centralized clearinghouse, where an applicant can apply for up to eight different programs in the same application. The vast majority of applicants (about 92 percent) apply through the first tier where applications are based solely on high school grade point average (GPA). Through a simple procedure, applicants create a rank-ordered list of education choices by logging in the website of the clearinghouse and choosing field of study, institution, and location. The remaining 8 percent of applicants apply based on their high school GPA, other merits (e.g., work experience and travel experience), and a written cover letter targeted at the specific education.

Unlike the American system, for example, where students specialize later, applicants in Denmark choose a specific field of study (e.g. economics, nursing, or anthropology) when applying for higher education. All education programs have a limited number of slots, and for about half of the programs, there are more applicants than available slots[2]. This means that some applicants will face restrictions in their first choice of education. Every year, a GPA-based admission cutoff is used to resolve the issue of having more applicants for specific programs than the available slots.

Table 1: Division of math fields

Nature math	Tech math	Health math	Social science math	Non-math
Chemistry	Engineering	Pharmacy	Economics	All other fields
Biology	Land-inspection[a]	Public Health	Economics and Politics	
Data Science		Medicine	Business and Math [b]	
Mathematics		Odontology		
Physics				
Geography				

[a]) Education not available in English, Danish name: Landinspektørvidenskab
[b]) Education not available in English, Danish name: Erhvervsøkonomi og matematik HA(mat.)

A centralized application system ranks the GPA of all the applicants for a given education program and assigns a slot to the applicants with the highest GPA until there are

[2]In 2018, 49 percent of the education programs had an admission cutoff.

no more slots. The GPA of the last applicant is the GPA admission cutoff within a given year. Because of the excess demand, some applicants will not succeed in enrolling into their preferred program. In such cases, they will then be assessed on the admission cutoff for their second choice, and the system continues until it locates an education program on the applicant's rank-ordered list where the applicant's high school GPA is higher than the specified admission cutoff. The GPA cutoffs are determined every year and are, therefore, unknown to the applicants at the time they submit their application. Because most educational programs only have annual enrollment, applicants are encouraged to state more than one education choice on their rank- ordered list if they wish to increase their chance of being admitted to an education.

The admission criterion of the GPA-based cutoff is solely driven by supply and demand and not by the ability needed for specific education programs. Instead, to assure a certain level of specific abilities, some education programs announce requirements for specific fields completed in high school. High school students can choose levels A–C, with A being the highest level. In this paper, we group fields of education into math fields, which we define as education programs that require the highest level of math from high school.

We group the math fields into different types of math using the definitions used by the Ministry of Education in Denmark. Table 1 shows our grouping of fields of education. The math fields are grouped into (1) nature math, which comprises fields, such as biology, math, chemistry, physics, and data science education; (2) tech math, which mostly includes engineering; (3) health math, which comprises medical school, odontology, pharmacy, and public health; (4) social science math, which includes different fields of economics. Finally, we have grouped all other fields that do not require math of level A. We use these groupings of fields when modeling the choice of education.

3 Data

3.1 Survey – eliciting beliefs

To elicit beliefs from Danish applicants for higher education, we conducted a survey among the applicants. The Danish Ministry of Education provided us with the cohort of applicants from 2018 and their rankings of educational programs. All the applicants received the survey after they had submitted their rankings to the central clearinghouse (as described in section 2). We wished to elicit the beliefs for the different education programs that an applicant had ranked in his or her application and were concerned that ex-post rationalizations might distort beliefs after the applicants learned about the education program they were admitted to (Brehm, 1956). This could, for example, cause an applicant that had been admitted to his or her second choice of education to rate that choice higher than his or her first choice. To avoid such problems, we finalized the survey just before the applicants would receive their admission letters. A sanity check of the data shows that less than 5 percent of the sample rate their second choice of education higher than their first choice when rating life satisfaction during and after studying.

To estimate the model, we need variation over the choice set, such that education $j \neq k$. To do so, we made use of the ranking structure in the applications. We inquired about the beliefs concerning the top two ranked education choices. For the applicants who had chosen one education program only, we asked for a hypothetical second choice of education from a list of all the possible field combinations and institutions available.

We used a combination of probability questions (see: Zafar, 2013; Boneva and Rauh, 2017; Manski, 2004) and Likert-scale questions ranging from 1-10 (for the use of Likert-scales in economics, see Benjamin et al., 2014; Falk et al., 2018). Following Benjamin et al. (2014), we asked all the questions specifically to the first, second and (if available) third choice. Because the choices were real and binding (i.e., when an applicant is offered a spot for his or her first ranked choice, he or she will not be offered a spot for the second ranked education), we are confident that the rankings represent the true preference ordering of the education programs.

3.1.1 Pecuniary questions

For pecuniary questions, we are especially interested in the expected returns 10 years after graduating from an educational program. We asked this question specifically for each of the programs that the individuals had applied for mainly because expectations for labor market returns will most likely differ between the choices. Specifically we asked the following question:

> *What do you think your typical monthly salary will be 10 years after graduating from education j [k] ?*

In Denmark, there is a large wage compression (Kahn, 2015), and we, therefore, hypothesized that the probability of employment would be an important input besides wages. The difference in employment is usually largest in the first two years after graduation. We, therefore, asked for the perceived probability of being in employment 1-2 years after graduation. Two years is also the limit for receiving unemployment benefits; therefore, after two years, many will simply change the sector to seek employment elsewhere. We specifically wanted to ask about full-time jobs since many take up part-time jobs in case they cannot find a full time job. Specifically we asked:

> *In percentage terms, what do you think your chance is that you will be employed in a full-time job 1–2 years after graduating from each of the education programs (j or k) below? Zero percent means there is no chance, and 100 percent means you are sure you will have a full-time job*

3.1.2 Non-pecuniary questions

Several studies have shown the importance of non-pecuniary factors in the choice of education (Zafar, 2013; Delavande and Zafar, 2019; Boneva and Rauh, 2017; Benjamin et al., 2014), ranging from questions about how the school curriculum is consistent with one's beliefs to the importance of location for one's partner.

Our survey, therefore, asked a number of non-pecuniary questions. For both the short run (while studying) and long run (10 years after graduation), we elicited the applicants' expectations about well-being related to job satisfaction, quality of personal life, and an overall measure of life satisfaction. We asked the questions beginning with the following introduction:

> *The next set of questions asks you to think about your life during your education [10 years after graduation]. Think about what you expect your life to be like during each education program. Each question asks you how satisfied you expect yourself to be with a different aspect of your life on a scale from 1 to 10. One always represents very dissatisfied, and 10 always represents very satisfied.*

Further, we also asked the individuals about their expectations for having children 10 years after graduation. This question is inspired by the evidence that having children influences one's career decisions (Nielsen et al., 2004; Wiswall and Zafar, 2017) and earnings (Kleven et al., 2018).

We take the overall life satisfaction while studying as a proxy for the expected consumption value of studying a specific education program, because it comprises all the possible inputs mentioned above. For the long run measures, we focus on work satisfaction, personal satisfaction, and the probability of having children.

Table 2 shows the questions and the scales used for the different questions (see Appendix E for the full questionnaire).

Table 2: Survey Questions

Question	Answer scale
During studies (or right after)	
(1) Your overall life satisfaction	1-10
(2) What do you think is the percent chance that you will be employed in a full time job 1-2 years after graduating from each of the education programs?	0-100 %
10 years after graduation	
(3) If you graduate, what is the percent chance that 10 years after graduating you will have children?	0-100 %
(4) Your health (physical and mental health)	1-10
(5) Satisfaction with your work (challenging, interesting, stress, etc.)	1-10
(6) Your satisfaction with your personal life (social, romantic, family relationships etc)	1-10
(7) Your satisfaction with your financial security	1-10
(8) What do you think your typical monthly salary will be 10 years after graduating from the education programs?	Number

3.2 Administrative register data

We are able to link the survey to administrative register data at Statistics Denmark. Therefore, we have individual links to background characteristics such as age, gender, ethnicity, high school grades, and parental education. We use gender directly in our choice model, to estimate the differences in the males' and females' choice of education. We use the other background characteristics to understand if male and female applicants are significantly different in their high school grades, for example. We do not wish to have the gender differences to be determined by some other characteristics that differs between the genders.

We do also obtain data on applicants for higher education in the years 1998–2001 and their earnings 15 years after applying for university. We use their earnings 15 years after graduation because we want the earnings to be as comparable as possible to the survey elicited expected beliefs about income 10 years after graduation. The higher education programs are 2–6 years long, which brings the earnings 15 years after applying for education as close as possible to the elicited expected earnings 10 years after graduation for the applicants from the survey. We group the actual earnings of the previous cohorts of applicants into the same field categories that we use for the survey respondents. We use these data to compare the actual earnings within each field to the survey respondents' expectations for future earnings within the same field category.

3.3 Sample selection

There were 76,369 applicants for higher education in 2018, and we invited all of them to participate in our survey. Of this number, 17,964 participated and answered at least 80 percent of the survey questions, which corresponds to a response rate of 24 percent. Compared to the full cohort of applicants, there are several differences between the survey respondents and non-respondents. The survey respondents are over-represented by females and applicants for university programs that would last 5 or 6 years (rather than 2 or 3.5 year college programs). We find no differences in age or number of programs applied for (see Appendix C for the full overview).

In this paper, we focus on the applicants who have applied for at least one math field in their ranked choices. We choose this group because they represent the applicants who have shown interest in math by attaining the highest level of math in high school and applying for at least one math field; thus, gender differences within this group cannot not solely be explained by different preferences for math. We need to restrict our sample to be able to estimate our choice model (see section 4.1). We cannot estimate the choice model for applicants that have only applied for education programs within the same field category. We need the choice set to consist of different ranked education programs. To ensure this variation over the choice set (i.e., education $j \neq k$), we only use applicants that have a different first and second choice based on the categories from Table 1. We also

Table 3: Gender differences on background characteristics

	Total	Females	Males	Gender diff.*
Age in year before applying for education	20.5	20.9	20.1	0.77***
High School GPA	8.42	8.68	8.16	0.51***
Father has education (HS or more)	0.90	0.90	0.90	-0.004
Mather has education (HS or more)	0.92	0.90	0.93	-0.022
Foreign origin	0.14	0.16	0.12	0.039**
Math on rank 1	0.73	0.72	0.74	-0.02
Observations	1,374	669	705	

*)Positive differences implies that values for females are larger than for males
Notes: The table reports means for the total sample, males and females. Stars indicate significant differences between male and female responses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

exclude all applicants that are 40 years or above; thus, our final sample comprises of 2,092 applicants who have two rankings out of the five field categories, with at least one rank being a math field. For the choice model estimations, we also exclude applicants who did not answer all survey questions that enters the utility function. This leaves us with an analysis sample of 1,374 applicants for higher educations, where 669 are females and 705 are male applicants.

3.4 Descriptive statistics

Table 3 shows the background characteristics of the male and female applicants. The table reports means and differences for the applicants' first choice of education. We find several differences between the genders. For example, females are on average about half a year older, have a higher GPA from high school, and are more often of foreign origin. However, there are no difference in parental education. There are also no immediate differences in the probability of choosing a math field as their preferred education. It is worth recalling that this sample is already restricted to the applicants who have applied for at least one math field in their top two ranked education choices. About three out of four applicants in this group have applied for a math field as their first choice. Yet, there are large differences

Table 4: Gender differences on the first choice field of study

	Females		Males	
	N	Percent	N	Percent
Nature math	147	21.97	150	21.28
Tech math	125	18.68	226	30.64***
Health math	170	25.41	84	11.91***
Social science math	40	5.98	72	10.21***
Non-math fields	187	27.95	183	25.96
Observations	669	100.0	705	100.0

Notes: Stars indicate significant differences between male and female responses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

in the type of math fields applied for by the female and male applicants. When looking at the different field categories (see Table 4), we notice that males and females prefer different fields of education. Males tend to apply more frequently for social science math and tech math, while females tend to apply more frequently for health math. For nature math and non-math field category, males and females apply almost equally.

3.5 Gender differences in elicited beliefs

In Table 5, we show the means of the elicited beliefs for the first and second choice education for males and females, respectively. We find considerable differences between males and females for expected income and the probability of having children for both first and second choice education. Males expect to have a monthly pre-tax income of DKR54,136 while females expect DKR44,499. On average, females expect to earn 17.8 percent less than males in their first choice of education and 20.6 percent in their second choice of education. For the expected probability of having children, males expect a 65 percent probability of having children, whereas females expect a 75 percent probability. We also find significant differences between male and female applicants' beliefs about personal satisfaction. The differences are much smaller than those pertaining to income and probability of having children. The beliefs about the different factors are on average always more significant for

Table 5: Elicited beliefs for first and second choice education. By gender

	First choice			Second choice		
	Total	Males	Females	Total	Males	Females
Pre-tax income 10 years after grad	49447.4	54136.0	44499.2***	47067.2	52302.7	41541.7***
Probability of employment efter grad	0.88	0.88	0.87	0.83	0.83	0.82
Life satisfaction while studying	8.46	8.45	8.47	7.30	7.27	7.34
Work satisfaction 10 years after grad	8.73	8.71	8.75	8.03	8.10	7.95*
Personal satisfaction 10 years after grad	8.49	8.38	8.61***	8.19	8.06	8.33***
Probability of children 10 years after grad	0.70	0.65	0.75***	0.69	0.64	0.75***
Observations	1,374	705	669	1,374	705	669

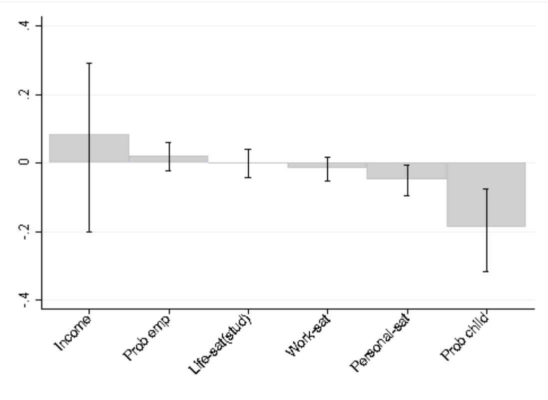
Notes: The table reports means for the total sample, males and females. Stars indicate significant differences between male and female responses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

first choice of education than the second choice.

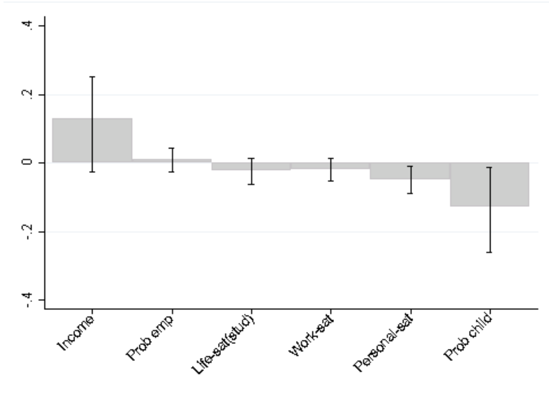
In Figure 1, we show the differences between the elicited beliefs of male and female applicants' first choice education. To compare the answers independent of scale, we calculate all differences relative to the average male answer. A positive difference means that males' expectations are higher than females'. Most notably, there are considerable gender differences in income expectations within fields. In social science math, for example, we find the largest difference where males expect to earn 34.5 percent more than females. Males applying for health math believe that they have a higher chance of employment; however, this is not significant for any of the other fields of education.

For the non-pecuniary questions, females state a higher probability of having children. This pattern is strong and significant for all fields of education except for social science math. The largest difference is in non-math fields where females have a 19.4 percent higher expected probability of having children than males. Females applying for nature math and tech math believe that they will have higher personal satisfaction than males.

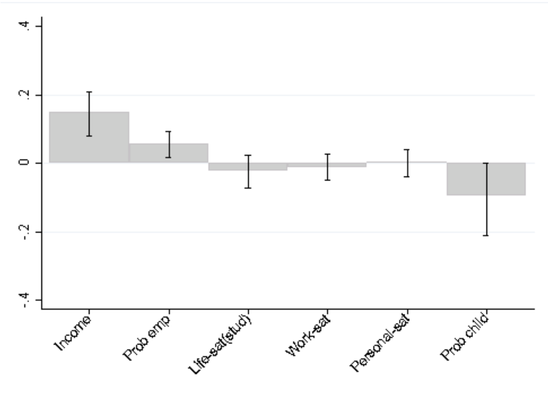
Figure 1: Relative gender differences for different fields



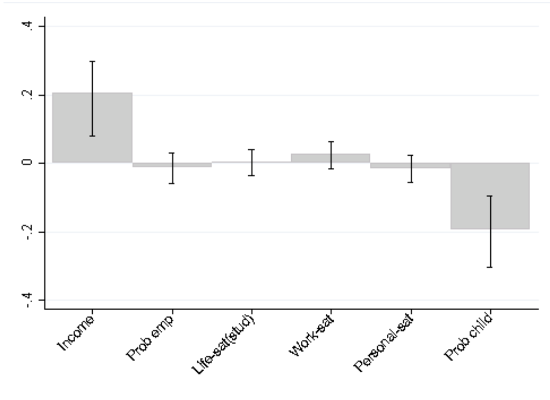
(a) Nature math



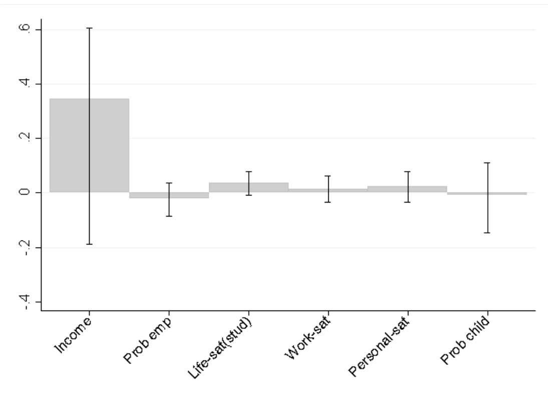
(b) Tech math



(c) Health math



(d) Non-math fields



(e) Social science math

Notes: The figure shows the relative gender differences in expectations for the different fields. The Y-axis measures relative differences of responses by males and females, that is: $\frac{male-female}{male}$.
The error bars are calculated as the 95% confidence interval on differences, divided by the confidence interval (high and low, respectively) for males

The remaining differences between males and females for the non-pecuniary beliefs are neither significantly different from each other nor substantial.

3.5.1 The elicited beliefs about the probability of having children

The elicited beliefs about the probability of having children can potentially suggest several things. It could suggest a preference for having children regardless of education choice, or it could illustrate the expectations of how doable it is to have children within a given field of education. If there is no variation in the expectations for having children across the education programs that individuals apply for, this could indicate that the variable captures an overall preference for having children. However, we cannot conclude that the education programs that the individuals did not list would have yielded a lower expected probability of having children for them (and that is why they did not apply for the education programs).

Variation across education programs concerning expectations for having children could suggest that it is a factor that applicants take into account in their choice of education. Table 6 shows that about 90 percent of both males and females have the same expectations of having children across their first and second choice education. Of the remaining 10 percent, about half reports a difference of 10 percentage points, and the other half reports a difference of more than 20 percentage points [3]. This suggests that about 10 percent of the sample directly factored in the probability of having children when choosing education. However, that does not mean that the remaining 90 percent do not take their preferences for children into account when choosing education. They could just choose not to apply for the education programs that they believe are hard to reconcile with their preferences for having children.

[3] In the survey, the applicants were asked to report their expected probability of having children on a 10-point scale, ranging between 0 percent and 100 percent; this means that the differences will by construction be 10 percentage points at a minimum.

Table 6: Difference in probability of having children across first and second choice educations

Difference	Male		Female	
	N	Percent	N	Percent
Difference=0	633	89.8	603	90.1
Difference=0.1	34	4.8	27	4.0
Difference larger than 0.2	38	5.4	39	5.8

Notes: Relative differences are calculated as $female_1 - female_2$ and $male_1 - male_2$

Taken together, our descriptive analysis of the choice of education of males and females reveals several fascinating insights for further analysis. First, we find statistically significant differences between the types of math fields that males and females apply for. Females tend to apply more frequently for health math, whereas males apply more frequently for tech math and social science math. Second, we find that females expect to earn significantly less than males both on average but also within specific fields. Besides, females expect a higher probability of having children than males. This difference is noticeable across all fields except social science math. Third, we find that a large proportion of the sample does not modify their expectations for having children across education choices. This could suggest that the factor captures an overall preference for having children among first and second choice education programs. However, we do not know if educations further down the rank- ordered list or the educations that the applicants chose not to apply for would be associated with a lower probability of having children.

4 The choice model

As stated in the previous section, applicants for higher education face the choice of rank ordering different education programs. As Arcidiacono et al. (2012), we assume that individuals specialize in higher education early in life and then after graduating enter the labor market for the remainder of their lives. In the Danish educational system this is a reasonable assumption because the majority of educations are specialized from the beginning. For example an applicant for engineering will apply at the bachelors level, and then continue the education until finishing and becoming an engineer after graduation.

Individuals choose among J education programs, which is a combination of field of study and institution. In this paper, we only focus on the choice of field and not on the choice of institution. We assume that applicants base their choice of educational program on the expected benefits and costs received while studying as well as the subsequent returns in the labor market. While studying, student i's utility for studying program j is a function of inputs such as ability and preferences for studying the specific field (Arcidiacono et al., 2012). We pool these factors together and define this as consumption value of studying.

Applicants also expect some future utility of graduating from an education program. We elicit a number of inputs for the utility function conditional on studying program j. First, there are pecuniary labor market returns. This we decompose into expected earnings for individual i in education j, and expected probability of employment. We also elicit a number of non-pecuniary inputs in the utility function. These are: Work-satisfaction, Personal life satisfaction, and expected probability of having children.

Applicants base their choice on pecuniary returns, $p = \{salary,\ probability\ of\ employment\}$ and non-pecuniary returns, $n = \{consumption\ value\ of\ studying,\ work\ satisfaction,\ personal\ satisfaction,\ probability\ of\ children\}$. Taken together, the subjective expected utility (SEU) of education program j for individual i is:

$$SEU_{ij} = U_i(p, n, X_i) \tag{1}$$

In the model we allow for heterogeneity based on observable characteristics, X, such as gender. We assume additive separability across pecuniary and non-pecuniary inputs, similar to Zafar (2013) and Boneva and Rauh (2017). This allows us to rewrite equation 1:

$$SEU_{ij} = \sum_{n=1}^{2} p_n \alpha_n(X_i) + \sum_{m=1}^{4} n_m \gamma_m(X_i) + \epsilon_{ij} \tag{2}$$

where p_n is the continuous pecuniary inputs and $\alpha_n(X_i)$ is a constant for the input p_n for an individual with characteristics X_i. Similarly, n_m is the continuous non-pecuniary inputs and $\gamma_m(X_i)$ is a constant for the input n_m for an individual with characteristics X_i. ϵ_{ij} is a random error term. With this formulation, the utility that individuals derive from (n, p) is the same for individuals with similar characteristics X up to the random error term. Both p_n and n_m are unknown at the time of applying for higher education, but individual i has subjective beliefs for the pecuniary inputs, $P_{ij}(p)$, and for the non-pecuniary inputs, $P_{ij}(n)$ for education program j. The expected utility for individual i of education program j can therefore be rewritten as:

$$SEU_{ij} = \sum_{n=1}^{2} \alpha_n(X_i) \int p_n dP_{ij}(p) + \sum_{m=1}^{4} \gamma_m(X_i) \int n_n dP_{ij}(n) + \epsilon_{ij} \tag{3}$$

For all inputs we elicit the expected value of each input instead of the full probability distribution, $\int p_n dP_{ij}(p) = E_{ij}[p]$ and $\int n_m dP_{ij}(n) = E_{ij}[n]$. From this we can rewrite equation 3:

$$SEU_{ij} = \sum_{n=1}^{2} \alpha_n(X_i) E_{ij}[p_n] + \sum_{m=1}^{4} \gamma_m(X_i) E_{ij}[n_m] + \epsilon_{ij} \tag{4}$$

The parameters that we need to estimate are $\{\alpha_n(X_i)\}_{n=1}^{2}$ and $\{\gamma_m(X_i)\}_{m=1}^{4}$. These are the parameters in the utility function for the pecuniary inputs, p_n, and the non-pecuniary inputs, n_m, for an individual with characteristics X_i. Individual i will choose education j over education k if the SEU of education j exceeds the SEU of education k, and the probability of individual i choosing education j can thus be written as:

$$\begin{aligned}
Pr(j|E_{ij}[p], E_{ij}[n]) &= Pr(SEU_{ij} > SEU_{ik}) \\
= Pr\Bigg(\Big(\sum_{n=1}^{2} \alpha_n(X_i)E_{ij}[p_n] &- \sum_{n=1}^{2} \alpha_n(X_i)E_{ik}[p_n] \Big) \\
+ \Big(\sum_{m=1}^{4} \gamma_m(X_i)E_{ij}[n_m] &- \sum_{m=1}^{4} \gamma_m(X_i)E_{ik}[n_m] \Big) \\
\geq \epsilon_{ik} - \epsilon_{ij} \Bigg), &\quad \forall\, j \in J,\ j \neq k
\end{aligned} \tag{5}$$

4.1 Choice model estimation

Under the assumption that the error terms, ϵ_{ij}, are independent for each individual, i and each education, j, and have a type I extreme value distribution, then $\epsilon_{ik} - \epsilon_{ij}$ has a standard logistic distribution. If the expected utility of individual i of choosing education j can be expressed as $SEU_{ij} = \beta x_{ij} + \epsilon_{ij}$, then the probability of choosing education j can be expressed as:

$$Pr_i(j) = \frac{exp(\beta x'_{ij})}{\sum_j exp(\beta x'_{ij})} \tag{6}$$

We allow for heterogeneity by gender and estimate the model separately for males and females. For estimation we use the elicited expected values of each of the inputs, and the parameters of interest are $\{\alpha_n(X_i)\}_{n=1}^{2}$ and $\{\gamma_m(X_i)\}_{m=1}^{4}$ from equation 5. In the model it is only differences in the elicited beliefs across the ranked education programs that is used for estimation. The model is estimated using maximum-likelihood.

As we do not observe the full choice set (each applicant applying for an education within the five education groups), our estimates are likely to be biased. In a simulation exercise, we find that our estimates are biased towards zero, and are estimated with more noise, resulting in higher standard errors, than if we had the full choice set (see appendix B for simulation and section 5.4.1 for a discussion of the repercussions).

4.2 Decomposition analysis

We perform a decomposition to understand the importance of different factors of choice of education for males and females. We use the same approach as Zafar (2013) who developed a decomposition method to explain the relative importance of the different factors in a similar model. We cannot say anything about the unobserved part of the model for either males and females, but the decomposition explicate how much of the observed variation in choice of the utility function for each gender is explained by specific factors in the utility function. Suppose that our choice model includes only two factors: x_1 and x_2, then given the parameter estimates β_1 and β_2, the contribution to the variation in choice from x_1 can be written as:

$$C_{x_1} = \sqrt{\sum_j \left[\sum_{n=1}^{N} \frac{(Pr(choice = j|\hat{\beta}_1, \hat{\beta}_2)}{N} - \sum_{n=1}^{N} \frac{(Pr(choice = j|\tilde{\beta}_1 = 0, \tilde{\beta}_2)}{N} \right]^2}$$

where the first term on the right hand side is the average predicted probability of choosing education j predicted by our choice model in equation 4 and the second term is the average predicted probability of choosing education j in a choice model, where the variable x_1 is not considered. The relative contribution of the variable x_1 to the choice can be calculated as $\frac{C_{x_1}}{C_{x_1}+C_{x_2}}$. For determining the significance level of each of the contributions we bootstrap the standard errors based on 1,000 repetitions.

Using this decomposition method allows us to calculate the contribution to the choice of education of several factors at the same time. Therefore, we divide the factors of our choice model into three categories: 1) consumption value of studying (life satisfaction while studying), 2) future pecuniary factors (pre-tax income and probability of employment), and 3) future non-pecuniary factors (work satisfaction, personal satisfaction and probability of having children).

5 Results

Table 7 shows the estimation results of our choice model presented in section 4. Column 1 displays the estimates for the full sample; in columns 2 and 3, we allow for heterogeneity by gender and estimate the choice model for subsamples of females (in column 2) and males (in column 3). For the total sample, we find that life satisfaction while studying, income, probability of employment, and work satisfaction are positive and significant determinants of the choice of education.

For females, we find that life satisfaction while studying, income, probability of employment (significant at a 10 percent level), and work satisfaction are determinants of education choice. Furthermore, we find that the probability of having children is associated with the choice of education for females (significant at a 10 percent level). For males, we find that life satisfaction while studying, work satisfaction, and the probability of employment are determinants of education choice. Neither income nor the probability of having children seem to have an effect on the choice of education for males.

Overall, both expectations for pecuniary and non-pecuniary factors are significant drivers of choice for both males and females. We find differences between the two genders in their education choices, with future pre-tax income and the probability of having children 10 years after graduation being associated with the female choice of education and not that of the male choice. To understand the importance of each of the factors, we

Table 7: Choice Model Estimation

	All	Females	Males
Life satisfaction while studying	0.865***	0.809***	0.921***
	(0.0840)	(0.127)	(0.109)
Pre-tax income (log)	0.607**	1.798***	0.256
	(0.304)	(0.643)	(0.337)
Probability of employment	0.220***	0.188*	0.237***
	(0.0669)	(0.0991)	(0.0895)
Work satisfaction 10 years after grad	0.351***	0.424***	0.274**
	(0.0801)	(0.112)	(0.117)
Personal satisfaction 10 years after grad	0.141	0.251	0.0540
	(0.161)	(0.220)	(0.243)
Probability of children 10 years after grad	0.177	0.293*	0.0918
	(0.109)	(0.172)	(0.154)
Nature math	-0.397***	-0.521***	-0.343**
	(0.116)	(0.175)	(0.161)
Tech math	-0.0874	-0.305	-0.0205
	(0.128)	(0.209)	(0.163)
Health math	0.176	0.0463	0.0766
	(0.160)	(0.204)	(0.284)
Social science math	0.0333	-0.259	0.180
	(0.162)	(0.260)	(0.210)
Observations	1,374	669	705

Notes: The omitted category is ”Non-math fields”. Standard errors clustered on individual level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

perform a decomposition of the results, presented in the following section.

5.1 Decomposition

Table 8 presents the results of the decomposition exercise for the full sample (column 1), females (column 2), and males (column 3). The results show that the consumption value of studying has the largest effect on males' choice of education, explaining 44 percent of the variation in the males' choice compared to 23 percent of the females' choice. We find that the importance of pecuniary factors is strikingly similar for males and females. The pecuniary factors explain about 26 percent of the variation in the choice of education for females and 22 percent for males. Pecuniary factors are significant at a 10 percent level for both males and females. The non-pecuniary factors 10 years after graduation are more important for the choice of education for females than males. The non-pecuniary factors explain 51 percent of choice for females and only 34 percent for males. This result is different from the finding of Zafar (2013) where males attributed more weight to pecuniary factors than females. In our model, males and females attach almost equal importance to pecuniary factors, despite varying degrees of attention to the consumption value of studying versus the future non-pecuniary benefits. The consumption value of studying explains 44 percent of the variation in males' choice and only 23 percent in females' choice. The future non-pecuniary factors explain half of the variation in females' choice compared to only 34 percent of that among males.

5.2 Understanding gender differences

Our results so far show two main differences in males' and females' education choices: differences in the role of expected earnings and differences in the role of the expected probability of having children. We, therefore, investigate how these expectations would

Table 8: Decomposition analysis by gender

	All	Females	Males
Life satisfaction while studying	0.123	0.232**	0.443***
	(0.088)	(0.101)	(0.165)
Pecuniary factors	0.300**	0.261*	0.216*
	(0.122)	(0.139)	(0.123)
Nonpecuniary factors	0.577***	0.506***	0.342**
	(0.140)	(0.140)	(0.168)
Observations	1,374	669	705

Notes: Bootstrap standard errors in parentheses calculated from 1,000 bootstrap repetitions
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

alter the gender gap in education choice if we assume that females have the same expectations as males. Furthermore, we conduct a heterogeneity analysis on subsamples with different expectations for having children 10 years after graduation.

5.2.1 Changing expectations

In this section, we investigate how changes in expectations would affect the gender gap in the choice of education. First, we ask how gender differences in the choice of education would change if females had the same income expectations as their male counterparts. From Figure 1, we know that males and females have different expectations for future income. For all the field categories, males expect a higher income than females. On average, males expect to earn 17.8 percent more than females for their first choice of education. Across the fields, we find that the narrowest gender gap in expected earnings is in tech math, and the widest gap is in social science math of more than 34 percent. Second, we ask how gender differences in the choice of education would change if females had the same expectations for having children as males. We ask this because there are large differences in these expectations across gender. Furthermore, the expected probability of having children enters the utility function of females positively.

To answer these two questions, we perform a forecast analysis where we replace the average female expectations for each specific education group with the average male expectation. We then calculate the predicted probabilities of choosing education j using the estimated parameters from Table 7. The predicted probabilities are evaluated at the sample averages within each combination of first choice j and second choice k for the subsamples of males and females. We assume that the top two chosen education programs by each applicant are the only two options that are relevant to the applicant. This implies that by changing females' expectations of future income, their choice of education could only switch between their first and second choice. The choice of education cannot switch to an education program that the applicant did not apply for. See Appendix D for details on how we calculate the choice probabilities when we only observe two ranked choices.

Table 9 shows the forecast analysis. Column 1 shows the gender gap in the predicted probability of choosing each field for the baseline model presented in Table 7. Column 2 shows the gender gap when setting the average female income expectation equal to the average male expectation; column 3 does the same thing but for the probability of having children. Finally, column 4 presents the predicted gender gap when setting both average expectations for income and the probability of having children for females equal to the male averages. The gender gap is defined as the male predicted probability subtracted from the female predicted probability of choice; thus, a positive coefficient correspond to more males than females, and a negative coefficient corresponds to more females than males. The predicted probabilities at baseline illustrate that there is a 9.7 percentage point higher probability of males choosing a tech math education as their first choice of education compared to females. Similarly, males are more likely to choose social science math. Females are 12.7 percentage points more likely to choose health math as their first choice than males. We do not find any statistically significant gender gaps within nature math fields or non-math fields. The bottom segment of Table 9 presents the differences

Table 9: Forecasting the gender gap: Setting female expectations equal to male expectations for income and children

	Baseline	Income	Children	Income + Children
Nature math	-0.00725	-0.00684	-0.00549	-0.00508
	(0.0190)	(0.0190)	(0.0185)	(0.0185)
Tech math	0.0967***	0.103***	0.0957***	0.102***
	(0.0199)	(0.0198)	(0.0196)	(0.0194)
Health math	-0.127***	-0.127***	-0.129***	-0.129***
	(0.0188)	(0.0188)	(0.0187)	(0.0187)
Social science math	0.0404***	0.0411***	0.0389***	0.0394***
	(0.0129)	(0.0129)	(0.0130)	(0.0129)
Other fields	-0.0180	-0.0162	-0.0139	-0.0120
	(0.0211)	(0.0205)	(0.0209)	(0.0204)
Observations	1,374	1,374	1,374	1,374
Differences in percent from baseline				
Nature math		5.697	24.368	29.917
Tech math		-6.177	1,090	-4,936
Health math		-0.073	-1.602	-1.668
Social science math		-1.757	3.742	2.302
Non-math fields		10.283	22.762	33.221

Notes: The top part of the table shows the predicted gender gap associated with each field (male probability - female probability). The bottom part of the table show the percent change in the gender gap relative to the baseline. A positive percentage corresponds to a decrease in the gender gap. Bootstrapped standard errors in parentheses. Based 1,000 bootstrap repetitions. *** p<0.01, ** p<0.05, * p<0.1

in percentage from the baseline gender gap. Generally, the changes in the gender gap are small when equalizing the expectations for income and the probability of having children. Females opt out of tech math when equalizing income between males and females, and females are more likely to choose social science math when equalizing the expected probability of having children, although the differences compared to the baseline gender gap are small. The gender gap in tech math increases by 6.2 percent by equalizing the male and female income expectations, and the gender gap decreases in social science math by 3.7 percent when equalizing the expectations for having children.

As stated above, we cannot forecast changes in the choice of education for educations that the applicants did not apply for. We only observe two choices out of the five fields, which means that we can only forecast how the choice would change within the two observed choices. And since there is a gender gap in the listed choices as well as in the first choice fields, there will to some extent still exist a gender gap regardless of how much we change the females' expectations. Therefore, the results are conservative estimates. Despite this, we still find that equalizing females' and males' expectations for both income and having children would encourage more females to choose social science math. This result is driven solely by the probability of having children where giving females the males' expectation of having children decrease the gender gap (with a male majority) by 3.7 percent.

5.2.2 Heterogeneity in expectations for having children

In the choice model results from Table 7, we find that the expected probability of having children is statistically significant and positively correlated with the choice of education for females, but not for males. We also show descriptively in Figure 1 that there are large differences when comparing the average females' and males' expectations of having children. Generally, males expect to have children to a much lower degree than females.

As discussed in section 3.5.1, it is difficult to distinguish if the expectations of having children is a general preference for having children or a determinant of the specific choice of education. Regardless of whether the expectations capture one or the other (or most probably a combination of the two), we wish to understand whether there are any differences among the applicants of the same gender depending on the expectations of having children. In this section, we present our choice model estimations from section 4.1 for subsamples of males and females that have different expectations of having children. We divide the sample into a group of males and females who expect a 90 percent or higher probability that they will have children from both their first and second choice education and a group of males and females who expect an 80 percent or lower probability of having children. About half of the females expect 90 percent or a higher probability of having children; for the males, this proportion is just under one-third of the sample.

Table 10 shows the choice of education for each of these groups. Columns 1 and 2 present females' choice of education where the first column "High" includes females with a high expected probability of having children and column "Low" includes females with a lower expected probability of having children. Similarly, columns 3 and 4 show the results for males. The choice of education across the probability of having children is significantly different for females.

The proportion of the applicants who apply for social science math as their first choice is significantly larger for females who have a lower expected probability of having children. Furthermore, the proportion of the applicants who apply for non-math fields is significantly lower for females who have a lower expected probability of having children compared to females with a high expected probability of having children. The absolute difference is large: twice as many females choose social science math in the sample of females with a lower probability of having children compared to females who expect a higher probability. The proportion of females who choose non-math fields are 28 percent

Table 10: First choice education by the probability of having children 10 years after graduation

	Females		Males	
	High	Low	High	Low
Nature math	0.21	0.24	0.2	0.22
Tech math	0.17	0.2	0.28	0.31
Health math	0.26	0.25	0.15	0.11*
Social science math	0.04	0.08**	0.12	0.1
Non-math fields	0.32	0.23***	0.25	0.26
Observations	330	307	468	219

Notes: Stars indicate significant differences between male and female responses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

smaller for the low probability of having children compared to the females with a high probability. This suggests that the females who reported a high expected probability of having children for both ranked choices of education are choosing more non-math education programs than the females who have a lower expected probability of having children. To make sure this is not driven by higher math abilities for the females who have a lower probability of having children, we check if there are any differences in high school math grades between the groups and find no differences across the groups (the results are available on request).

For males, the choice of education displays a more similar pattern across the expected probability of having children. We only find a difference in the proportion of males who apply for health math (significant at a 10 percent significance level).

The estimation results of our choice model for subsamples by gender and the expected probability of having children are presented in Table 11. Columns 1 and 2 present the results of females where the first column "High" contains the estimates for females with a high expected probability of having children and column "Low" contains the estimates for females with a lower expected probability of having children. Similarly, columns 3 and 4 show the results for males.

Table 11: Choice model estimates: By gender and expected probability of having children

	Females		Males	
	High	Low	High	Low
Life satisfaction while studying	1.080***	0.719***	1.017***	0.935***
	(0.159)	(0.181)	(0.249)	(0.122)
Pre-tax income 10 years after grad (log)	0.901	2.322**	-0.835	0.308
	(1.033)	(0.956)	(1.207)	(0.371)
Probability of employment	-0.0342	0.352***	0.615***	0.146
	(0.200)	(0.111)	(0.225)	(0.106)
Work satisfaction 10 years after grad	0.564***	0.526***	0.319	0.266*
	(0.184)	(0.169)	(0.214)	(0.147)
Personal satisfaction 10 years after grad	0.681	-0.356	-0.238	0.240
	(0.459)	(0.397)	(0.456)	(0.329)
Nature math	-0.429	-0.712***	-0.00600	-0.404**
	(0.285)	(0.237)	(0.336)	(0.201)
Tech math	-0.342	-0.402	0.173	-0.0929
	(0.342)	(0.282)	(0.311)	(0.204)
Health math	0.198	0.0155	0.208	0.329
	(0.352)	(0.300)	(0.512)	(0.383)
Social science math	-0.354	-0.173	0.921**	-0.0445
	(0.447)	(0.350)	(0.399)	(0.258)
Observations	307	330	219	468

Notes: The omitted category is ”Non-math fields”. Standard errors clustered on individual level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

For females with a high expected probability of having children, only the consumption value of studying and future work satisfaction influence their choice of education. For the females with lower expectations of having children, we also find that the consumption value of studying and future work satisfaction matter for the choice of education, furthermore, the expected future income and probability of employment of this group of females also have a significant impact on their choice. These results suggest that the significant effects of pecuniary factors for females from our main specification in Table 7 is driven solely by the group of female applicants who have a lower expectation for having children 10 years after graduation. The results for males illustrate that those with a high expected probability of having children have the consumption value of studying and the probability of employment as significant factors for their choice of education. Males with a lower expected probability of having children also appreciate the consumption value of studying; however, they do not seem to take any pecuniary factors into account but, instead, find future work satisfaction as an important factor of choice.

5.3 Investigating errors in income expectations

The literature on the choice of education generally finds errors in expected earnings when comparing them to actual earnings (Baker et al., 2018; Wiswall and Zafar, 2014; Jensen, 2010). Because expected earnings is a significant determinant of choice of education we investigate whether the applicants in our sample could have wrong income expectations, as suggested by the previous literature. In case the answer should be positive, we investigate if their choice of education would change when replacing their expectations with actual average income within each field of education.

We compare our elicited income expectations with income from previous cohorts of applicants for higher education. We obtain income data from applicants for higher education in 1998-2001 and calculate their average monthly pre-tax income 15 years after applying for higher education to approximate our elicited beliefs of average monthly pre-tax income 10 years after graduation. We compare our elicited expectations for monthly pre-tax income 10 years after graduation with the monthly pre-tax income from previous cohorts of applicants 15 years after they had applied for higher education [4]. We deflate all earnings to be expressed in 2015 price levels. Figure 2 shows that both males and females have a higher expected income than the actual income of the previous cohorts. The differences are most pronounced in social science math and nature math and smallest for tech math and health math.

We also find that males and females have different actual earnings, with females earning significantly less than males. The largest gender difference is in social science math, where females earn 24 percent less than males, and the smallest gender gap is in nature math, where females earn 15 percent less. Gender differences are also reflected in the expectations for earnings where females expect lower earnings compared to males. The largest gender gaps are in social science math and non-math fields, where females expect to earn 35 percent and 22 percent less than males. The smallest difference is in health math, where females expect to earn 11 percent less than males.

Because we find large differences between actual income and the income expectations in our sample of applicants, we carry out a simulation exercise similar to the forecast analysis from the previous section. Here, we correct the errors in beliefs by replacing the average income expectation within each field of study category with the actual average income within each field category for the previous cohorts of applicants for higher education.

Tables 12 and 13 show the forecast analysis for females and males, respectively. Column 1 presents the predicted probabilities for choosing each field category using the

[4] The actual income measure is associated with some uncertainty. There are a fairly large number of education programs that existed in the years 1998-2001 that do not exist today, and also a number of new education programs of today that did not exist in 1998-2001

Table 12: Setting female expected income equal to actual income of previous female cohorts

	Expected income	Actual income	Difference
Nature math	0.161***	0.152***	-0.00873***
	(0.0139)	(0.0135)	(0.00212)
Tech math	0.156***	0.156***	0.000272
	(0.0135)	(0.0135)	(0.000976)
Health math	0.230***	0.235***	0.00559***
	(0.0160)	(0.0162)	(0.00133)
Social science math	0.0526***	0.0512***	-0.00136
	(0.00846)	(0.00826)	(0.000907)
Non-math fields	0.212***	0.197***	-0.0154***
	(0.0162)	(0.0155)	(0.00260)
Observations	669	669	669

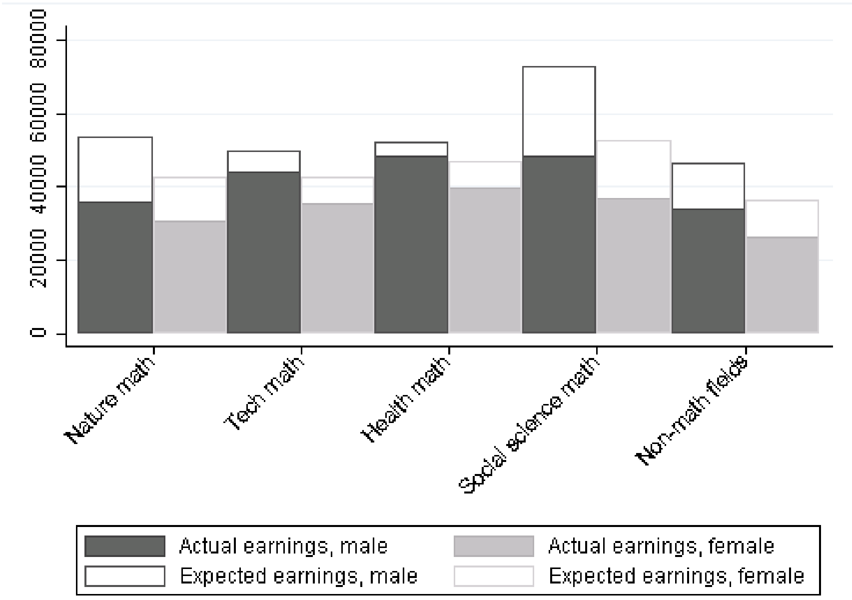
Notes: The table shows the predicted female probability of choosing each field. Column 3 presents the difference between column 1 and column 2. Bootstrapped standard errors in parentheses. Based 1,000 bootstrap repetitions. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 13: Setting male expected income equal to actual income of previous male cohorts

	Expected income	Actual income	Difference
Nature math	0.156***	0.141***	-0.0143***
	(0.0131)	(0.0127)	(0.00248)
Tech math	0.242***	0.254***	0.0124***
	(0.0157)	(0.0154)	(0.00282)
Health math	0.109***	0.111***	0.00156**
	(0.0110)	(0.0110)	(0.000694)
Social science math	0.0948***	0.0946***	-0.000150
	(0.0103)	(0.0103)	(0.000721)
Non-math fields	0.199***	0.190***	-0.00923***
	(0.0153)	(0.0146)	(0.00259)
Observations	705	705	705

Notes: The table shows the predicted male probability of choosing each field. Column 3 presents the difference between column 1 and column 2. Bootstrapped standard errors in parentheses. Based 1,000 bootstrap repetitions. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 2: Actual and expected earnings for different fields



Notes: The figure shows the actual earnings and expected earnings by gender.

elicited expected income. In column 2, we substitute the average expected income for each field category with the actual average income within that field. Column 3 shows the difference between columns 1 and 2. Focusing on column 3 of Table 12, we find statistically significant differences in the predicted probability of choosing nature math, health math, and non-math fields for females. Adjusting for actual income decreases the probability of choosing nature math by 5.6 percent and non-math fields by 7 percent. It increases the probability of choosing health math by 2.2 percent. Looking at the income adjustment for males in Table 13 we find statistically significant differences in choice by correcting expected income with the actual income for all field categories except social science math. The predicted probability of choosing nature math and non-math fields decrease by 9.6 percent and 4.5 percent. The predicted probability of choosing tech math increases by 5 percent and 1.8 percent for health math.

The findings from Table 12 and 13 show that correcting the income expectations for males and females would cause a small group of applicants to choose differently. The patterns are almost similar across the genders: both males and females would opt out of nature math and non-math fields, whereas more will opt in for health math. Finally, more males would choose tech math. Correcting the expectations for income does not seem to explain the gender differences in education choice. Again, this forecast analysis is based on the very restrictive assumption that the top two chosen education programs by each applicant are the only two choices that are relevant for the applicant. This means that the results should be interpreted with some cation.

5.4 Sensitivity checks

There are three threats to the identification of our choice model estimates. First, we do not observe the complete choice set of educations but only the first two ranked choices. Second, as explained in section 2, some education programs have a binding GPA cutoff for admission, which is based on supply and demand for a given education program. Therefore, given the GPA cutoff, not all the applicants face the same choice set. A high school graduate with a GPA of 8 would not perhaps have considered a program with a GPA cutoff of 11 in the previous years. Third, we could have bias from omitting a determining variable. Specifically, the literature on gender differences in education and earnings suggests that abilities, competitiveness, and risk preferences differ between the genders and tend to influence the choice of education. In the following sections, we discuss each of these threats to identification.

5.4.1 Incomplete choice set

Most applicants did not apply for more than two education programs even though they could apply for eight programs. This indicates that the education programs not listed are associated with a much lower expected utility than the ones that the applicants had applied for. In fact, the utility of the education programs not listed must be lower than the outside option of not applying for higher education. However, when we estimate a choice model with five possible choices with individuals who have only listed two of the five, we worry that our estimates will be biased. To investigate this, we conduct a Monte Carlo simulation whereby we simulate a choice model similar to ours and compare it to a model where we observe the full choice set (see Appendix B for details). The conditional logit estimates of the two simulated data sets are different in size. The estimates from the choice model on data with only two observed choices is biased toward zero. A second simulation using the same, albeit negative, coefficients confirms that the estimates are biased toward zero rather than being negatively biased. The estimates, while having only two observed choices, are estimated with more noise and hence yields larger standard errors. As our parameter estimates are biased toward zero and have larger standard errors than if we had data on the full choice set, the results in this paper are conservative estimates of the gender differences in the choice of education.

The existing literature on education choices mostly uses small and convenient samples from a single university where they can ask students about hypothetical choices of all educations to obtain a full choice set (For example: Arcidiacono et al., 2014; Zafar, 2013). As we have just shown in our simulation study, these samples benefit from a richer choice set compared to ours. However, our data are more accurate because our choices are actual choices from the applicants who have just gone through the process of applying and researching on important factors of choice. Furthermore, asking students to share their beliefs about many hypothetical education programs is likely to cause considerable uncertainty about the validity of the responses.

Table 14: Sensitivity check: Restricting applicant GPA to top quartile

	All	Females	Males
Life satisfaction while studying	1.111***	0.897*	1.894***
	(0.359)	(0.459)	(0.517)
pre-tax income 10 years after grad (log)	0.908	0.953	1.751
	(0.977)	(1.785)	(1.303)
Probability of employment	0.496***	0.529***	0.289
	(0.179)	(0.184)	(0.292)
Work satisfaction 10 years after grad	1.286***	0.949**	2.550**
	(0.415)	(0.391)	(1.286)
Personal satisfaction 10 years after grad	0.00469	0.0975	-0.160
	(0.318)	(0.518)	(0.533)
Probability of having children	0.435	0.452	0.471
	(0.470)	(0.780)	(0.498)
Nature math	-0.898**	-0.626	-1.640**
	(0.379)	(0.538)	(0.708)
Tech math	-1.188**	-1.019	-1.624**
	(0.493)	(0.666)	(0.780)
Health math	-0.340	-0.0300	-1.367
	(0.470)	(0.537)	(1.036)
Social science	0.213	0.138	0.0377
	(0.433)	(0.744)	(0.665)
Observations	450	230	220

Notes: The omitted category is ”Non-math fields”. Standard errors clustered on individual level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

5.4.2 Gender differences for the top GPA quartile

An important sensitivity check involves comparing male and female applicants with a similar high school GPA for the following two reasons: (1) the descriptive statistics show that females have a statistically significant higher GPA than males and (2) the admission GPA cutoffs into higher education could make applicants with different GPAs face different choice sets of education. Therefore, Table 14 presents choice model estimates of a subsample of male and female applicants who have a high school GPA in the top quartile of the GPA distribution in the sample (corresponding to a GPA of at least 10.2). If there are no gender differences in this sample where high school GPA and choice set are equal for males and females, then it suggests that the gender differences from our choice model are driven by the differences in high school GPA.

The sensitivity check in Table 14 illustrates that future work satisfaction and life satisfaction while studying are statistically significant determinants of choice for both males and females. The probability of employment only enters the utility function of females' choice of education and not into the male utility function. Thus, we conclude that when comparing males and females with high and similar GPA from high school, we still find gender differences in the choice of education. This suggests that gender differences in education choice are not solely driven by differences in GPA scores of males and females.

5.4.3 Abilities, competitiveness, and risk preferences

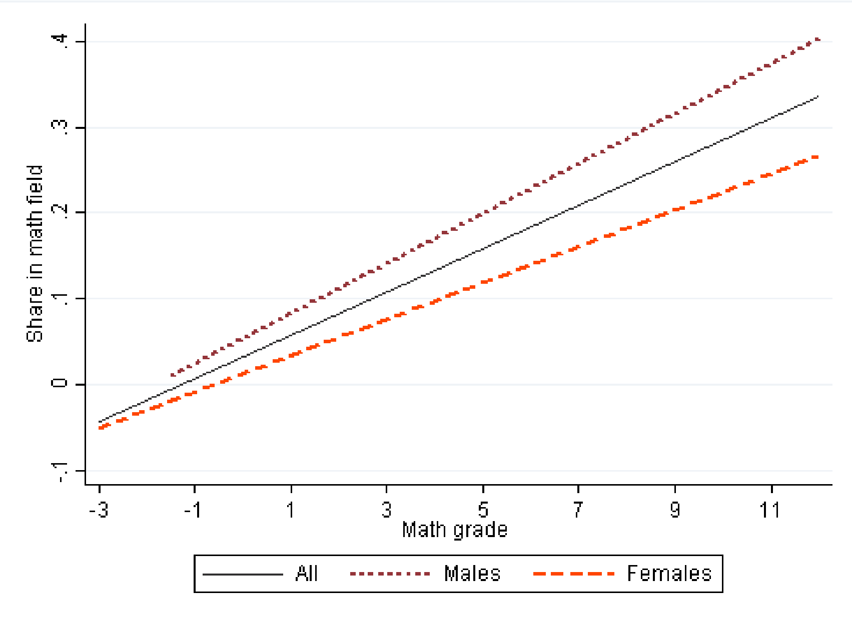
The existing literature on gender differences presents several possible explanations for why males and females choose different education programs. In this section, we discuss how ability, competitiveness, and risk preferences influence the choice of education for both genders in our sample.

Abilities. One driver of males choosing math educations to a higher degree than females could be a gap in (innate) math abilities. If males are better at math, it is natural that males choose math fields more so than females. That males are better at math than females is a persistent stereotypical belief (Reuben et al., 2014; Nosek and Smyth, 2011). This stereotype has been nurtured by the fact that males have been performing better in math for many years. Yet, females' performance in math is now equal to that of males in many countries (Guiso et al., 2008), and there are no signs of innate gender differences in math ability. Examining gender differences in the US math tests for grades 2–11, Hyde et al. (2008) find no differences on average. However, they notice that the variability is slightly larger for males, meaning that there are slightly more males in the 99th percentile than females. This result only holds for white students, while it is reversed for Asian/Pacific Islander students. In the Danish context, on average, females perform slightly better than males across almost all math assessments. This pattern is seen both in elementary school and high school (Eriksen and Kjærsgård-Jensen, 2018).

In our sample, we find that females have slightly better math abilities, as measured by math grades in high school. Nonetheless, we find that over the entire math grade scale, males are more likely to choose a math field when applying for higher education. In Figure 3, we plot the share of females and males who choose a math field given a certain math grade from high school (grades in Denmark are distributed from -3 to 12). The figure shows that males are more likely than females to choose a math field independently of their high school math performance. In the Danish setting, math abilities do not seem to drive the gender gap in the choice of education.

Competitiveness and risk preferences. Two other well-established differences between males and females are risk preferences and competitiveness (Croson and Gneezy, 2009). For competitiveness, a long line of research originating from Niederle and Vesterlund

Figure 3: Share of females and males choosing a math field on their first priority



Notes: The figure shows the share of males and females applying for a math field on their first priority, by math grades. The lines are constructed by making a best line of fit by math grade and share applying to math.

(2007) shows that females are typically less competitive than males. This is hypothesized to affect vertical segregation in the labor market, where only males pursue high-profile competitive jobs as managers or CEO's. Competitiveness can also affect horizontal segregation in the labor market, which is influenced by the choice of education. Specifically, Buser et al. (2014) show that competitiveness is predictive of choosing a prestigious math-related high school track in the Netherlands. The math-related track is more competitive in the Netherlands, and competitiveness of both males and females is highly predictive of whether students choose a math related high school track. Having competitiveness as a measure thereby explains some part of the gender difference in educational choices.

We do not have a direct measure of competitiveness or know of any studies on perceived competitiveness in different fields of education in Denmark. In Buser et al. (2014)'s work, the most prestigious study tracks are chosen by students with the highest GPA, and

Table 15: The educations with the highest GPA cutoffs

University	Field of study	GPA cutoff	Math A req.	Female share
Copenhagen Business school	International Business	12.2	No	20.8
University of Aarhus	Cognitive Science	11.8	No	52.9
Copenhagen Business school	International Business and politics	11.7	No	42.0
University of Copenhagen	Psychology	11.7	No	75.1
University of Copenhagen	Medicine	11.4	Yes	66.9
University of Copenhagen	Molecular Bio-medicine	11.4	Yes	68.8
College Metropol	Midwife	11.4	No	97.7
University of Aarhus	Psychology	11.4	No	78.3
Copenhagen Business school	Business Economics, and project-mng	11.3	No	42.7
Zealand - Academy of Technologies and Business	Multimedia design and communication	11.3	No	42.0
University of Copenhagen	Actuarial Sciences	11.3	Yes	39.0

Notes: The table displays the educations with the top 5 GPA cutoffs, the Math A requirement and the share of female applicants with this education as their first choice

students' survey answers confirm that the best students choose the math track. We do not have the same measure, and instead, we define competitiveness by the difficulty in being accepted at a certain field of study. By this definition, the most competitive fields are those that require the highest GPA-admission cutoffs. For these education programs, only the applicants with the highest GPAs are accepted [5]. As seen in Table 15, there is no sign of these education programs having a majority of male students. Also, only a few of these programs require high levels of math. In fact, only medical school, molecular bio-medicine, and actuarial sciences have this requirement. Of these three education programs, only a majority of males are in actuarial sciences. We take this as suggesting that math and competitiveness are not as closely linked in the Danish context as in other countries. We do not deny that competitiveness might influence one's subsequent career choices. .

The literature on risk preferences also points to gender differences, with females being more risk averse than males (Croson and Gneezy, 2009; Charness and Gneezy, 2012). Dif-

[5]Just the same as the definition of competitive fields in (Buser et al., 2014)

ferences in risk preferences could lead females to choose less risky education pathways than males. A way to measure the riskiness of an education is to look at unemployment in a given education. Being unemployed could be interpreted as a zero return on the educational investment, and the higher the chance of being unemployed after graduation, the more riskier the investment. For the applicants of 2018, we test if there is a correlation between educations with higher unemployment rates and the share of females enrolling in the those programs [6]. Using a simple ordinary least squares regression, we find that the parameter estimate on gender is practically zero and insignificant (see Appendix E for the results).

The validity rests on the assumption that students have a least a vague idea about the unemployment rates within different education programs. While this has been highly debated in the media, and there exists a website where all this information is available, we believe that students do have an idea of the average unemployment rate in these educations. Furthermore, the applicants' beliefs about the unemployment rate in the survey are in line with the actual unemployment rates. We take these results as an indicator of risk not being a leading cause of females choosing differently compared to males.

6 Conclusion

Gender differences in earnings persist in all developed countries. Recent literature has attributed the gender wage gap to persist because of a child penalty that happens only to mothers and not fathers. Furthermore, the literature on the choice of education shows that males are much more likely to choose high-paying careers in math fields. One proposed reason is that females are willing to pay more for workplace flexibility, which affects their

[6] We use the average unemployment rate in the first two years following the completion of an education program. We choose five education programs with the highest unemployment rate and five with the lowest unemployment rate. Further, we only used education programs from which more than 20 people had graduated.

sorting into different education programs.

This paper presents the first attempt to relate the gender gap in the choice of education to the gender wage gap and the expected female child penalty in the labor market. For this purpose, we developed a unique survey and invited all higher education applicants in 2018 in Denmark to participate. The survey inquires about the expectations of pecuniary and non-pecuniary factors from the top two education programs in each applicant's actual rank-ordered list of education programs. With more than 17,000 respondents, this is the biggest survey seen in the literature on education choice. Because of the scale of the survey, we are able to select a very specific group of applicants who have chosen to apply for a math field. This makes the applicants more comparable across the genders, as this group of applicants both have the abilities and are interested in math, as they chose to graduate with the highest level of math from high school and have chosen to apply for a math field. Therefore, gender differences in the choice of education among this group cannot be explained by differences in preferences for math.

We highlight three important findings below. First, our findings suggest that females and males attach nearly equal importance to pecuniary factors. Males appreciate the consumption value of studying more highly than females, who, instead, pay more attention to future non-pecuniary factors, such as work satisfaction being more important than for males. This contrasts the findings in the existing literature. Zafar (2013) uses a sample of 161 students (62 males and 92 females) from Northwestern University and find that males, in general, value pecuniary factors much more than females. A possible explanation for the different results could be the higher female participation rates in Denmark compared to the United States: Pecuniary factors associated with work are likely to matter more if you expect to work more hours.

Second, we are able to compare the expected income of the applicants to the actual income of previous cohorts of applicants within the same fields of education. We find that females are already expecting the gender wage gap when applying for higher education.

Third, we find suggestive evidence of the existence of two types of female applicants for higher education. Females who expect to have children seem to be less driven by pecuniary factors of choice compared to females who expect a lower probability of having children. Furthermore, females with high expectations of having children are also more likely to choose non-math related fields of education. Females with a lower expected probability of having children are more likely to choose math-related fields.

Taken together, these three results relate to the literature on gender differences in the choice of education and contribute two new perspectives. First, higher education applicants do expect a gender wage gap within all fields of education. Second, females seem to take future childbearing decisions into account when applying for education. We are able to focus on a sample of applicants who have an interest in and ability for math. Even among this group of applicants, females already expect to earn less than their male counterparts when applying for higher education. Furthermore, females who think that they will have children are more likely to choose a non-math field of study compared to females who report a lower probability of having children.

Our findings suggest that the child penalty shown in the existing literature might fail to account for educational decisions based on expected childbearing. Specifically, some females might opt for education programs with child-friendly labor market prospects, although they may also earn less compared to math fields. In this sense, the child penalty strikes twice.

For policy purposes, our results indicate that females might believe that prospective careers in math fields are hard to reconcile with having children. As such, policies on promoting math fields should seek to inform prospective applicants about family-friendly careers rather than just informing about a good study environment or earnings.

This paper outlines several avenues for future research. The estimates from this paper are not causal, as we only show that there is an association between the choice of education and the expectation for having children for females. We find that females with a high probability of having children are significantly different from females with a lower probability of having children. An important area for future research would be to dig deeper into this relationship. Do females anticipate a child penalty when choosing higher education? Besides, how does this relate to occupation and industry choice?

References

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1-2):343–375.

Arcidiacono, P., Hotz, V. J., and Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, 166(1):3–16.

Arcidiacono, P., Hotz, V. J., Maurel, A., and Romano, T. (2014). Recovering ex ante returns and preferences for occupations using subjective expectations data. Technical report, National Bureau of Economic Research.

Baker, R., Bettinger, E., Jacob, B., and Marinescu, I. (2018). The effect of labor market information on community college students' major choice. *Economics of Education Review*, 65:18–30.

Benjamin, D. J., Heffetz, O., Kimball, M. S., and Rees-Jones, A. (2014). Can marginal rates of substitution be inferred from happiness data? evidence from residency choices. *American Economic Review*, 104(11):3498–3528.

Black, D. A., Haviland, A. M., Sanders, S. G., and Taylor, L. J. (2008). Gender wage disparities among the highly educated. *Journal of human resources*, 43(3):630–659.

Blau, F. D. and Kahn, L. M. (2000). Gender differences in pay. *Journal of Economic perspectives*, 14(4):75–99.

Boneva, T. and Rauh, C. (2017). Socio-economic gaps in university enrollment: The role of perceived pecuniary and non-pecuniary returns. *CESifo Working Paper Series*.

Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *The Journal of Abnormal and Social Psychology*, 52(3):384.

Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, competitiveness, and career choices. *The Quarterly Journal of Economics*, 129(3):1409–1447.

Charness, G. and Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, 83(1):50–58.

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic literature*, 47(2):448–74.

Delavande, A. and Zafar, B. (2019). University choice: The role of expected earnings, non-pecuniary outcomes and financial constraints. *Journal of Political Economy*, forthcoming.

Eriksen, M. and Kjærsgård-Jensen, M. E. (2018). Pigerne er bedst til matematik i gymnasiet. Technical report, Dansk Erhverv.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., and Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, 133(4):1645–1692.

Guiso, L., Monte, F., Sapienza, P., and Zingales, L. (2008). Culture, gender, and math. *Science*, 320(5880):1164–1165.

Gupta, N. D., Oaxaca, R. L., and Smith, N. (2006). Swimming upstream, floating downstream: Comparing women's relative wage progress in the united states and denmark. *ILR Review*, 59(2):243–266.

Hastings, J., Neilson, C. A., and Zimmerman, S. D. (2015). The effects of earnings disclosure on college enrollment decisions. Technical report, National Bureau of Economic Research.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., and Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, 321(5888):494–495.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, 125(2):515–548.

Kahn, L. M. (2015). Wage compression and the gender pay gap. *IZA World of Labor*.

Kleven, H., Landais, C., and Søgaard, J. E. (2018). Children and gender inequality: Evidence from denmark. Technical report, National Bureau of Economic Research.

Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72(5):1329–1376.

Mostafa, T. (2019). Why don't more girls choose to pursue a science career? *OECD Publishing*.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? do men compete too much? *The quarterly journal of economics*, 122(3):1067–1101.

Nielsen, H. S., Simonsen, M., and Verner, M. (2004). Does the gap in family-friendly policies drive the family gap? *Scandinavian Journal of Economics*, 106(4):721–744.

Nosek, B. A. and Smyth, F. L. (2011). Implicit social cognitions predict sex differences in math engagement and achievement. *American Educational Research Journal*, 48(5):1125–1156.

OECD (2015). The abc of gender equality in education: aptitude, behaviour, confidence, pisa. OECD Publishing.

OECD (2018a). *Education at a Glance 2018*. OECD Publishing.

OECD (2018b). Labour force statistics by sex and age-indicators. `https://www.oecd-ilibrary.org/employment/data/labour-market-statistics/labour-force-statistics-by-sex-and-age-indicators_data-00310-en`. Accessed: 2019-05-30.

Reuben, E., Sapienza, P., and Zingales, L. (2014). How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences*, 111(12):4403–4408.

Wiswall, M. and Zafar, B. (2014). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, 82(2):791–824.

Wiswall, M. and Zafar, B. (2017). Preference for the workplace, investment in human capital, and gender. *The Quarterly Journal of Economics*, 133(1):457–507.

Zafar, B. (2013). College major choice and the gender gap. *Journal of Human Resources*, 48(3):545–595.

A Additional Material

A.1 Invitation letter

KØBENHAVNS UNIVERSITET
DET SAMFUNDSVIDENSKABELIGE FAKULTET

NAVN

Deltag i et spørgeskema om dit uddannelsesvalg

10. JULI 2018

ØKONOMISK INSTITUT
CEBI

ØSTER FARIGMAGSGADE 5,
BYGNING 26
1353 KBH K

TLF 3532 4411

studiesurvey@econ.ku.dk
www.econ.ku.dk/cebi

Kære NAVN
Du har netop søgt ind på en videregående uddannelse i Danmark, og derfor inviteres du til at deltage i et forskningsprojekt om uddannelsesvalg. Man ved meget lidt om, hvordan unge som dig vælger uddannelse, og derfor beder vi om din hjælp til at svare på et spørgeskema. I spørgeskemaet vil du for eksempel blive spurgt om hvor glad du forventer at blive med dit studie, og hvad du forventer af dit liv efter studiet. Alle der har søgt om videregående uddannelse i år er blevet inviteret til at svare på spørgeskemaet.

Det tager ca. 10-15 minutter at gennemføre spørgeskemaet. Når du har svaret vil du automatisk deltage i en konkurrence om et af fem supergavekort fra GoGift på 1000 kr. Vinderne vil få direkte besked efter svarfristen.

Du kan deltage til og med den 27. juli 2018. Invitationen er personlig, og vi beder derfor om, at du ikke videregiver den til andre. Undersøgelsen starter, når du klikker på dette link:

Link

Dine svar i undersøgelsen vil blive behandlet anonymt. Københavns Universitet har godkendt forskningsprojektet, og vores procedurer opfylder Persondataforordningen og Databeskyttelseslovens krav om databehandling. Vi har dine kontaktoplysninger fra Styrelsen for Forskning og Uddannelse. På næste side kan du læse mere om din sikkerhed og dine rettigheder.

Du er velkommet til at kontakte os, hvis du har spørgsmål til undersøgelsen. Du kan ringe til projektkoordinatorer Helene Willadsen og Anne Toft Hansen på telefonnummer 3532 4411 mandag-fredag kl 10-12 eller skrive til adressen studiesurvey@econ.ku.dk.

Med venlig hilsen
Helene Willadsen og Anne Toft Hansen

Persondatabeskyttelse

SIDE 2 AF 2

Københavns Universitet (KU) har fået udleveret data fra Styrelsen for Forskning og Uddannelse, jf. databeskyttelsesforordningens artikel 6 og forvaltningslovens § 31.

- KU overholder Persondataforordningen pr. 25. maj 2018.
- Oplysningerne bruges alene i videnskabeligt øjemed.
- Der formidles aldrig resultater således, at enkeltpersoner kan identificeres. Alle resultater videreformidles i aggregeret form.
- Efter undersøgelsens afslutning vil det følgende ske:
  - Population og spørgeskemadata overføres til Danmarks Statistik, og anonymiseres af Danmarks Statistik. Alt andet data slettes, således at der ikke findes lokale kopier.
  - Oplysningerne gives ikke til tredjemand.

Dine rettigheder som deltager i undersøgelsen

1. Deltagelse i spørgeskemaet er frivilligt. Hvis du svarer ja til at svare på spørgeskemaet, giver du samtykke til, at vi kan bruge de oplysninger du giver os i forskningsprojektet. Du har til enhver tid ret til at tilbagekalde samtykket. KU kan ikke fortsætte med at behandle oplysningerne efter dit samtykke er trukket tilbage.
2. Som deltager har du ret til at få slettet oplysninger fra spørgeskemaet, hvis oplysningerne ikke længere er nødvendige til det formål de blev indsamlet til. Oplysningerne vil også blive slettet, hvis du tilbagekalder samtykket til behandlingen eller hvis oplysningerne ved en fejl er blevet behandlet ulovligt. Du har ikke krav på sletning af oplysninger arkiverede efter arkivlovens regler i universitetets arkivsystem. Oplysninger som indgår i et forskningsprojekt, skal ikke slettes, hvis det sandsynligvis vil betyde, at det vil umuliggøre eller i alvorlig grad hindre udførslen af forskningsprojektet.
3. Hvis du mener, at der er registreret forkerte oplysninger, kan du bede universitetet om at berigtige oplysningerne. Det vil sige, at universitetet retter oplysningerne eller noterer, at oplysningerne er forkerte og registrerer de rigtige oplysninger. Du kan læse privatlivspolitikken her: https://informationssikkerhed.ku.dk/persondatabeskyttelse/privatlivspolitik/
4. Data vil senest blive slettet 31/12-2028.
5. Du kan kontakte følgende databeskyttelsesrådgivere:
   - Københavns Universitets databeskyttelsesrådgiver:
     - Email: databeskyttelsesraadgiver@adm.ku.dk.
     - Telefon: +45 3533 5688
   - Styrelsen for Forskning og Uddannelse databeskyttelsesrådgiver:
     - E-mail: dpo@ufm.dk
     - Telefon: +45 7231 8909
     - Brev: Brev til ministeriet "att. Databeskyttelsesrådgiver".

Dataansvarlig for forskningsprojektet

Helene Willadsen, Københavns Universitet, Økonomisk Institut.

B Monte Carlo Simulations

- N applicants
- with 5 possible education choices
- there are two choice specific explanatory variables: $x_1 \sim \mathcal{N}(0, 16)$ and $x_2 \sim \mathcal{N}(0, 4)$
- the utility derived from each of the 5 educations is a function of the two explanatory variables:

  (i) $u = 0.5 + 1.5x_1 + 0.5x_2 + \varepsilon$ for choice=1

  (ii) $u = 2 + 1.5x_1 + 0.5x_2 + \varepsilon$ for choice=2

  (iii) $u = 1 + 1.5x_1 + 0.5x_2 + \varepsilon$ for choice=3

  (iv) $u = 1.5 + 1.5x_1 + 0.5x_2 + \varepsilon$ for choice=4

  (v) $u = 1.5x_1 + 0.5x_2 + \varepsilon$ for choice=5 , where ε is T1EV distributed

- We estimate the choice model for two different scenarios:
  - Spec 1: we have elicited preferences for the complete rank of all five education choices for each applicant
  - Spec 2: we only have elicited preferences for the two most preferred of the five education choices for each applicant
- We run 5,000 replications with $N = 1000$.
- We then check how well the choice model recovers the true parameters in each of the two scenarios

Figure 4: A simulated example

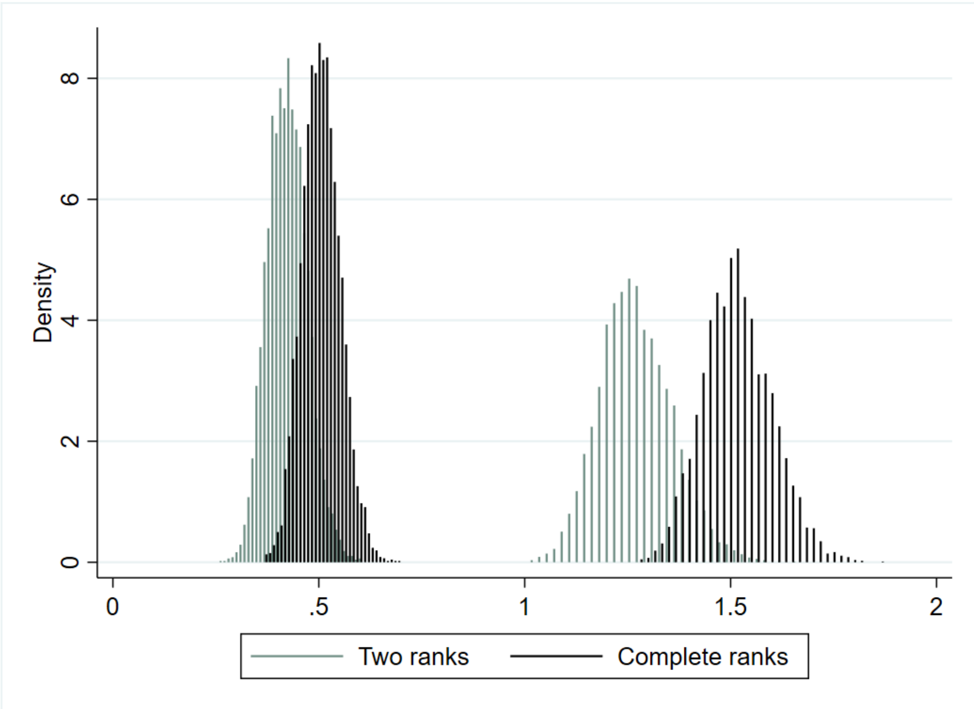


Table 16: Summary statistics from simulation

	Two choices		Complete choice set	
VARIABLES	mean	sd	mean	sd
β_1	1.269	0.0885	1.517	0.0841
β_2	0.423	0.0481	0.506	0.0461
Constant choice 1	-0.841	0.226	-1.008	0.218
Constant choice 2	0.428	0.206	0.509	0.201
Constant choice 3	-0.422	0.216	-0.506	0.210
Constant choice 4	-1.264	0.242	-1.514	0.231
Observations	5,000		5,000	

Notes: The omitted category is choice 5

C Summary statistics

Table 17: Summary statistics of all applicants for higher education

	Respondents	Non-respondents	All
Age	24.01	23.96	23.97
	(6.58)	(5.72)	(5.93)
Programs applied for	2.59	2.60	2.60
	(1.87)	(1.88)	(1.88)
Female	0.65	0.56	0.58
	(0.48)	(0.50)	(0.49)
University of Copenhagen as first choice	0.180	0.11	0.13
	(0.00)	(0.00)	(0.00)
Long higher education as first choice	0.56	0.49	0.50
	(0.50)	(0.50)	(0.50)
Medium higher education as first choice	0.33	0.37	0.36
	(0.47)	(0.48)	(0.48)
Short higher education as first choice	0.11	0.14	0.14
	(0.31)	(0.35)	(0.34)
Observations	17,964	58,405	76,369

D Calculating choice probabilities

We calculate the predicted probability of choosing education j as first choice and education k as second choice as:

$$Pr(d_1 = j, d_2 = k) = \phi(X\hat{\beta})Pr([d_1, d_2] \in [j, k]) \tag{7}$$

where X is a vector of means for each of the factors from our choice model presented in section 4.1. The means are calculated for each combination of education choices, for example, the mean for pre-tax income is calculated separately for applicants with the combination (1) tech math and (2) health math and applicants with the combination (1) tech math and (2) nature math. We calculate the mean for both each of the first choice education (e.g tech math) and each of the second choice educations (e.g health math and nature math). The $\hat{\beta}$ is a vector of the parameter estimates from our choice model presented in Table 7. The last term on the right hand side is the probability of first choice d_1 is education j and second choice is education k. This probability is given directly by the share of the sample that has this specific combination of first and second choice education [7]. We calculate the predicted probability of choosing education j as first choice as:

$$Pr(d_1 = j) = \sum_{\substack{k=1 \\ k \neq j}}^{5} Pr(d_1 = j, d_2 = k) \ , \ \forall \ j = 1, .., 5 \tag{8}$$

Hereby, we obtain a predicted probability of choosing each of the five different educations as first choice with all possible combinations of second choice fields. The predicted gender gap in each education is calculated as the difference in the predicted probability of

[7]Some of these combinations of fields are very small. When bootstrapping the standard errors some of these shares are set to zero, and therefore do the estimated gender gaps not always sum to zero

choosing education j between males and females:

$$Pr_{male}(d_1 = j) - Pr_{female}(d_1 = j) \tag{9}$$

With this method, we estimate changes in the gender gap by setting females' expectations for future income and probability of children equal to males' expectations. However, the gender gap forecast will give a conservative estimate of changes in the gender gap because the predicted gender gaps only consider the top two ranked choices of each applicant as possible choices. This means, that the change cannot exceed the number of females that have an education in their choice set. The estimates should therefore be interpreted as fairly restrictive, because it simply cannot take into account that some females might want to include a "new" education in their choice set when their expectations are set to equal the male expectations.

E Regression on unemployment

The educations with the highest unemployment rate:

1. European Ethnology, unemployment rate: 46.13 percent
2. Digital design, unemployment rate: 46.0 percent
3. Sociology and culture, unemployment rate: 35.71 percent
4. Engineering - export and technology, unemployment rate: 34.19 percent
5. Literature (history), unemployment rate: 30.38 percent

The educations with the lowest unemployment rate:

1. Optometry, unemployment rate: 0.69 percent
2. Medicine, unemployment rate: 0.83 percent
3. Engineering - software, unemployment rate: 1.03 percent
4. Math and economics, unemployment rate: 1.12 percent
5. Sign language and interpretation, unemployment rate: 1.17 percent

Data Source:

`https://ufm.dk/uddannelse/statistik-og-analyser/faerdiguddannede/aktuel-ledighed`

Table 18: Correlation between female and unemployment rate

	Unemployment
Female	-0.0044
	(0.0033)
Observations	4603

Notes: Outcome is unemployment rate
*** p<0.01, ** p<0.05, * p<0.1

Chapter 2

Should I Stay or Should I Go? The Effect of University Location on Danish Graduates' Residential Choice and Place of Work

Should I Stay or Should I Go? The Effect of University Location on Danish Graduates' Residential Choice and Place of Work *

Anne Toft Hansen †

Abstract

This paper examines the effect of university location on the work and residence choices of Danish university graduates. The admission process into higher education in Denmark creates credible instruments from discontinuities that randomize applicants near an unpredictable GPA-based admission cutoff into different locations. Using a regression discontinuity (RD) design that takes into account the choice of university location as an unordered choice, this paper estimates the causal effect of studying either in a metropolitan or a non-metropolitan region on where graduates live and work eight years after applying for university. The findings suggest that the location of study does have an effect on where graduates choose to live and work, even for graduates who do not get accepted for their first choice university location but instead graduate from a university in their second choice location.
colleagues

*I thank Bo Honore, Mette Ejrnæs, Mette Gørtz, participants at the 2016 CEN workshop at Copenhagen Business School, at the 2016 VIVE conference, and seminar participants at Kraks Fond - Institute for Urban Economic Research for comments and suggestions. Finally, I thank the Ministry of Education for providing detailed information about the application system and the data structure.

†The Danish Centre for Social Science Research & the Department of Economics, University of Copenhagen. E-mail: ath@vive.dk.

1 Introduction

Human capital and employment mobility are positively correlated (Herzog Jr et al., 1985). This implies that university graduates are a very mobile part of the labor force. The group is furthermore a group of workers that most regions wish to attract. Human capital spillovers to local regions can increase aggregate productivity more than just in terms of the direct effect on productivity (Ahlin et al., 2014). Also, increases in the educational level of the local population can increase social returns, such as reduced crime and increased political participation (Moretti, 2004). Consequently, to attract and retain university graduates is a key goal for policy makers. The general trend is that regions that manage to attract and retain the highly educated perform better, while others fall behind. This knowledge constitutes the importance of identifying factors that attract and retain university graduates.

In recent years in Denmark and several other Western countries, the political agenda has been to try to contain urbanization and regional brain drains. In 2015, the Danish Government decided to move more than 10 percent of all public workplaces from the capital to various smaller regions of Denmark. [1] The motivation for doing so was to create a "better balance" to ensure economic growth and development in all parts of Denmark. This forced many workers living in the Copenhagen area to either relocate or quit their jobs. Another policy tool to increase the proportion of highly skilled labor in non-metropolitan regions could be to push more prospective university students out of the metropolitan universities and into the non-metropolitan ones. However, this relies on the universities being able to retain their students as they graduate, and there exists minimal evidence of whether this has been the case.

This paper investigates the effect of university location on graduates' location choice. More specifically, I estimate the effect of graduation from a geographic university location

[1] The workplaces mainly consist of public administration agencies, such as the Danish tax agency.

on the probability of living or working in the specific university location or in another region outside the metropolitan regions eight years after having applied for university. I define a metropolitan region as the two largest cities in Denmark, Aarhus and Copenhagen, which also have the two largest universities; non-metropolitan regions include all the other regions of Denmark. [2]

A major identification issue when trying to estimate the effects of the geographic location of education is that students' select themselves into field and geographic location of study making the location of study endogenous. A mere comparison of university graduates from different locations will yield misleading conclusions on graduates' residential choice. To eliminate the selection bias, I need a sample of university students who are randomly assigned to different university locations. Fortunately, the admission process into higher education in Denmark creates credible instruments from discontinuities that randomize applicants near an unpredictable GPA admission cutoff into different locations. This allow me to estimate the effect of having (randomly) studied in a specific university region on the probability of working and settling down in the same region or in a non-metropolitan region after graduating from university using a fuzzy regression discontinuity (RD) design.

My findings suggest that universities can retain some of their graduates. I find that both metropolitan and non-metropolitan universities retain their graduates. Even some of the graduates who had preferred to study in a metropolitan region but did not get accepted for their preferred location and graduated from a university in their second choice non-metropolitan region end up living and working in the university region eight years after the year of application (YOA). The findings show that there is a 39 percentage point higher probability of living and 61 percentage points of working in the same region as the second choice non-metropolitan university if a student graduated from there compared

[2] A region in Danish is called a "kommune" or a municipality. There are 98 regions in Denmark.

to graduating from their first choice university location (i.e., in a metropolitan region for more than 80 percent of the sample). Furthermore, there is an increased probability of living and working outside a metropolitan region of 46 and 48 percentage points for graduates from their second choice non-metropolitan university region compared to if they had graduated from their first choice university location. Subgroup analysis shows that this effect is driven by students who lived outside the metropolitan regions in the YOA for university. No effects exist for students who lived in a metropolitan region when they applied for university.

From a policy perspective, the results point to that universities can be used as a policy tool to attract human capital to non-metropolitan regions. This paper provides evidence that pushing university applicants to study in a non-metropolitan regions will result in some of these applicants staying and entering the labor market in these regions after graduation. Therefore, it could be a useful policy tool to increase the share of highly educated graduates in local non-metropolitan regions.

The remainder of this paper is organized as follows. Section 2 provides an overview of the existing literature. Section 3 describes the higher education system in Denmark. Section 4 presents the data. Section 5 elaborates the empirical strategy, and section 6 presents the results. Section 7 concludes this paper.

2 Existing literature

It is well-established in the literature that individuals living close to a college or university are more likely to enroll (Card, 1993; Currie and Moretti, 2003; Spiess and Wrohlich, 2010; Öckert, 2012). Universities can attract potential human capital, which may increase the educational level of the local population, and if students choose to stay after graduating, this could have an effect on the local labor market.

The decisions of where to settle down and enter the labor market of university graduates has been an area of research that has attracted more attention in the last couple of decades. The literature reports mixed results on universities being able to retain their graduates across countries; however, there is agreement that universities in metropolitan regions can retain significantly larger proportions of their graduates compared to universities in less urbanized areas (Krabel and Flöther, 2014; Faggian and McCann, 2009; Venhorst et al., 2011; Krabel and Flöther, 2014; Kotavaara et al., 2018; Winters, 2011).

A couple of papers investigate the migration patterns of university graduates in Finland (Haapanen and Tervo, 2012; Kotavaara et al., 2018). They find that most Finish graduates do not move from their region of study within 10 years of graduation. Haapanen and Tervo (2012) find that of the graduates who study in their home region (where they lived before applying for university) in the capital of Finland, Helsinki, 87 percent are still living in the region 10 years after graduation. Of those studying away from their home region in Helsinki 64 percent live in the region 10 years after graduation. The numbers are much lower when looking at the graduates who studied in the other regions of Finland. They estimate a survival rate of 61 and 37 percent for the group studying at home and the group studying away, respectively. Migration most often inclines toward the metropolitan regions while universities in non-metropolitan regions struggle to retain their graduates.

Winters (2011) examines on American data, why smart cities are growing. He defines smart cities as cities that have a university, and investigates the in- and out-migration of human capital of these cities. As expected, there is a large in-migration to smart cities due to individuals pursuing higher education. Even though there is a large out-migration of these individuals once they graduate, he finds that the net gain is positive, meaning that the university regions attract more high skilled individuals than they lose.

Faggian and McCann (2009) find that graduates from the UK are very mobile. Seventy percent of all prospective students relocate over large distances to study, and when including graduates who move after attending university in their home region, 80 percent of all graduates have moved away from their home region by the time they enter the labor market.

Krabel and Flöther (2014) wish to understand the determinants of graduates' regional mobility when entering the labor market by estimating a two-stage selection model by first analyzing the factors determining the likelihood of finding a job and in the second step by considering the determinants of regional mobility when finding a job. They do this on a sample of German university graduates from 2006 and 2007. They find that two out of five graduates stay in the region where they studied, and three out of five stay in the larger geographic areas. When investigating the factors affecting the probability of employment and hereby the probability of moving, they find that social ties to employers impact the chance of finding a job and reduce the likelihood of leaving the university region. Furthermore, compared to metropolitan areas, the graduates are more likely to leave non-metropolitan areas.

In line with the literature presented above, Venhorst et al. (2011) study the relationship between migration of Dutch college and university graduates and both regional economic circumstances and individual characteristics. Their findings show that regional economic growth and employment rates are important factors for retaining university graduates. There are large inflows of graduates to the economic center of the Netherlands; however Venhorst et al. (2011) conclude that graduates over time are migrating less, which can be explained by the economic developments in various regions.

It is important to note, that none of the studies above has probed the impact of the choice of where to study on the choice of where to live and work after graduation. To asses where university graduates choose to live and work, it is important to consider that the choice of where to live and work is likely affected by the initial decision of where to attend university, which this paper aims to explore.

This paper estimates the effect of the choice of where to study on the choice of where to live and work. I use the admission system into higher education in Denmark as exogenous variation of university location choice to estimate the causal effect of university location on the choice of residence and place of work after graduation. In recent years, several papers have used admission systems into higher education as a source of exogenous variation to estimate the effect of education on various outcomes. Öckert (2010) uses the college admission system in Sweden to estimate the effect of higher education on earnings. He finds for a cohort of higher education applicants in 1982 that the effect of getting accepted is 0.2 years of college, albeit without any significant returns to education. Similarly, Heinesen (2018) estimates returns to education in Denmark for higher education applicants between 1994 and 2002. His results are consistent with findings of Öckert (2010) in that getting accepted for an education increases the probability of completing any higher education by 6-9 percentage points and increases the total years of education by 0.2 percent. He also does not find any significant effects on earnings. Another Danish study using admission data is by Humlum et al. (2017) who estimate the effect of higher education on fertility decisions. Their findings suggest that being above the admission cutoff speeds up education completion and labor market entry, which thereby speeds up the process of starting a family; being above the admission cutoff increases the number of children by 40 percent eight years after the YOA.

Lastly, Kirkeboen et al. (2016) use the Norwegian admission system to estimate the effects of field of study and institution on future earnings. They look at the intensive margin of choice and model the decision of which field of study and which institution to attend. They estimate the returns of specific fields and institutions in an unordered choice model using that some educations have binding admission cutoffs and some do not. This means that being just above an admission cutoff for a given education makes you more likely to complete that specific education, and if you are just below the cutoff, it will make you more likely to complete your second choice of education. Using this strategy, Kirkeboen et al. (2016) estimate the payoff of each first choice field (or institution) compared to the second choice field. Their findings suggest that there are large returns to the field of education dominating the returns to institution.

3 Institutional setting

In Denmark, education, including higher education, is free. Students enrolled in higher education receive a monthly grant of USD1000, and students can also apply for governmental student loans in combination with the student grant. [3] High school graduates apply for college and university education through a two-tiered centralized clearinghouse, where an applicant can apply for up to eight different programs in the same application. The vast majority of applicants (about 92 percent) apply through the first tier, where applications are based solely on high school GPA. Through a simple procedure, applicants create a rank-ordered list of education choices by logging in the website of the clearinghouse and choosing field of study, institution, and location. This paper focuses on the applicants who apply through the first tier only. The remaining 8 percent of applicants apply based on their high school GPA, other merits (e.g., work experience and travel experience), and a written cover letter targeted at the specific education choice.

After the admission deadline, the clearinghouse ranks the GPA of all applicants for a given program, and assigns a slot for the applicants with the highest GPAs until there are no more slots. The GPA of the last applicant is the GPA admission cutoff within a given year. Because of the excess demand, some applicants will not get into their preferred choice of program. In such cases, they will then be assessed on the admission cutoff for their second choice, and the system continues until it locates an education program on the

[3] Denmark is one of the countries that has the highest public spending on education in the world; it allocated 6.3 percent of its GDP to it in 2014 OECD (2018)

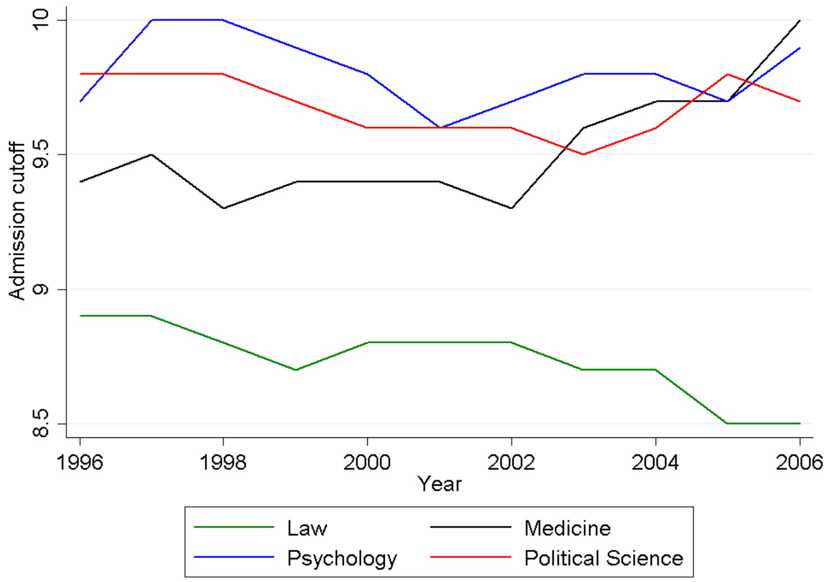


Figure 1: Admission cutoffs over time for selected educations at University of Copenhagen

applicant's rank-ordered list where the applicant's high school GPA is higher than the admission cutoff. The GPA cutoffs are formed every year and are, therefore, unknown to the applicants at the time they submit their application. Because most educations only have yearly enrollment, applicants are encouraged to state more than one education choice on their rank-ordered lists, if they wish to increase their chances of being admitted to an education program. The admission criteria of the GPA cutoffs is solely driven by supply and demand, and not by a specific ability criteria for specific education programs.

Figure 1 shows variations in admission cutoffs over the period 1996 to 2006 for four of the large fields of education at University of Copenhagen, namely law, psychology, medicine, and political science. The admission cutoffs have varied the least (0.3-grade points) for psychology and political science, and most (0.7-grade points) for medicine over 10 years.

Table 1: Sample Restrictions

	No.	Percent
All applicants for higher education in 1996-2007	211,596	100
Applicants who are accepted for an education	183,456	86.70
Applicants who apply for more than one education program	81,066	38.31
Applicants who apply for university education	28,066	13.26
Applicants who apply within same field of studies	19,538	9.23
Applicants who did not apply for the same university	7,871	3.72
First choice has a binding cutoff	5,290	2.50
Information about GPA and geographic placement	4,337	2.05
Restricted sample: +/- 0.5 around admission cutoff	2,520	1.19

4 Data

I apply data of all applicants for higher education in Denmark for the years 1996-2007 and merge them with Danish administrative registers to obtain a very detailed set of individual characteristics. The background information includes gender, age, ethnicity, parental education level, and high school GPA of all the applicants. Additionally, I include variables concerning where applicants live in the YOA.

I create the outcome variables using information about each individual's residence and work region over time merged with information about the region of university that each individual graduates from. Furthermore, I obtain information about the applicants' unemployment rates and annual pre-tax earnings eight years after YOA.

4.1 Data restrictions

To create a sample of applicants for university education conducive to the fuzzy RD design with exogenous variation around the admission cutoff stemming only from the choice of location, I have to restrict the sample of applicants for higher education. Table 1 shows the sample restrictions. I consider all applicants for higher education, who apply for the first time between 1996 and 2007. I exclude the applicants who applied for one education program only, and applicants who did not receive an offer letter. Furthermore, I only consider university applicants and not the college applicants mainly because there are many colleges across smaller cities in Denmark, allowing many students not to relocate when attending them.

I group the specific fields of study into five broader categories: (1) humanities, (2) social science, (3) health, (4) business, and (5) nature and technology. The labor markets across these field categories are very different in terms of geographic location, whereas a medical doctor can work from every city in Denmark, which might not be the case for a graduate in business. To prevent the effect from becoming contaminated by possible labor market differences across the fields, I only consider graduates who applied for education programs within the same field category.

I also only consider the applicants who apply for two different geographic university locations where the first choice of location needs to have an admission cutoff, and the second choice must have a lower cutoff (or no cutoff). Thereby, the variation only stems from the applicants' ranking of geographic university locations. Lastly, the applicants with missing information about their high school GPA and/or geographic placement are excluded. This leaves me with a sample of 2 percent of the total pool of applicants. In my main estimation specification, I restrict the sample further to only include applicants that have a high school GPA that is +/- 0.5-grade points around the admission cutoff for their first choice of university location. The analysis sample, therefore, consists of 2,520 university applicants.

The sample consists of a certain type of university applicants. The applicants have applied across university locations and are probably among the most mobile of the pool of university applicants. A lot of applicants do consider one location only and do, instead, apply for different fields of education. Pushing this group of students to study in another

Table 2: Descriptive Statistics

	Sample		All	
	mean	sd.	mean	sd.
Male	0.36	(0.48)	0.40	(0.49)
Foreign origin	0.05	(0.22)	0.11	(0.31)
Mother has higher education	0.53	(0.50)	0.38	(0.48)
Father has higher education	0.46	(0.50)	0.33	(0.47)
Age	21.16	(2.77)	22.32	(5.22)
Total number of rankings applied for	2.09	(0.38)	1.74	(1.11)
High school GPA (std.)	0.00	(1.00)	-0.86	(1.40)
Rank of final offer	1.25	(0.49)	1.16	(0.54)
Live in university region in YOA	0.30	(0.46)	0.31	(0.46)
Live in metropolitan region in YOA	0.19	(0.39)	0.19	(0.39)
Observations	4,337		211,538	

location will not necessarily have the same effect as for the sample selected for this study. Another group of applicants who are not considered in this paper is the group that apply across both locations and fields. This group might be as mobile as the applicants in the sample, and my findings can possibly be extended to this group. However, for the purpose of this paper, I do not wish to consider these applicants as it would be very difficult to distinguish the pure location effect from the effect that a specific field of study has on the choice of location (some fields have a few large concentrated labor markets, while labor markets for other fields are very geographically dispersed).

From Table 2, it is evident that my sample is not representative of the entire pool of applicants. This is expected since the sample is restricted to university students. The total pool of applicants also comprises college applicants. The applicants in the sample have higher educated parents and perform better in high school. There are slightly fewer males, and only about 5 percent of the sample are of foreign origin, whereas 11 percent of all the applicants are foreigners.

Table 3 presents the share of applicants within each of the five field of studies across

Table 3: Applicants' choice of field of education across university location

	Copenhagen	Aarhus	Aalborg	Odense
Humanities	29.6	30.6	59.3	28.3
Social Science	35.1	35.0	37.7	20.8
Business	10.1	6.4	-	-
Health	14.6	21.8	-	36.9
Nature and Technology	10.5	6.2	3.1	14.0
Observations	2,119	1,424	162	280

Notes: The table reports column percentages

the largest four university regions. In Copenhagen[4]and Aarhus, which is the two largest university regions, it is possible to study all five fields. The largest share of applications comes from humanities and social sciences. The three fields, humanities, social science, and nature and technology can be studied in all of the four university locations.

Table 4 shows applicants' location choices. The table shows all first and second choice combinations of university locations. The two largest university regions, Aarhus and Copenhagen, are close substitutes. Almost half of the applicants who choose Copenhagen as their preferred university location choose Aarhus as their second choice. Likewise, almost 55 percent of the applicants who prefer to study in Aarhus have Copenhagen as their second choice. Roskilde and Odense are also commonly a second choice location for the applicants with Copenhagen as first choice. They are the two university regions that are geographically closest to the Copenhagen region. About 20 percent have Odense as second choice and 25 percent have Roskilde as their second choice. The same pattern appears for the other first choice university locations. Aarhus and Copenhagen are often second choice locations; however large shares of the applicants choose as second choice the university regions that are closest to their first choice university region.

In the fuzzy RD models, I aggregate the individual university regions into metropoli-

[4]The Danish Technical University is included in the Copenhagen region, even though it is technically in the neighboring region of Lyngby.

Table 4: Applicants' first and second choice university location

	Second choice location						
	Copenhagen	Aarhus	Aalborg	Odense	Roskilde	Other	Total
First choice location							
Copenhagen	0.0	47.2	5.1	20.2	24.9	2.6	100.0
Aarhus	54.7	0.0	15.9	27.7	0.7	1.0	100.0
Aalborg	30.9	53.7	0.0	10.5	4.9	0.0	100.0
Odense	36.7	54.3	5.4	0.0	3.6	0.0	100.0
Roskilde	83.8	7.4	4.3	4.5	0.0	0.0	100.0
Total	28.3	29.2	8.4	19.7	12.8	1.6	100.0

Notes: The table reports row percentages

tan and non-metropolitan regions. I define the metropolitan university regions to include the Copenhagen region [5] and Aarhus region. Non-metropolitan regions are all other regions in Denmark. I aggregate the data in this way because the previous literature reports differences in the moving patterns of graduates from metropolitan and non-metropolitan regions, as presented in section 2. Ideally, I would estimate the fuzzy RD model for each university region separately; however the sample size is too small for such an analysis.

To get a better idea of how the applicants rank the two metropolitan and non-metropolitan university regions, Table 5 presents the choice of university location aggregated on metropolitan vs. non-metropolitan regions. Almost 82 percent of the sample prefer to study in a metropolitan region. Forty-one percent of the total sample apply for a metropolitan region both as their first choice and their second choice location (this will by construction be combinations of applying for Aarhus or Copenhagen as their first and second choice). Another 41 percent of the sample apply for a metropolitan university as their first choice and a non-metropolitan as their second choice.

Eighteen percent prefer to study in a non-metropolitan region. Less than 2 percent prefer to study in a non-metropolitan region and has a metropolitan region as second

[5] Including Frederiksberg region, which is a small region that is located in the middle of the Copenhagen region

Table 5: Applicants' first and second choice university location. Aggregated to metropolitan and non-metropolitan regions

	Second choice location					
	Metropolitan region		Non-metropolitan region		Total	
	no.	percent	no.	percent	no.	percent
First choice location						
Metropolitan region	1,763	40.65	1,780	41.04	3,543	81.69
Non-metropolitan region	83	1.91	711	16.39	794	18.31
Total	2,491	57.44	1,846	42.56	4,337	100

Notes: The table shows the cell percentage for metropolitan and non-metropolitan first and second choice, which shows the shares of the sample with each possible combination of choice

choice. Sixteen percent has a non-metropolitan region both as their first and second choice location.

An important insight from Table 5 is that since almost all the applicants have a metropolitan first choice location when evaluating the completion of university in the second choice location, it will almost exclusively be compared to graduates who completed their university education in a preferred metropolitan university location.

Figures 2 and 3 show how many students actually live in the same region where they study. Figure 2 shows the share of applicants who get accepted for and live in their first choice university location divided into metropolitan and non-metropolitan regions from the YOA and eight years on. Figure 3 shows the share of applicants who get accepted for and live in their second choice university location. The vertical line on the x-axis indicates year one, where the students begin their studies. As for the first choice university locations, 9 percent of the applicants who end up graduating from a non-metropolitan location already lives there in the YOA. For the metropolitan universities, 14 percent of the students already live there in the YOA. The numbers increase significantly from the YOA to the first year after applying, where the students are now enrolled in university. Forty-three percent of the non-metropolitan students now live in the same region as the

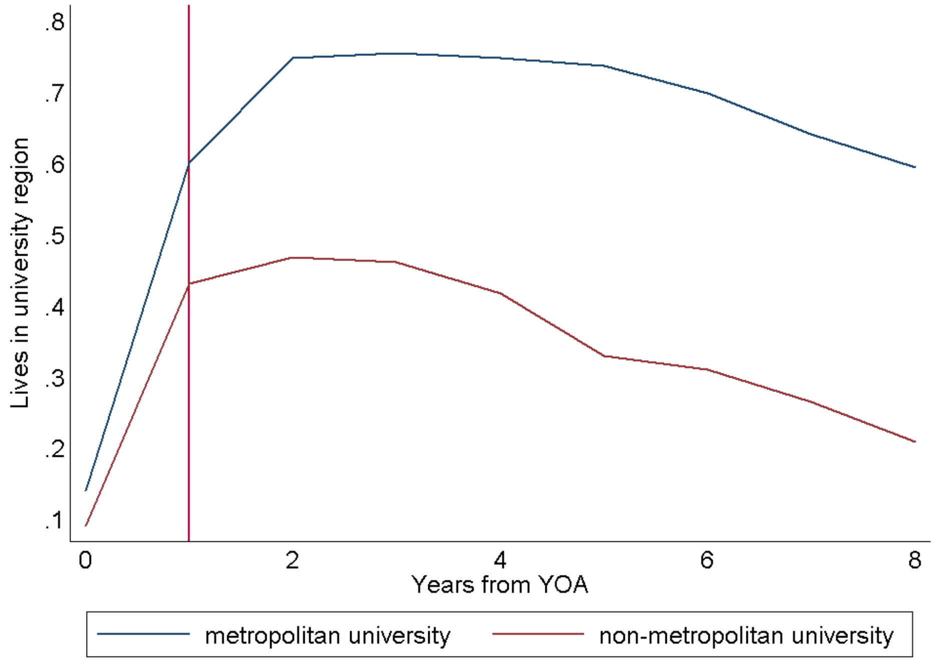


Figure 2: Share of students who lives in the university region over time. First choice university location

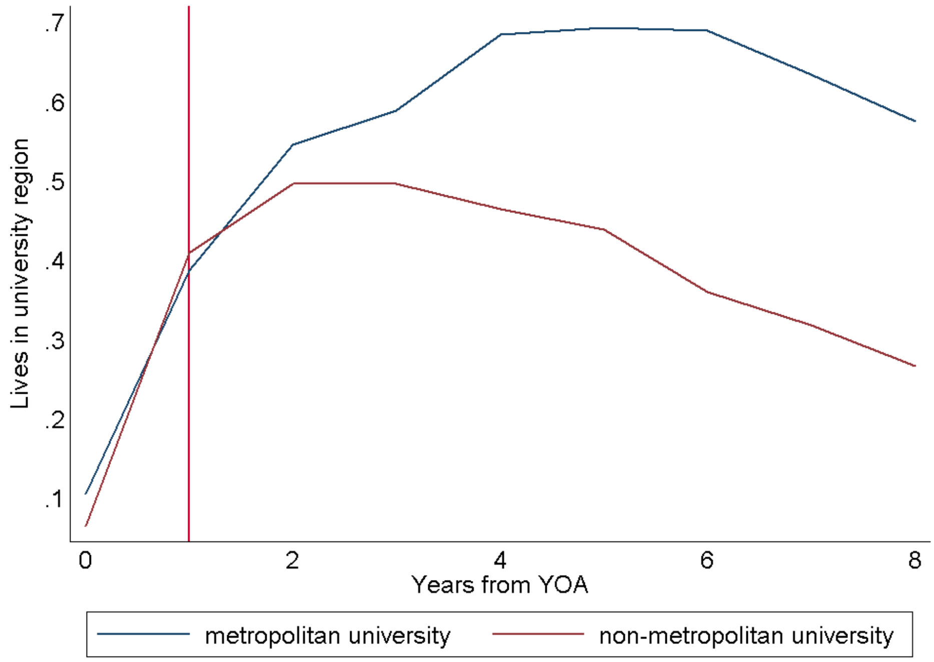


Figure 3: Share of students who lives in the university region over time. Second choice university location

university, and 60 percent of the students who study in a metropolitan university now live in the same region. It should be noted that the remaining share of applicants could have moved to a region close to the university region. Especially for the Copenhagen region where the cost of living is high, some students choose to live in a region just outside Copenhagen. By year two after YOA, 75 percent of the students live in the region of the metropolitan universities and 47 percent in region of the non-metropolitan universities. The numbers remain fairly constant up until year five, whereas they start to decline. Eight years after applying, 59 percent still resides in the metropolitan university region, and 21 percent still resides in the non-metropolitan region.

For the second choice locations a similar pattern arises. Slightly fewer students live in the region in the YOA (year zero) compared to the applicants studying in their first choice locations. Ten percent already live in the metropolitan second choice location, and 6 percent in the non-metropolitan regions. Similar to the first choice university regions, the share of applicants increase significantly over the first years, and start declining from around year five. Eight years after YOA, 57 percent still lives in the metropolitan university regions and 27 percent in the non-metropolitan university regions.

For both first and second choice university locations and in both metropolitan and non-metropolitan regions, more students live in the university region eight years after YOA compared to the year the students apply for education (year zero). Nevertheless, the shares are much larger for metropolitan university locations for both first and second choice. Taken together, this suggests that universities, especially the metropolitan ones, attract and retain some university applicants.

5 Empirical strategy

I wish to estimate the effect of university location on the choice of where to live and work after graduation. One major challenge is that the choice of where to study is endogenous. In other words, for many higher education applicants, the decision on where to live and work is likely to influence where they intend to study. By just estimating the effect of university location on the future choice of where to live and work, the estimated effect will not take into account that students select themselves into locations where they wish to study. An ideal experiment would be to randomly assign a university location to higher education applicants and force the applicants to complete their education in that given location. However, this is of course not possible. Instead, I will use the Danish admission system for higher education. The admission system creates credible instruments from discontinuities for specific education programs (and hereby locations), which randomizes applicants near an unpredictable GPA admission cutoff into different locations. This allows me to estimate the effect of a specific university location on the choice of where to live and work eight years after applying for university.

It is likely that the effect of the university location is not the same for the applicants' first and second choice of location. Applicants who graduate from their preferred university location could be more likely to stay in the university region compared to applicants who graduate from their second choice university location (simply by the fact that the applicants prefer their first choice location to their second choice location). Therefore, I estimate the effect separately for graduates from first choice locations and graduates from second choice locations. Furthermore, as the previous literature has established, the universities in metropolitan regions have higher retention rates than the universities in less metropolitan regions. Therefore, I also estimate the models separately for graduates from the metropolitan and non-metropolitan regions.

Taken together, I estimate the model separately for four different subgroups of graduates: (1) graduates from first choice metropolitan university locations, (2) graduates from first choice non-metropolitan university locations, (3) graduates from second choice metropolitan university locations, and (4) graduates from second choice non-metropolitan university locations. This allows me to understand if the effects are different for metropolitan universities compared to non-metropolitan ones, and furthermore, if the effects are different across first and second choice locations.

The choice of university location is not an ordered choice like years of education, for example. Years of education can easily be ordered; six years of education over five years of education over four years of education and so forth. However, it is not straightforward to do the same with university locations. As a consequence, the choice of university location cannot be estimated using a standard fuzzy RD design.

When estimating a model using fuzzy RD with an unordered choice, it is important to take the graduates' margin of choice into account (Kirkeboen et al., 2016). This means that for graduates from first choice locations, I need to take the second choice location into account. And for graduates from second choice locations, I need to take the first choice location into account. If I were to follow the methods developed and applied by Kirkeboen et al. (2016) directly, I would need to estimate a matrix of coefficients for graduates with each combinations of first and second choice location. However due to data limitations, I instead control for the graduates' margin of choice directly in the models. Because I divide the data into four groups, as explained above, I include a control variable in the regressions for whether the applicant has applied to a metropolitan or non-metropolitan second choice university location for applicants who have applied for a first choice university location. Similarly, in the regressions for the applicants who graduate from their second choice location, I include a control variable in the models for whether their first choice location is metropolitan or non-metropolitan. [6]

The fuzzy RD design is estimated as a two-stage least-squares (2SLS) model, as follows:

$$y_{ij} = \beta c_{ij} + \gamma x_i' + f(r_{ij}) + \eta a_{ik} + \lambda_j + \delta_t + \varepsilon_{ij} \tag{1}$$

$$c_{ij} = \pi d_{ij} + \alpha x_i' + f(r_{ij}) + \rho a_{ik} + \tau_j + \phi_t + u_{ij} \tag{2}$$

where equation (1) is the second stage and equation (2) is the first stage equation. y_{ij} is the outcome of interest, either living or working in the university region or living or working outside a metropolitan area eight years after applying for university for applicant i in location j, respectively. c_{ij} is an indicator variable for individual i graduating (not graduating) from university in the first choice location j; x is a vector of individual characteristics, including age, gender, ethnicity, mother's and father's education level, and an indicator variable for applicant i living in a metropolitan region at the YOA. a_{ik} is an indicator variable for the alternative university location k for individual i being in a metropolitan region. λ_j and δ_t are field of study and YOA fixed effects, respectively. $f(r_{ij})$ is a function of the distance to the GPA cutoff, which is the difference in high school GPA of individual i and the GPA admission cutoff for education in university location j [7]. ε_{ij} is the error term.

In the second stage equation 2, I instrument the indicator variable for graduation from university location j with the dummy variable d_{ij}, which indicates if individual i's GPA is above (below) the admission cutoff for education j. The second stage also includes the vector of individual characteristics as in the first stage along with fixed effects for the field

[6]See section 6.5.3 for estimations following Kirkeboen et al. (2016) methods. The results are similar in magnitude to the results from estimating the fuzzy RD model, as in equations 1 and 2.

[7]The preferred specification includes the distance to the admission cutoff, which is supported by the Akaike information criterion as the preferred specification when comparing models with a second- to a fourth-order polynomial of the running variable. This selection process is from Lee and Lemieux (2010)

of study and YOA, τ_j and ϕ_t.

The assumption underlying this strategy is that applicants who end up being just above or below the admission cutoffs are essentially identical ex-ante. Assuming that the applicants cannot perfectly predict the admission cutoff, I estimate the local average treatment effect (LATE) of graduation in a specific location on the choice of where to live and work eight years after applying for university.

The model can also estimate the effect of receiving an offer for an education program in a specific university location without restricting on graduation from the location. This will instead of the estimated LATE give me the intention to treat estimates (ITT), as all applicants who get an offer are not necessarily going to accept it, and some student will also drop out of a given education (see section 6.5.1. for the ITT estimates).

6 Results

6.1 Validity of the research design

This section examines the validity of the fuzzy RD design. First, I present the discontinuities in the offer and completion of university education across the distance between applicants' high school GPA and the admission cutoff. Second, I present figures of predicted probabilities across the distance to the admission cutoff on observables, and third, I present a balancing check for individual background characteristics across the distance to the admission cutoff. If the design is valid, I do not expect any discontinuities on observables or the predicted probabilities of living or working in a non-metropolitan region.

Figure 4 shows the share of applicants who receive an offer or graduates from their second choice university location. The figure for the first choice university location is included in Appendix A.1 [8]. The x-axis displays the distance between an applicant's high

[8] The discontinuity is the same; however the instrument is just reversed in Figure 10 in Appendix A.1

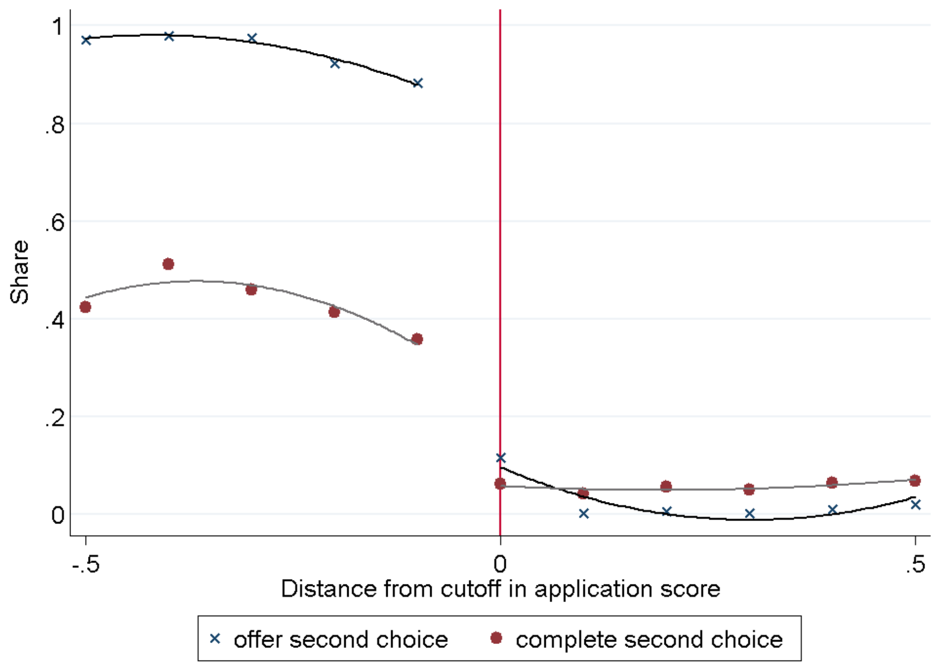
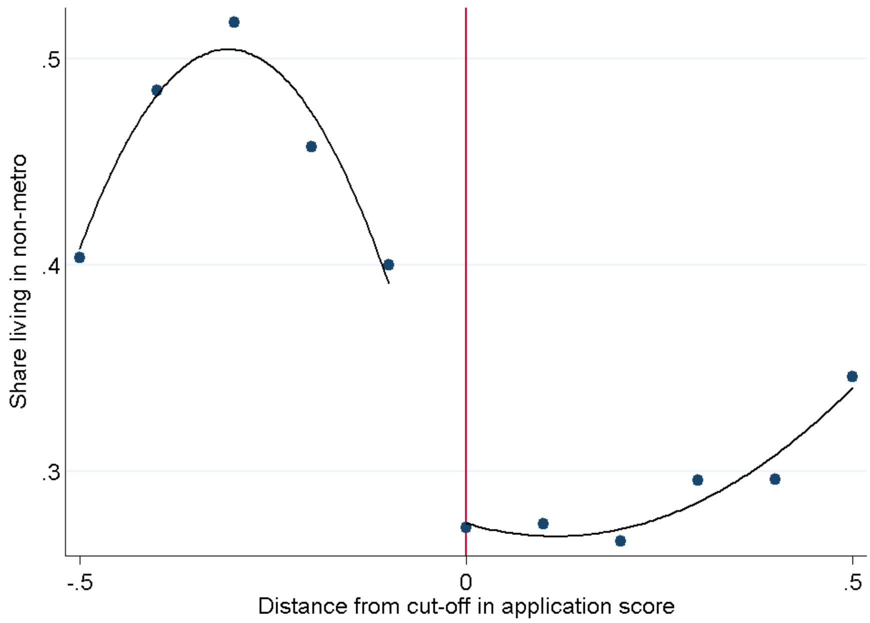


Figure 4: Admission cutoffs and next-best location offer and completion

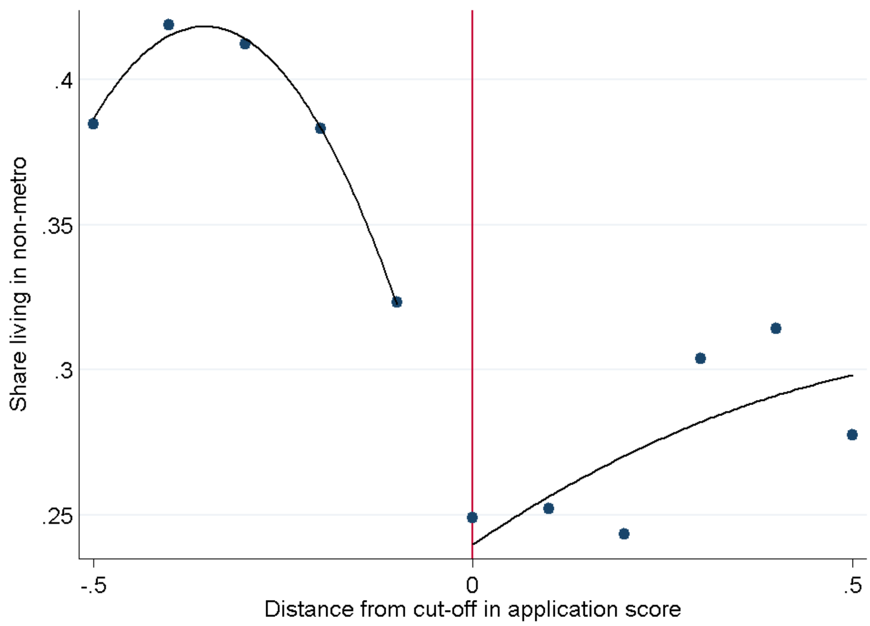
school GPA and the admission cut-off for the first choice university location. I pool the admission cutoffs and normalize them. A zero on the x-axis corresponds to having an application score that exactly equals the admission cutoff.

The figure shows, that the probability of receiving an offer for the second choice location decreases by around 80 percent at the admission cut-off. The probability of completing a university degree in the second choice university location decreases by more than 30 percent at the cutoff. The large differences between the offer and completion suggest that about 45 percent of the applicants who receive an offer for their second choice location choose not to accept it and, therefore, do not complete their studies in their second choice university location. Estimating the fuzzy RD model on the completion of second choice education will yield the estimates of the LATE, whereas estimating the effect of offer of second choice location will yield the ITT estimates of university location. Section 6.2 presents the effect of completion (the LATE), and the ITT estimates are presented in



Notes: This figure shows the applicants who have applied for a second choice non-metropolitan university location, and the share of this group living in a non-metropolitan region eight years after applying for university

Figure 5: Admission cutoffs and share living in non-metropolitan region eight years after for non-metropolitan second choice



Notes: This figure shows the applicants who have applied for a first choice metropolitan university location, and the share of this group living in a non-metropolitan region eight years after applying for university

Figure 6: Admission cutoffs and share living in non-metropolitan region eight years after for metropolitan first choice

section 6.5.1.

Figures 5 and 6 show the outcome variables of living in a non-metropolitan region eight years after applying for university for two subsamples. Figure 5 shows the share of the applicants who apply to a second choice non-metropolitan university location, regardless of whether the first choice location is metropolitan or non-metropolitan. The figure clearly shows that a higher share of the applicants are living in a non-metropolitan region if they are below the admission cutoff for their first choice education and, therefore, receive an offer for their second choice.

Figure 6 shows the share of applicants living in a non-metropolitan region for applicants who apply for a metropolitan university as their first choice, regardless of whether their second choice is metropolitan or non-metropolitan. Being above the cutoff and, therefore, receiving an offer for their first choice metropolitan education clearly reduce the share of the applicants living in a non-metropolitan region eight years after applying for university.

Figure 7 shows a balancing test of the average age, the share of males, and the share of highly educated mothers and fathers. Figure 8 shows the predicted probabilities of moving and working in a non-metropolitan area eight years after YOA. The predicted probabilities are a weighted average of observables. This weighted average is constructed by regressing the outcome of living in a non-metropolitan region on all observables. Therefore, finding no discontinuities over the distance from cutoff in application score indicates that the weighted average of all observables is balanced across the cutoff.

Figures 8 and 7 indicate that the applicants are close to identical on observables on each side of the admission cutoff. Conditioned on observables, there does not seem to be a higher probability of moving or working in a non-metropolitan area across the distance between GPA and the admission cut-off. Similarly, there do not seem to be any differences in characteristics such as gender, age and parental education of having a GPA above or

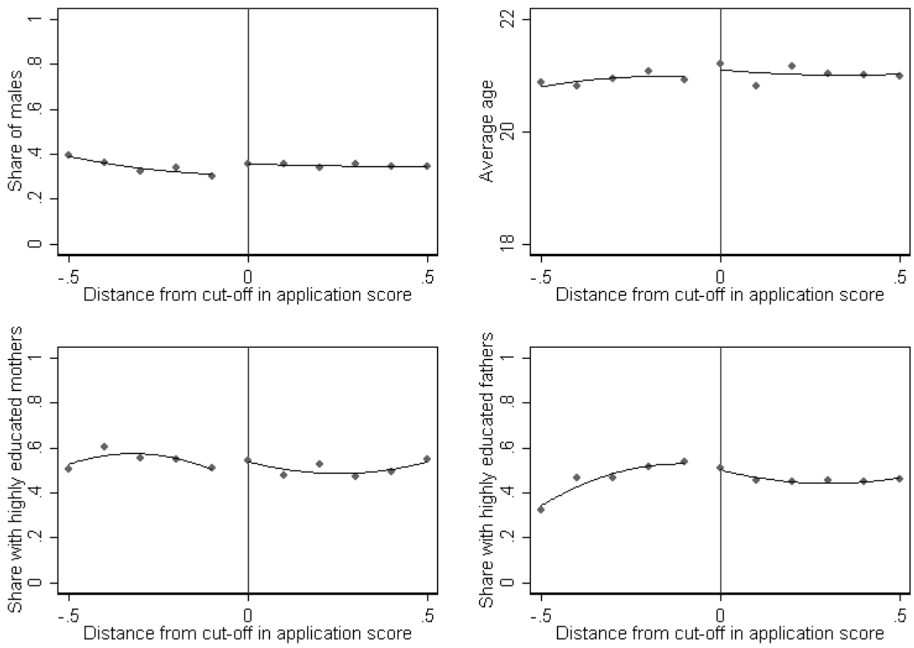
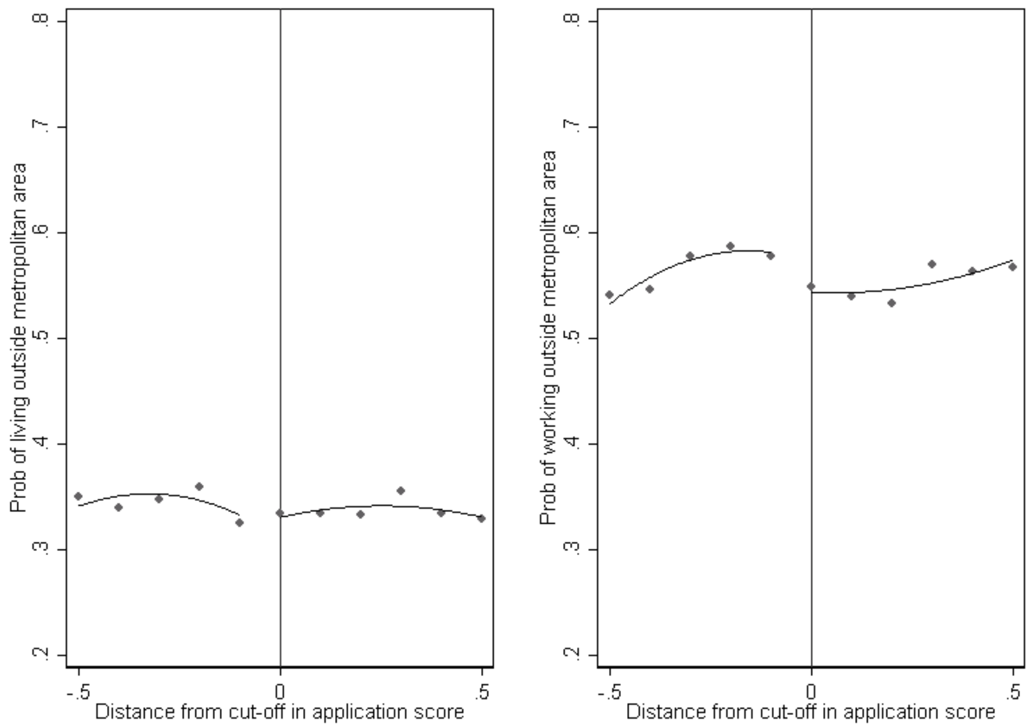


Figure 7: Predicted probability of living and working in non-metropolitan area eight years after applying for university



Notes: This figure shows the predicted probability of living in a non-metropolitan region(left) or working in a non-metropolitan region (right) eight years after applying for university. The probabilities are predicted using an OLS regression of the probability of either living or working in a non-metropolitan region on gender, age, high school GPA, parental education and an indicator for living in a metropolitan region at the year of application

Figure 8: Predicted probability of living and working in non-metropolitan area eight years after applying for university

Table 6: The effect of completing university in first choice location

	Metropolitan first choice			
	live in non-metro	work in non-metro	live in uni region	work in uni region
Grad from first choice	-0.388***	-0.264*	0.589***	0.386***
	(0.112)	(0.140)	(0.125)	(0.117)
F-stat	72.81	57.53	72.28	72.28
Observations	2,086	1,093	2,104	1,095
	Non-metropolitan first choice			
Grad from first choice	0.133	0.364	0.0231	-0.177
	(0.255)	(0.319)	(0.212)	(0.251)
F-stat	12.73	8.16	13.56	8.40
Observations	367	184	370	186

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, location of first choice education, municipality of residence in YOA, field of study and YOA fixed effects. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

below an admission cut-off.

Taken together, this section provides evidence of the validity of the research design.

6.2 The effect of university location on choice of where to live and work

This section presents the regression results of the effect of university location. First, the section presents the results of completing university in the first choice metropolitan university region and the first choice non-metropolitan region. Second, the section presents results for completing university in the second choice metropolitan university location and the second choice non-metropolitan university location.

The results from my main specification from equations (1) and (2) are presented in Tables 6 and 7. Table 6 shows the results of completing preferred education in the metropolitan and non-metropolitan locations, respectively. Column 1 shows the effect of the university location on living in a non-metropolitan region eight years after applying for university, and column 2 shows the effect of working in a non-metropolitan region. Columns 3 and 4 show the results of living or working in the first choice university location.

The sample sizes vary greatly between preferring to study in a metropolitan area (the top part of Table 6) and in a non-metropolitan region (the bottom part of Table 6). This is because most university applicants in the sample apply for a metropolitan university location as their first choice as explained in the data section.

Focusing on the graduates who complete university in their first choice metropolitan region, I find that they are less likely to live and work in a non-metropolitan region eight years after applying for university compared to graduates who completed their studies in their second choice location. Looking at columns 3 and 4, I find that the graduates are significantly more likely to live and work in the region of their preferred choice than the graduates who did not complete their studies in their preferred metropolitan region. This suggests that even though two students initially have the same preferences for a specific metropolitan university location, the probability of living and working in that location is much higher for the student who end up completing his or her studies in that location. More specifically, the students who graduate from their first choice university location are almost 60 percentage points more likely to live in the region and almost 40 percentage points more likely to work in the region compared to a student who also preferred to study in that specific metropolitan region but was not accepted and instead graduated from his or her second choice location.

I do not find any similar differences between the students who complete university in their first choice non-metropolitan region. However, it should be noted that sample sizes are small, and the instrument is not performing as well as for the sample with a metropolitan university as their first choice.

Table 7 shows the results for completion of university in the second choice location. The table is structured in the same way as Table 6; however instead of focusing on the group of students who graduate from their first choice university location, Table 7 presents the results for students who were accepted for their preferred university location but instead enrolled in and graduated from their second choice university location.

Recall that more than 80 percent of the sample prefer a metropolitan university location. This means that when evaluating the results from Table 7, the comparison group will consist primarily of graduates from a metropolitan first choice university location. For example, when looking at the top part of Table 7 for graduates from their second choice location in a metropolitan region, the significant estimates from columns 3 and 4 should be interpreted as it is 52 percentage points and 66 percentage points more likely for a graduate from his or her second choice metropolitan region to live and work in the metropolitan second choice region compared to a graduate who graduated from his or her first choice metropolitan region. This will, therefore primarily consist of different combinations of Aarhus and Copenhagen. If a graduate, for example, had Aarhus as first choice and Copenhagen as second choice, the effect of 0.523 from column 3 should be interpreted as: The effect of graduating from a second choice location (Copenhagen) is a 52.3 percentage point higher probability of living in Copenhagen compared to a graduate who graduated from the first choice university location in Aarhus. This, thereby, shows that two applicants who initially had the same location preferences now after graduating from differently ranked location choices are more likely to live where they studied. This may not be very surprising when looking at studying in metropolitan regions as they have larger labor markets than the non-metropolitan regions, and the metropolitan regions are better at retaining their graduates (Ahlin et al., 2018; Venhorst et al., 2011; Winters, 2011; Haapanen and Tervo, 2012; Faggian and McCann, 2009).

Nonetheless, the results presented in the bottom part of Table 7 show that graduating from university in a second choice non-metropolitan region compared to graduating from first choice location (most likely in a metropolitan area) is associated with an almost 50

Table 7: The effect of completing university in second choice location

	Metropolitan second choice			
	live in non-metro	work in non-metro	live in uni region	work in uni region
Grad from second choice	0.00760	-0.176	0.523***	0.662***
	(0.137)	(0.193)	(0.129)	(0.155)
F-stat	66.41	43.44	68.61	48.53
Observations	1,354	694	1,365	690
	Non-metropolitan second choice			
Grad from second choice	0.463***	0.476***	0.390***	0.609***
	(0.145)	(0.182)	(0.0983)	(0.143)
F-stat	79.32	59.18	76.41	61.47
Observations	1,099	583	1,109	582

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, location of first choice education, municipality of residence in YOA, field of study and YOA fixed effects. * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

percentage points higher probability of living and working outside the metropolitan regions. Also, the probability of living and working in the specific second choice university region is respectively 39 and 61 percentage points higher for students who graduated from that specific location. This is suggestive evidence of non-metropolitan universities being able to retain graduates. Even though most graduates preferred to study in a metropolitan region, some graduates will still live and work in the non-metropolitan second choice university region eight years after applying for university.

6.3 Heterogeneous effects

The findings so far suggest that graduates from second choice university locations in non-metropolitan regions are more likely to live and work outside the metropolitan regions eight years after applying for university compared to graduates who completed university in their first choice location. Also, I find evidence of universities being able to retain some of their graduates even though it was not necessarily the first choice location of the graduates. This section investigates whether these results are heterogeneous across gender, municipality of residence at YOA, and parental education. For convenience, I only present the results from graduating from second choice university location; however the results for graduating from preferred university location can be found in Appendix A.2.

Table 8 presents the regression results across the subgroups. The sample is stratified by place of residence in YOA in panels A.1 and A.2 by gender in panels B.1 and B.2, and on parental education in panels C.1 and C.2. The full outputs of each subgroup analysis are included in Appendix A.2.

The subsamples across the place of residence in the YOA are divided into the group of graduates who lived in a metropolitan region in the YOA and the group of graduates who lived in a non-metropolitan region in the YOA. Focusing on the results of these subgroups presented in panels A.1 and A.2, I find no differences between the groups for graduating from a university in the second choice metropolitan region. The results for graduating from a second choice university in a non-metropolitan region show that the graduates who initially lived in a non-metropolitan region have a higher probability of living and working in a non-metropolitan region than the graduates who also initially lived in a non-metropolitan region but graduated from their first choice university location (mostly in metropolitan areas). Likewise, the graduates from the second choice non-metropolitan location are also more likely to live in the region of the second choice university compared to graduates from their first choice location.

For the graduates who initially lived in a metropolitan region, I find a significant positive effect of graduating from the second choice non-metropolitan university location on the probability of living in a non-metropolitan region compared to graduates who completed university in their preferred university location. For the outcomes working in a

Table 8: Heterogeneous effects

	Metropolitan second choice			
	live in non-metro	work in non-metro	live in uni region	work in uni region
Graduate from second choice				
Baseline estimates from Table 7	0.00760	-0.176	0.523***	0.662***
	(0.137)	(0.193)	(0.129)	(0.155)
B.1 Place of residence in YOA				
Living in non-metropolitan area	0.0118	-0.265	0.524***	0.683***
	(0.161)	(0.221)	(0.139)	(0.172)
Living in metropolitan area	-0.0913	-0.298	0.575**	0.742***
	(0.227)	(0.378)	(0.257)	(0.292)
C.1 Gender				
Female	0.103	-0.0387	0.386***	0.538***
	(0.160)	(0.233)	(0.143)	(0.193)
Male	-0.299	-0.499	0.892***	0.969***
	(0.269)	(0.319)	(0.273)	(0.267)
D.1 Parental education				
Parents have higher educ	-0.0348	-0.126	0.497***	0.704***
	(0.156)	(0.213)	(0.151)	(0.164)
Parents do not have higher educ	0.147	-0.0197	0.588**	0.803**
	(0.330)	(0.447)	(0.0291)	(0.390)
	Non-metropolitan second choice			
Baseline estimates from Table 7	0.463***	0.476***	0.390***	0.609***
	(0.145)	(0.182)	(0.0983)	(0.143)
B.2 Place of residence in YOA				
Living in non-metropolitan area	0.403**	0.572***	0.456***	0.651***
	(0.173)	(0.206)	(0.119)	(0.160)
Living in metropolitan area	0.815***	-0.317	0.160	0.374
	(0.288)	(0.436)	(0.137)	(0.269)
C.2 Gender				
Female	0.552**	0.516*	0.563***	0.656***
	(0.237)	(0.291)	(0.169)	(0.240)
Male	0.348**	0.210	0.236***	0.587***
	(0.176)	(0.210)	(0.117)	(0.148)
D.2 Parental education				
Parents have higher educ	0.398**	0.384	0.215*	0.609***
	(0.156)	(0.233)	(0.114)	(0.196)
Parents do not have higher educ	0.547**	0.842***	0.650***	0.561***
	(0.240)	(0.287)	(0.164)	(0.198)

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, location of first choice education, municipality of residence in YOA, field of study and YOA fixed effects. * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

non-metropolitan region or living or working in the same region as the second choice university is placed, I find no significant effects of university location.

The evidence combined suggest that the finding of non-metropolitan universities being able to retain graduates even though they are not the preferred location choice (from Table 7), is driven by the graduates who initially lived outside the metropolitan regions. The large effect of second choice university location for students who lived in a metropolitan region in YOA from panel A.2 could be caused by these students moving back to the metropolitan regions to work, but instead of living in the metropolitan regions, they choose to live in the less urbanized regions surrounding the metropolitan regions. I do not find any evidence of this group of graduates being more likely to stay in the non-metropolitan university regions, suggesting that the graduates who used to live in the metropolitan regions either do not move to study or move when they graduate from a non-metropolitan university.

Focusing on the subgroups by gender, I discover that for both male and female applicants, the probability of living in a second choice metropolitan university region increases if they graduate from a university in the second choice location compared to graduating from their preferred choice of university location. The effect sizes are almost twice as large for males than for females, suggesting that males are more likely to be retained in the metropolitan regions. The results are reversed when looking at the coefficients from graduating from a second choice non-metropolitan university location. Females seem more likely to live and work outside the metropolitan regions if they graduate from their second choice non-metropolitan university location compared to graduating from their preferred location. They are also more likely to live and work in the location of the second choice university that they graduate from compared to the graduates who complete their education in a preferred university location. The differences are smaller in magnitude for males, suggesting that females are more likely to stay in the non-metropolitan regions compared to their male counterparts.

These results are surprising as the trend usually inclines toward females living in the cities and males in the less urbanized parts of Denmark. It could be a story of different marriage markets but it is unfortunately outside the scope of this paper to delve into these gender differences.

For the subgroups by parental education, I stratify the sample by the graduates who have at least one parent with a completed higher education and graduates who have parents with no higher education. According to panel C.1, there are no differences between graduates with highly educated parents and graduates with parents without higher education. The effect of graduating from a metropolitan second choice university location on the probability of living and working in this specific second choice university region compared to graduates from their preferred university location is similar in magnitude and positive and significant for both subgroups.

The results of graduating from a non-metropolitan second choice location in panel C.2 are generally mixed with no clear pattern. The graduates with parents without higher education are more likely to live and work outside the metropolitan regions than parents with higher education, and they are also more likely to live in the same region as their second choice university location compared to the graduates from their first choice university location. The graduates across the parental education categories are equally as likely to work in the second choice university location if they have graduated there compared to the graduates from their preferred university location.

To sum up, the heterogeneity analysis suggests that male graduates are more easily retained in the metropolitan region than their female counterparts. Furthermore, the evidence from the main regressions in Table 7 of non-metropolitan universities being able to retain their graduates seems to be driven by graduates who initially lived outside the metropolitan regions.

Table 9: Labor market returns to completing university in first choice location

	Metropolitan first choice		Non-metropolitan first choice	
	log earnings	unemployment	log earnings	unemployment
Grad from first choice	-0.00919	-0.0106	0.996	-0.0769
	(0.266)	(0.0280)	(0.733)	(0.100)
F-stat	94.69	103.17	7.46	13.62
Observations	1,696	1,796	293	325

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, location of second choice education, municipality of residence in YOA, field of study and YOA fixed effects. * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

6.4 Labor market returns to university location

The literature on urban wage premiums suggests that the most competent people move and work in the large metropolitan regions. Ahlin et al. (2018) find that the ability sorting toward the metropolitan regions happens primarily when choosing where to study. In this paper, I compare university graduates on both sides of the admission cutoff in different university locations. As I focus on the graduates who have GPAs close to the admission cutoff, they should by construction be very similar (at least in terms of abilities measured by GPA). Therefore, estimating equations (1) and (2) and changing the outcome variable to labor market returns, such as earnings and unemployment, I study the effect of university location on labor market returns. For example, it could be that there is a penalty associated with graduating from your second choice education or by graduating from a non-metropolitan region where the labor market for university graduates might not be as good as in the metropolitan regions.

Table 9 (graduating from the first choice university location) and Table 10 (graduating from the second choice location) show the results of such regressions. I estimate the effect for both metropolitan and non-metropolitan regions and focus on log earnings and unemployment eight years after applying for university. The results from the two tables

Table 10: Labor market returns to completing university in second choice location

	Metropolitan second choice		Non-metropolitan second choice	
	log earnings	unemployment	log earnings	unemployment
Grad from second choice	-0.144 (0.379)	-0.00665 (0.0374)	-0.154 (0.344)	-0.0415 (0.0415)
F-stat	60.03	68.36	87.76	91.75
Observations	1,113	1,198	876	923

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, location of first choice education, municipality of residence in YOA, field of study and YOA fixed effects. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

show that there are no returns to university location from either metropolitan or non-metropolitan university locations or from graduating from the first or second choice university location. The results are in line with the findings of Kirkeboen et al. (2016), who find that once they control for the field of study, there is little evidence of any returns to graduating from a more selective institution.

6.5 Robustness of results

6.5.1 The effect of receiving an offer

One limitation of the main results presented in section 6.2 is that individuals can choose to enroll when they receive an offer. As Figure 8 from section 6.1 shows, nearly 45 percent of the students receive an offer for their second choice but does not complete their studies. When estimating the LATE of completing education in individuals' second choice location, some students could have opted out simply by not choosing to accept the offer at their second choice university location, or perhaps they dropped out of education at a later point. The effect of receiving an offer would for policy purposes be the more relevant margin as it is not possible to distinguish how many students who receive an offer will actually accept and graduate from a given university. Therefore, I estimate the model for

Table 11: The effect of getting offer from university in first choice location

	Metropolitan first choice			
	live in non-metro	work in non-metro	live in uni region	work in uni region
Offer from first choice	-0.159***	-0.127*	0.239***	0.167***
	(0.0454)	(0.0666)	(0.0490)	(0.0636)
F-stat	1532.32	994.81	1561.35	1012.48
Observations	2,086	1,093	2,104	1,095
	Non-metropolitan first choice			
Offer from first choice	0.0950	0.397	0.0171	-0.187
	(0.180)	(0.311)	(0.156)	(0.271)
F-stat	47.89	13.58	48.22	14.85
Observations	367	184	370	186

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, location of second choice education, municipality of residence in YOA, field of study and YOA fixed effects. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

only having been offered a slot and not subject to completion (which instead, yield an ITT estimate).

The results are presented for receiving an offer for the first choice location in Tables 11 and 12 for receiving an offer for the second choice location. Compared to the estimates from Tables 6 and 7, the results look similar, although they are slightly smaller in magnitude.

6.5.2 Restricting the sample around the admission cutoff

All the results presented in section 6 so far are conducted on a sample around the admission cutoff of +/- 0.5-grade points. There is a trade-off between precision and statistical power when it comes to choosing the bandwidth in an RD design. On the one hand, choosing a bandwidth really close to the cutoff will give very precise estimates as the groups on each side are very similar (here in regards to high school GPA); however, the

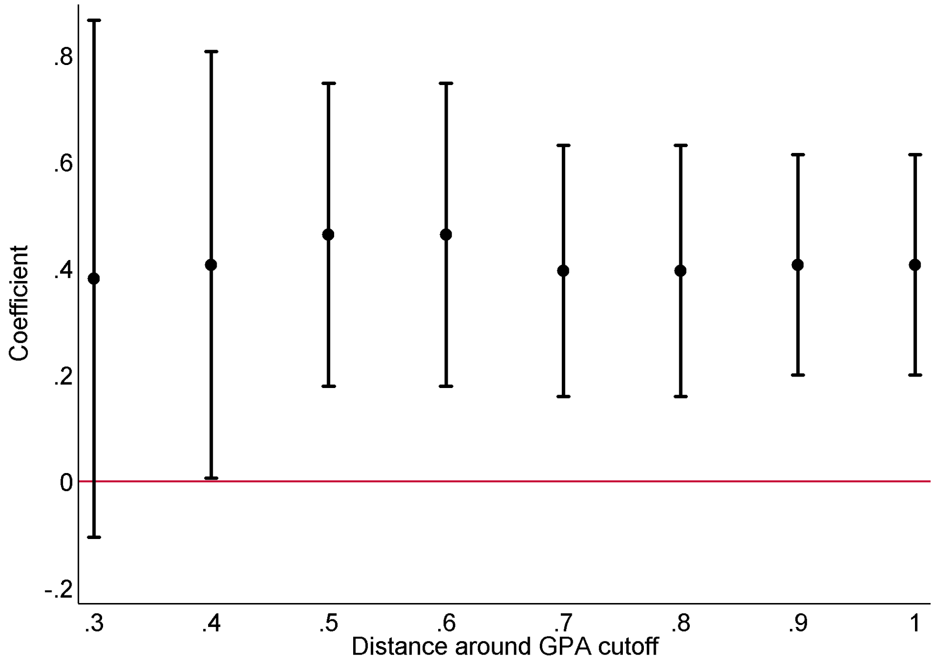
Table 12: The effect of getting offer from university in second choice location

	Metropolitan second choice			
	live in non-metro	work in non-metro	live in uni region	work in uni region
Offer from second choice	0.00323	-0.0859	0.225***	0.349***
	(0.0580)	(0.0935)	(0.0576)	(0.0903)
F-stat	869.0	465.94	877.38	419.03
Observations	1,354	694	1,365	690
	Non-metropolitan second choice			
Offer from second choice	0.218***	0.254***	0.179***	0.325***
	(0.0677)	(0.0959)	(0.0441)	(0.0723)
F-stat	556.16	328.09	570.95	338.00
Observations	1,099	583	1,109	582

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, location of second choice education, municipality of residence in YOA, field of study and YOA fixed effects. * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

sample will be small, and the standard errors are likely to be larger. On the other hand, choosing a wider bandwidth around the cutoff will give a larger sample; however, it could also give less precisely estimated coefficients.

In Figure 9, I estimate the effect of graduating from a second choice non-metropolitan university on the probability of living in a non-metropolitan region eight years after applying for university for different bandwidths around the admission cutoff (the estimates correspond to the results presented in column 1 of Table 7). The coefficients are fairly similar for bandwidths from around +/- 0.3-grade points up to +/- 1-grade point. The standard errors are larger for the smallest bandwidth of 0.3-grade points around the cutoff and decrease as the bandwidth gets larger, as expected. The results of the effect of graduating from a second choice non-metropolitan university on working in a non-metropolitan region as well as the effect of graduating from a first choice metropolitan university on living and working in a non-metropolitan region are similar across different bandwidths



Notes: The figure shows the regression estimates and the 95-percent confidence interval using the bandwidths around the GPA admission cutoff shown on the horizontal axis as sample restriction. The outcome variable is living in a non-metropolitan region eight years after YOA.

Figure 9: Bandwidth analysis: The effect of completing university in second choice non-metropolitan location on the probability of living in non-metropolitan region

(see Appendix A.3).

6.5.3 Replicating the methods by Kirkeboen et al. (2016)

In their paper, Kirkeboen et al. (2016) develop a method for estimating fuzzy RD models with an unordered choice as the dependent variable. With an unordered choice, it is not enough to estimate the fuzzy RD model with a valid instrument using 2SLS. The margin of choice also needs to be taken into account. To do so, Kirkeboen et al. (2016) estimates a matrix of payoffs for all combinations of first choice and second choice fields of study (and institutions). Because of data limitations, I cannot replicate their analysis directly for all samples and, instead, I control for the margin of choice by including an indicator variable for the alternative university location from which the students graduate.

Table 13: The effect of university location on living in a non-metropolitan region eight years after YOA. Following the methods by Kirkeboen et al. (2016)

First choice Second choice	Non-metro Non-metro	Non-metro Metropolitan	Metropolitan Non-metro	Metropolitan Metropolitan
Complete first	-	-	-0.402***	0.00466
choice	-	-	(0.0797)	(0.0960)
Complete second	-	-	0.400***	-0.00550
choice	-	-	(0.0802)	(0.113)
Observations	-	-	1,595	1,609

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, location of alternative choice education, municipality of residence in YOA, field of study and YOA fixed effects. * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

To test how much my results differ compared to if I had applied the methods by Kirkeboen et al. (2016) I replicate the main results from Tables 6 and 7 using their methods. Because of data limitations, I can only apply their method to the two largest samples of students, namely (1) the students who apply to a first choice metropolitan university and a second choice metropolitan university (combinations of Aarhus and Copenhagen) and (2) students who apply for a metropolitan university as first choice and a non-metropolitan university as their second choice. Table 14 shows the results for the outcome variable living in a non-metropolitan region eight years after applying for university. The top part of the table presents the estimates for completing first choice education and the bottom part for completing second choice education. The first row shows the combination of first and second choice university locations. For example, ”metropolitan / non-metropolitan” illustrates the estimate for having a metropolitan university as first choice and a non-metropolitan university as second choice.

The results show that graduating from a first choice metropolitan university compared to graduating from a second choice non-metropolitan university is associated with a 40 percentage points lower probability of living in a non-metropolitan region eight years after applying for university. The results are exactly reversed when looking at graduates from a second choice non-metropolitan university compared to graduates from a metropolitan first choice university. The estimates are similar in magnitude to the main results from Tables 6 and 7. The other combinations of first and second choice university locations are not associated with any significant differences in the probability of living in a non-metropolitan region eight years after applying for university. See Appendix A.4 for the analyses following these methods with the other outcome variables from Tables 6 and 7.

6.5.4 Commuting to university

There is a possibility that some students simply move to the metropolitan regions and commute to the university in non-metropolitan regions if they do not get accepted for their preferred metropolitan university location. For example, 25 percent of the applicants who apply for University of Copenhagen have Roskilde University as their second choice education. The infrastructure between Roskilde and Copenhagen is very convenient, and the

Table 14: The Effect of Completing University in Second Choice Location. Excluding Roskilde University

	Metropolitan second choice			
	live in non-metro	work in non-metro	live in uni region	work in uni region
Graduate from second choice	0.0563	-0.0403	0.510***	0.648***
	(0.141)	(0.185)	(0.122)	(0.150)
F-stat	64.86	46.91	67.05	50.54
Observations	1,222	638	1232	635
	Non-metropolitan second choice			
Graduate from second choice	0.547***	0.574***	0.633***	0.724***
	(0.208)	(0.203)	(0.158)	(0.174)
F-stat	44,53	46.86	43.11	50.61
Observations	735	411	741	410

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, location of first choice education, municipality of residence in YOA, field of study and YOA fixed effects. * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

commuting time by train is about 20 minutes. Therefore, most of the students at Roskilde University live in Copenhagen. If the students live outside the university region, it is less likely that they will develop specific regional human capital or social ties to employers as Krabel and Flöther (2014) conclude as being important factors for entering the labor market in the university region. To check how these commuting students affect the main results from Table 7, I exclude the students from Roskilde University and rerun the model.

Table 14 presents the results of the model, excluding students from Roskilde University. The results of graduating from a metropolitan second choice education do not change compared to the main results, whereas the results of graduating from a second choice non-metropolitan university location are larger when excluding Roskilde University. The effect of graduating from a non-metropolitan second choice location compared to graduating from a first choice (metropolitan) university location is 54.7 and 57.4 percentage point higher probability of living and working in a non-metropolitan region eight years after applying for university. Similarly, the results of living and working in the same region as the second choice non-metropolitan university are 63.3 and 72.4 percentage point higher compared to graduates from a first choice university location. The results confirm my hypothesis that graduates who do not live in the university region while studying are less likely to enter the labor market in the university region.

7 Conclusion

This paper examines the effect of university location on the choice of where to live and work of Danish university graduates. The fact that the annual admission cutoffs into higher education are unpredictable allows me to account for self-selection into a given university location. Using a fuzzy RD design that takes into account that the choice of field and location of education is an unordered choice, I estimate the causal effect of studying either in a metropolitan or a non-metropolitan region on where graduates live and work eight years after applying for university. My findings suggest that the location of study does have an effect on where graduates choose to live and work. Prospective students who did not get accepted for their preferred university location and instead graduate from their second choice non-metropolitan university are more likely to live and work outside the metropolitan areas eight years after applying for university compared to graduates from their first choice university location.

The results are based on university applicants who apply for the education programs within the same field but in different locations. I do not consider the share of university applicants who are more willing to compromise on the field of study rather than the location because the identification of the model relies on the exogenous variation across locations. The moving patterns of these university students are yet to be examined.

Another finding of this paper is that universities can retain some of their graduates.

Some of the graduates who did not get accepted for their first choice university location but instead graduate from a non-metropolitan second choice university are retained in their university location, despite the fact that most of these graduates prefer to study in a metropolitan region.

From a policy perspective, this study provides evidence that pushing university applicants to study in a non-metropolitan region will result in some of these graduates staying and entering the labor market in these regions. Therefore, it could be a useful policy tool to increase the share of highly educated students in the local non-metropolitan regions.

There are several avenues for further research. Because of the empirical strategy of this paper, it is possible to estimate the causal effect of university location on the choice of where to live and work. However, in such a setup, it is not possible to investigate what makes some graduates stay and others leave. Furthermore, combining the results from this paper with the previous findings in the literature that an attractive labor market is key to both attract and retain university graduates, it would be interesting to dig deeper into the synergies of educating specific types of graduates that are targeted at specific local labor markets. Perhaps an even stronger policy tool would be to place universities and large private companies geographically close so that they can complement each other (e.g., data science educations and tech companies).

References

Ahlin, L., Andersson, M., and Thulin, P. (2014). Market thickness and the early labour market career of university graduates: An urban advantage? *Spatial Economic Analysis*, 9(4):396–419.

Ahlin, L., Andersson, M., and Thulin, P. (2018). Human capital sorting: The "when" and "who" of the sorting of educated workers to urban regions. *Journal of regional science*, 58(3):581–610.

Card, D. (1993). Using geographic variation in college proximity to estimate the return to schooling. Technical report, National Bureau of Economic Research.

Currie, J. and Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*, 118(4):1495–1532.

Faggian, A. and McCann, P. (2009). Human capital, graduate migration and innovation in british regions. *Cambridge Journal of Economics*, 33(2):317–333.

Haapanen, M. and Tervo, H. (2012). Migration of the highly educated: Evidence from residence spells of university graduates. *Journal of Regional Science*, 52(4):587–605.

Heinesen, E. (2018). Admission to higher education programmes and student educational outcomes and earnings–evidence from denmark. *Economics of Education Review*.

Herzog Jr, H. W., Hofler, R. A., and Schlottmann, A. M. (1985). Life on the frontier: migrant information, earnings and past mobility. *The Review of Economics and Statistics*, pages 373–382.

Humlum, M. K., Kristoffersen, J. H., and Vejlin, R. (2017). College admissions decisions, educational outcomes, and family formation. *Labour Economics*, 48:215–230.

Kirkeboen, L. J., Leuven, E., and Mogstad, M. (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics*, 131(3):1057–1111.

Kotavaara, N., Kotavaara, O., Rusanen, J., and Muilu, T. (2018). University graduate migration in finland. *Geoforum*, 96:97–107.

Krabel, S. and Flöther, C. (2014). Here today, gone tomorrow? regional labour mobility of german university graduates. *Regional Studies*, 48(10):1609–1627.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2):281–355.

Moretti, E. (2004). Human capital externalities in cities. In *Handbook of regional and urban economics*, volume 4, pages 2243–2291. Elsevier.

Öckert, B. (2010). What's the value of an acceptance letter? using admissions data to estimate the return to college. *Economics of Education Review*, 29(4):504–516.

Öckert, B. (2012). On the margin of success? effects of expanding higher education for marginal students. *Nordic Economic Policy Review*, 1(2012):111–157.

OECD (2018). *Education at a Glance 2018*.

Spiess, C. K. and Wrohlich, K. (2010). Does distance determine who attends a university in germany? *Economics of Education Review*, 29(3):470–479.

Venhorst, V., Van Dijk, J., and Van Wissen, L. (2011). An analysis of trends in spatial mobility of dutch graduates. *Spatial Economic Analysis*, 6(1):57–82.

Winters, J. V. (2011). Why are smart cities growing? who moves and who stays. *Journal of regional science*, 51(2):253–270.

A Additional material

A.1 Discontinuities around admission cutoff for first choice offer and completion

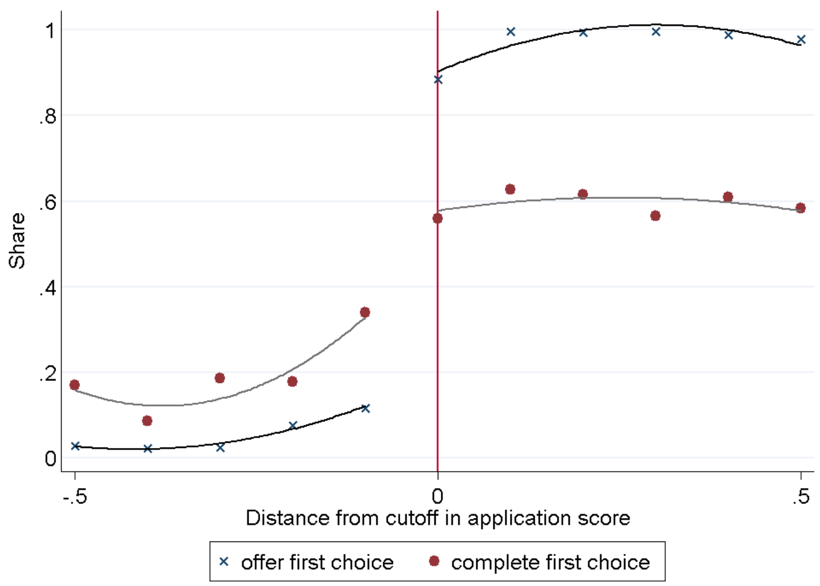


Figure 10: Admission cutoffs and first choice location offer and completion

A.2 Heterogeneity analysis

Table 15: Heterogeneous effects for second choice location. Gender

	Metropolitan second choice			
	live in non-metro	work in non-metro	live in uni region	work in uni region
Graduate from second choice				
Female	0.103	-0.0387	0.386***	0.538***
	(0.160)	(0.233)	(0.143)	(0.193)
F-stat	52.72	28.54	53.95	31.41
Observations	909	470	916	467
Male	-0.299	-0.499	0.892***	0.969***
	(0.269)	(0.319)	(0.273)	(0.267)
F-stat	14.63	14.04	15.56	15.34
Observations	445	224	449	223
	Non-metropolitan second choice			
Female	0.552**	0.516*	0.563***	0.656***
	(0.237)	(0.291)	(0.169)	(0.240)
F-stat	30.87	21.22	29.09	21.79
Observations	692	379	697	378
Male	0.348**	0.210	0.236***	0.587***
	(0.176)	(0.210)	(0.117)	(0.148)
F-stat	51.24	51.24	51.37	51.24
Observations	407	204	412	204

Notes: Robust standard errors in parentheses. All models include controls for age, ethnicity, mother and father's education, municipality of residence in YOA, location of first choice education, field of study and YOA fixed effects. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table 16: Heterogeneous effects for second choice location. Living in a metropolitan area when applying for university

	Metropolitan second choice			
	live in non-metro	work in non-metro	live in uni region	work in uni region
Graduate from second choice				
Living in non-metropolitan area	0.0118	-0.265	0.524***	0.683***
	(0.161)	(0.221)	(0.139)	(0.172)
F-stat	52.97	35.49	55.14	39.35
Observations	1,075	554	1,085	550
Living in metropolitan area	-0.0913	-0.298	0.575**	0.742***
	(0.227)	(0.378)	(0.257)	(0.292)
F-stat	17.28	11.34	17.68	11.34
Observations	279	140	280	140
	Non-metropolitan second choice			
Graduate from second choice				
Living in non-metropolitan area	0.403**	0.572***	0.456***	0.651***
	(0.173)	(0.206)	(0.119)	(0.160)
F-stat	60.24	46.21	58.59	49.24
Observations	878	483	887	482
Living in metropolitan area	0.815***	-0.317	0.160	0.374
	(0.288)	(0.436)	(0.137)	(0.269)
F-stat	16.24	9.42	16.07	8.90
Observations	221	100	222	100

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, location of first choice education, field of study fixed effects, and YOA fixed effects. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table 17: Heterogeneous effects of second choice location. Parental education

	Metropolitan second choice			
	live in non-metro	work in non-metro	live in uni region	work in uni region
Graduate from second choice				
Parents have higher educ	-0.0348	-0.126	0.497***	0.704***
	(0.156)	(0.213)	(0.151)	(0.164)
F-stat	47.07	34.04	47.07	34.92
Observations	876	425	876	375
Parents do not have higher educ	0.147	-0.0197	0.588**	0.803**
	(0.330)	(0.447)	(0.0291)	(0.390)
F-stat	14.02	6.64	14.02	7.81
Observations	415	228	415	227
	Non-metropolitan second choice			
Parents have higher educ	0.398**	0.384	0.215*	0.609***
	(0.156)	(0.233)	(0.114)	(0.196)
F-stat	48.47	34.84	48.47	34.92
Observations	711	376	711	375
Parents do not have higher educ	0.547**	0.842***	0.650***	0.561***
	(0.240)	(0.287)	(0.164)	(0.198)
F-stat	30.56	25.03	30.56	23.39
Observations	338	177	338	178

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, location of first choice education, municipality of residence in YOA, field of study and YOA fixed effects. * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

Table 18: Heterogeneous effect of first choice location. Gender

	Metropolitan first choice			
	live in non-metro	work in non-metro	live in uni region	work in uni region
Graduate from first choice				
Female	-0.513***	-0.473**	0.697***	0.420**
	(0.178)	(0.218)	(0.198)	(0.194)
F-stat	31.23	24.60	30.68	24.91
Observations	1,340	714	1,351	716
Male	-0.230*	0.0213	0.471***	0.184
	(0.135)	(0.176)	(0.152)	(0.161)
F-stat	49.02	37.27	49.46	38.42
Observations	746	379	753	379
	Non-metropolitan first choice			
Female	0.0681	0.645	0.220	-0.337
	(0.316)	(0.601)	(0.277)	(0.436)
F-stat	8.37	3.63	8.32	3.54
Observations	261	135	262	136
Male	0.300	0.113	-0.234	0.0825
	(0.561)	(0.290)	(0.425)	(0.258)
F-stat	2.25	4.56	3.10	4.59
Observations	106	49	108	50

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, field of study, location of first choice education, year of application, and municipality of residence in YOA. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table 19: Heterogeneous effects for first choice location. Living in a metropolitan area when applying for university

	Metropolitan first choice			
	live in non-metro	work in non-metro	live in uni region	work in uni region
Graduate from first choice				
Living in non-metropolitan area	-0.322***	-0.299**	0.483***	0.338**
	(0.115)	(0.148)	(0.124)	(0.136)
F-stat	72.35	53.29	70.79	53.63
Observations	1,653	887	1,669	888
Living in metropolitan area	-0.880*	0.0925	1.287**	0.0697
	(0.511)	(0.376)	(0.602)	(0.288)
F-stat	4.51	6.04	5.08	6.46
Observations	433	206	435	207
	Non-metropolitan first choice			
Graduate from first choice				
Living in non-metropolitan area	0.163	0.701	0.0224	-0.245
	(0.359)	(0.480)	(0.306)	(0.330)
F-stat	7.32	5.32	7.84	6.12
Observations	300	150	303	151
Living in metropolitan area	0.0829	0.271	0.0561	-0.0869
	(0.422)	(0.410)	(0.232)	(0.132)
F-stat	1.42	5.12	1.42	2.76
Observations	67	34	67	35

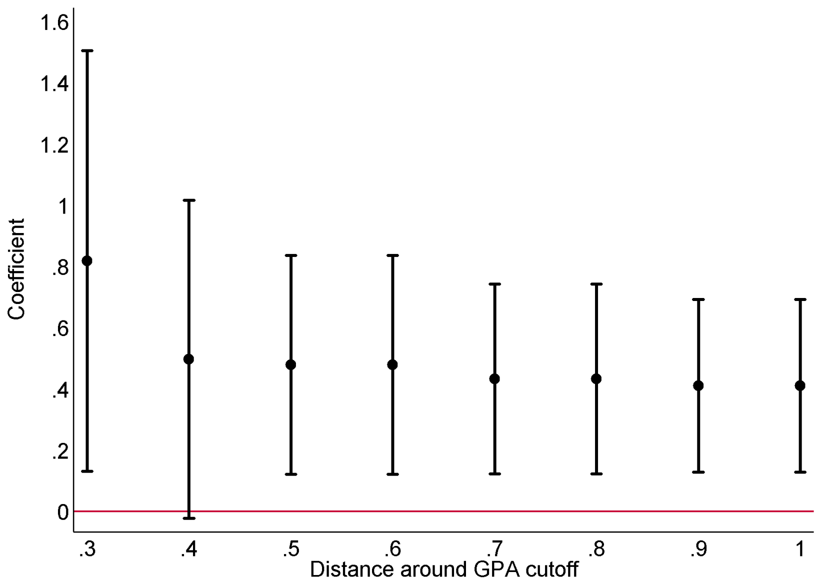
Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, field of study, location of first choice education, year of application, and municipality of residence in YOA. * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

Table 20: Heterogeneous effect of first choice location. Parental education

	Metropolitan first choice			
	live in non-metro	work in non-metro	live in uni region	work in uni region
Graduate from first choice				
Parents have higher educ	-0.399***	-0.258	0.698***	0.192
	(0.150)	(0.199)	(0.182)	(0.188)
F-stat	35.83	28.29	35.83	28.07
Observations	1,367	695	1,367	698
Parents do not have higher educ	-0.447***	-0.476**	0.579***	0.520***
	(0.181)	(0.193)	(0.117)	(0.175)
F-stat	38.17	32.76	28.17	34.43
Observations	616	333	616	332
	Non-metropolitan first choice			
Parents have higher educ	0.430	0.740*	0.0620	-0.426
	(0.349)	(0.406)	(0.283)	(0.279)
F-stat	8.82	6.28	8.82	5.87
Observations	220	106	220	107
Parents do not have higher educ	-0.381	1.052*	-0.00918	0.467
	(0.350)	(0.622)	(0.284)	(0.572)
F-stat	7.71	3.45	7.71	3.65
Observations	137	72	137	73

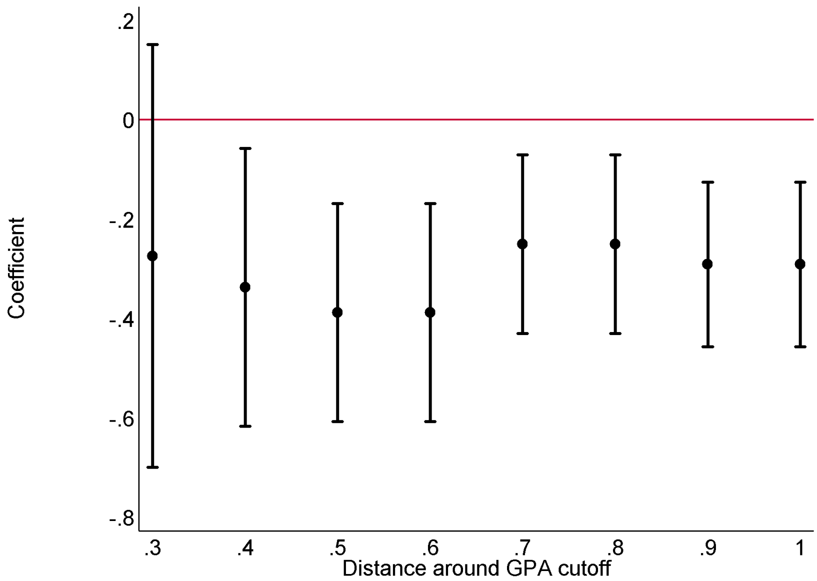
Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, field of study, location of first choice education, year of application, and municipality of residence in YOA. * p< 0.1, ** p< 0.05 and *** p< 0.01.

A.3 Bandwidth analysis



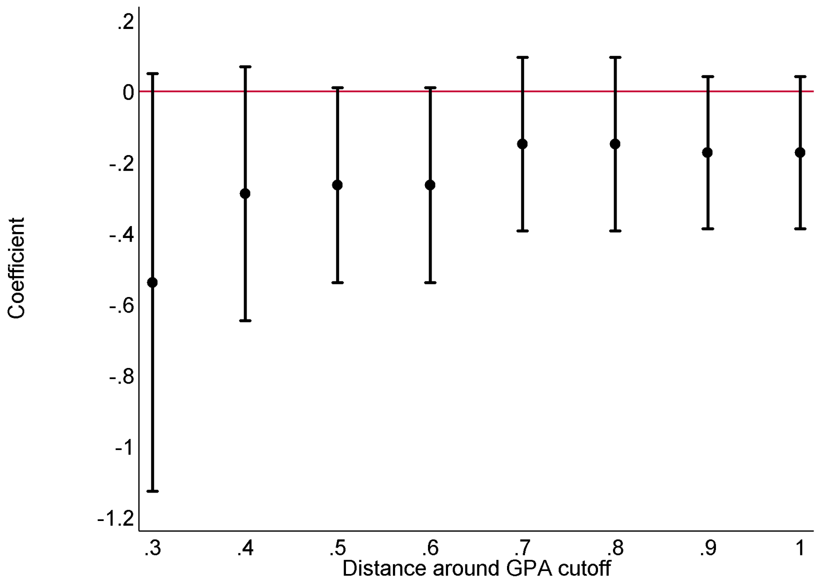
Notes: The figure shows the regression estimates and the 95-percent confidence interval using the bandwidths around the GPA admission cutoff shown on the horizontal axis as sample restriction. The outcome variable is working in a non-metropolitan region eight years after YOA.

Figure 11: Bandwidth analysis: The effect of completing university in second choice non-metropolitan location on the probability of working in non-metropolitan region



Notes: The figure shows the regression estimates and the 95-percent confidence interval using the bandwidths around the GPA admission cutoff shown on the horizontal axis as sample restriction. The outcome variable is living in a non-metropolitan region eight years after YOA.

Figure 12: Bandwidth analysis: The effect of completing university in first choice non-metropolitan location on the probability of living in non-metropolitan region



Notes: The figure shows the regression estimates and the 95-percent confidence interval using the bandwidths around the GPA admission cutoff shown on the horizontal axis as sample restriction. The outcome variable is working in a non-metropolitan region eight years after YOA.

Figure 13: Bandwidth analysis: The effect of completing university in first choice non-metropolitan location on the probability of working in non-metropolitan region

A.4 Additional analyses following Kirkeboen et al. (2016)

Table 21: The effect of university location on working in a non-metropolitan region eight years after applying for university. Following the methods by Kirkeboen et al. (2016)

First choice Second choice	Non-metro Non-metro	Non-metro Metropolitan	Metropolitan Non-metro	Metropolitan Metropolitan
Complete first	-	-	-0.179*	-0.0115
choice	-	-	(0.104)	(0.154)
Complete second	-	-	0.183*	0.00131
choice	-	-	(0.108)	(0.176)
Observations	-	-	778	789

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, location of alternative choice education, municipality of residence in YOA, field of study and YOA fixed effects. * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

Table 22: The effect of university location on living in the same region eight years after applying for university. Following the methods by Kirkeboen et al. (2016)

First choice Second choice	Non-metro Non-metro	Non-metro Metropolitan	Metropolitan Non-metro	Metropolitan Metropolitan
Complete first	-	-	0.379***	0.424***
choice	-	-	(0,0818)	(0,108)
Complete second	-	-	0.375***	0.484***
choice	-	-	(0.0544)	(0.108)
Observations	-	-	1,615	1,620

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, location of alternative choice education, municipality of residence in YOA, field of study and YOA fixed effects. * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

Table 23: The effect of university location on working in the same region eight years after applying for university. Following the methods by Kirkeboen et al. (2016)

First choice Second choice	Non-metro Non-metro	Non-metro Metropolitan	Metropolitan Non-metro	Metropolitan Metropolitan
Complete first	-	-	0.286***	-0.0525
choice	-	-	(0.0983)	(0.144)
Complete second	-	-	0.486***	0.401***
choice	-	-	(0.0893)	(0.146)
Observations	-	-	778	787

Notes: Robust standard errors in parentheses. All models include controls for age, gender, ethnicity, mother and father's education, location of alternative choice education, municipality of residence in YOA, field of study and YOA fixed effects. * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

Chapter 3

Grades and Employer Learning

Grades and Employer Learning*

Anne Toft Hansen[1,2], Ulrik Hvidman[3], and Hans Henrik Sievertsen[1,4]

[1]The Danish Center for Social Science Research - VIVE

[2]University of Copenhagen

[3]Aarhus University

[4]University of Bristol

Abstract

We identify the signaling value of a grade point average (GPA) on labor market outcomes and how fast employers learn about variation in GPA that is unrelated to labor market productivity. Exploiting a reform-induced variation in university graduates' GPA, we find that a higher GPA causes higher earnings in the first two years after graduation, after which the signaling effect goes to zero. We provide suggestive evidence that the signaling effect is more relevant to majors related to larger wage dispersion and the private sector and that the earnings adjustment happens both within and across firms. We conduct a range of sensitivity tests to validate that the identified relationship represents a signaling effect. Importantly, we show that the recoding scheme that causes variation in the GPA is unrelated to labor market outcomes for non-treated cohorts. Our findings are consistent with a learning model whereby employers initially screen graduates based on observable signals but update their beliefs about worker productivity.

*We thank conference participants at the 2017 ESPE meeting in Glasgow, at the 2018 RES annual meeting in Sussex, at the 2018 CEN meeting at the Copenhagen Business School, at the 2018 IWAEE in Catanzaro, at the 2018 IZA workshop on the Economics of Education in Bonn, at the 2018 DGPE meeting in Sønderborg, and at the 2019 SOLE annual meeting, as well as seminar participants in Aarhus, Bristol, Copenhagen and Trondheim for valuable comments. Sievertsen acknowledges financial support from the Danish Council for Independent Research through grant DFF: 4182-00200.

Hansen: ath@vive.dk
Hvidman: uhvidman@ps.au.dk
Sievertsen: h.h.sievertsen@bristol.ac.uk.

1 Introduction

How workers are allocated across jobs has implications for inequality and efficiency at an aggregate level (Roy, 1951; Sattinger, 1975) and involves large potential costs for employers and employees at the microlevel (Fredriksson et al., 2018). A major challenge in the matching process is that labor market productivity is imperfectly observed. According to job-market signaling theory, employers use completed schooling as a signal of labor market productivity to screen workers (Spence, 1973). However, as degrees have become increasingly common, these credentials constitute very crude signals and mask valuable information about the applicants' ability.[1] Consequently, employers are often faced with a choice between applicants with similar levels of educational attainment (e.g., a university degree) and may therefore look for other signals of productivity, such as information on educational achievement. Signals of educational achievement in terms of grade point averages (GPAs) are common in many countries. For example, US colleges typically use letter grades (A through F), which are then converted to a numerical GPA, whereas UK universities assign scores on a 100-point scale, which are then translated into a degree classification (e.g., first-class honors).[2]

In this paper, we examine the signaling value of a GPA on labor market outcomes and how fast employers learn about workers' true productivity. We, thereby, revisit a central question in labor economics: Is a causal effect of schooling on earnings best explained by

[1]Between 2000 and 2016, the proportion of the students aged 25 to 34 who had attained a tertiary education increased on average across the OECD countries from 26 percent to 43 percent (OECD 2017). For example, in 2016, more than 50 percent of the students aged 25 to 34 completed tertiary education in the UK.

[2]The role of grades in the labor market constitutes a long-standing debate in media coverage and academia. In 2006, the New York Times ran a story on the impact of grades under the headline "Those Low Grades in College May Haunt You." In 2013, Forbes ran the headline "Do Employers Really Care about your College Grades?" with the following opening remark, "Yes, is the short answer!" Interview and survey data also provide evidence that supports the notion that employers and colleges in the United States are aware of measures of achievement when they hire. According to a survey conducted in 2016 by the National Association of Colleges and Employers, 70 percent of the surveyed companies report that they use GPA to screen potential job candidates (National Association of Colleges and Employers, 2016).

the human capital theory or the job-market signaling theory? According to the human capital theory (Schultz, 1961; Becker, 1962), a higher GPA is related to higher earnings because schooling increases productivity, and a higher GPA reflects a better mastering of the taught material. In contrast, according to the job-market signaling (Spence, 1973), a higher GPA is related to higher earnings because a higher GPA reflects higher innate ability.[3] [4]

It is challenging to empirically separate the human capital and job-market signaling theory because they lead to similar predictions (e.g., Huntington-Klein (2018): "Human Capital vs. Signaling is Empirically Unresolvable"). Even if one were able to randomly manipulate years of schooling and follow students' labor market trajectories, one would not be able to isolate signaling from human capital effects of schooling. Instead of manipulating education at the extensive margin (i.e., what degree you completed), recent experimental studies apply CV designs to manipulate signals at the intensive margin (i.e., how well you completed the degree in terms of your GPA) (Koedel and Tyhurst, 2012; Protsch and Solga, 2015; Piopiunik et al., 2018). Although these studies provide valuable insights into how better credentials lead to a higher probability of being invited for an interview, they only study the initial phase of the hiring process and not actual job market outcomes (i.e., hirings and wages).

We exploit a grading-reform that mimics the ideal experiment by causing as-good-as random variation in university graduates' GPA. A Danish grading reform from 2007 introduced a new grading scale across all educational programs. Students who were enrolled in university during the implementation had their existing grades recoded to the

[3]It is worth noting that in contrast to the distinction between the human capital theory and job-market signaling theory based on completed years of education, the GPA to some extent works as a signal in both cases. In the "human capital theory", the GPA signals a better understanding of the concepts taught in school (which increases productivity).

[4]While the association between years of schooling and earnings is established, the association between GPAs and earnings is less studied. In Appendix B, we document that a one standard deviation increase in university GPA is associated with a four to seven percent increase in earnings in the first five years after graduation.

new scale based on a scheme provided by the Ministry of Education. The recoding caused substantial variation in post-reform GPAs: Two individuals with identical pre-reform GPA could end up with more than a standard deviation difference in post-reform GPA. We use this reform-induced variation in grades to identify the effects of GPA on labor market outcomes.

Using a novel data set containing all students at the two largest universities in Denmark, corresponding to around half of the total population of university students in Denmark, we show that a one standard deviation increase in the reform-induced variation in the GPA led to 21 percent higher earnings in the first year and 19 percent higher earnings in year two after graduation. However, this effect declines with experience and there is no detectable effect three years after graduation, suggesting a rapid employer learning process. To assess the validity of the design, we provide a set of supplementary analyses. Importantly, we show that the variation caused by the recoding is not associated with individual characteristics that predict labor market outcomes (e.g., high school GPA, parental income, and parental schooling). Moreover, we conduct placebo tests that demonstrate that the recoding algorithm does not predict future labor market outcomes for non-treated cohorts. These supplementary analyses provide strong evidence that the relationship between reform-induced variation in GPA and labor market outcomes is not driven by the recoding benefiting specific grade combinations that in itself are rewarded on the labor market.

We find no effects of reform-induced variation in GPA on sector choice (private vs. public employer); however, our subgroup analysis suggests that the earnings effects are strongest for graduates from fields associated with high wage dispersion and the private sector, which provides suggestive evidence of the crucial moderating role of the wage-determination process in the relevant labor market. We hypothesize that the signaling effect should be largest for students without informal contacts in the relevant labor market.

Based on this hypothesis, we would expect that children of parents without a university degree are more reliant on the GPA as a signaling device, which is also confirmed by our subgroup analysis. Finally, we also find stronger signaling effects for men, which could be explained by a more compressed wage structure for women, or by gender differences in sector choice.

Looking into the wage adjustment process, we find no evidence of a link between reform- induced variation in GPA and overall job changes in the first five years after graduation. However, we find a slower earnings growth for individuals who experienced a positive reform-induced change in GPA in the second to the third year after graduation. Although the slower earnings growth is detectable for workers who stay in the same firm, supplementary analyses suggest that adjustment is fastest among workers who change firms. These findings suggest that the earnings adjustments occur both within and across firms.

These results have direct policy implications. First, grades are relevant in the labor market matching process for university graduates. If we give a student a different grade, all else equal (including exam performance), the student will have a different labor market outcome in the short run. This finding suggests that the grading system affects matching efficiency. Second, employer learning happens rapidly. An initially substantial earnings premium to variation in a signal of educational achievement that is not related to labor market productivity disappears within three years, and the adjustment occurs both within firms and across firms.

Although the previous research on the signaling value of education has focused on educational attainment, a few papers have studied the signaling value of educational credentials among students with similar length of education. Freier et al. (2015) use a difference-in-differences approach to estimate the effect of graduating with an honors degree. They exploit that only graduates from a law major receive an honors labeling for being in the top of the distribution. They compare the difference in the earnings across majors in the top (i.e., above the honors cutoff for law graduates) of the achievement distribution to the difference below the honors cutoff. They find a sizable earnings effect of graduating with honors, five to six years after graduation. Khoo and Ost (2018) also focus on the top of the achievement distribution and estimate the return to an honors degree in Ohio. To identify the causal effect, they exploit that an honors degree requires the final cumulative GPA to be above a certain threshold in a regression discontinuity setting. In line with our findings, they find labor market returns to be above the cutoff in the first two years after graduation. Another way to signal productivity is by university prestige. Bordón and Braga (2017) exploit university admission test score cutoffs in a regression discontinuity approach and find a substantial initial wage premium of graduating from a prestige university, which decreases over time. Our paper provides several contributions to this literature. First, we identify the signaling effect of university GPA across the entire achievement distribution and across a wide range of university majors. Second, we use an identification strategy that allows us to conduct a wide range of robustness checks and placebo tests. Third, we assess the earnings adjustment mechanism and heterogeneity in effects to understand where signaling is most relevant and how the earnings adjustment occurs.

Our findings contribute to the literature on signaling and employer learning (Arrow, 1973; Wise, 1975; Farber and Gibbons, 1996; Altonji and Pierret, 2001; Riley, 2001; Schönberg, 2007; Arcidiacono et al., 2010; Clark and Martorell, 2014; Jepsen et al., 2016; Di Pietro, 2017) and documents sorting and employer learning by a new dimension of educational credentials. Perhaps even more importantly, our study reinforces the existing evidence that employers learn rather quickly about actual productivity (Lange, 2007; Aryal et al., 2019).

The remainder of this paper is organized as follows. Section 2 presents the research design. Section 3 describes the data, Section 4 presents the results. Section 5 concludes this paper.

2 Research design

Before we describe our empirical strategy, it is useful to clarify how we define the signaling value of GPA. Our approach is that GPA has non-zero signaling value if the GPA ***per se*** has a causal effect on earnings. This definition implies that the signaling value of GPA is the effect of the GPA on earnings holding everything else constant, including how well the students performed in university. An ideal experiment would randomly allocate shocks to the GPA on diplomas for university graduates and follow their labor market trajectory. We exploit a grading reform that created a setting that very closely resembles this ideal experiment.

2.1 The 2007 Danish grading reform

Until 1 August 2007, school achievement in the Danish school system from lower secondary schooling to post-secondary schooling was evaluated on a scale from 0 to 13 (called the "13 scale").[5] In early 2004, the Danish government launched an initiative to modernize tests and grades. As part of this initiative, a Commission for Examining the Danish Grading Scale was established. In November 2004, the commission recommended the introduction of a new 7-point grading scale from -3 to 12 (called the "7-point scale"). In 2006, the government decided to introduce the new scale, and on 1 August 2007, the 7-point grading scale replaced the 13 scale grading system in all educational programs in Denmark.

How the grading reform affected an individual depended on where the individual was in the educational system on August 1. If the individual had completed a degree,

[5] See also Hvidman and Sievertsen (2018) for a description of the grading reform.

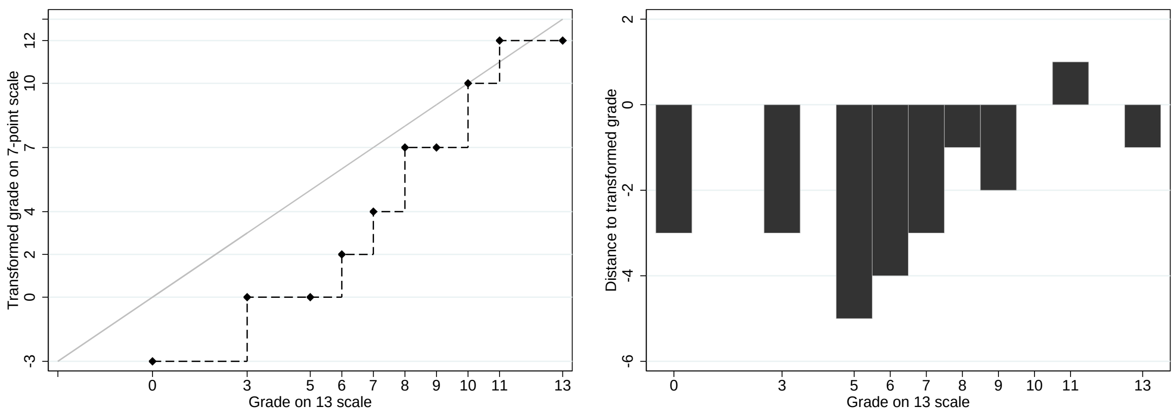
the entire GPA would be transformed into a new GPA according to a mapping scheme provided by the Ministry of Education. However, for students enrolled in an educational program on August 1, all grades given on the old scale were transformed to the new scale based on the scheme shown in Table 1, which gave rise to the variation in GPAs.

Table 1: The Danish Grading System: Transformation from the Old to New Scale

Old 13-scale	New 7-point scale	ECTS	Description
00	-3	F	For a performance which is unacceptable in all respects.
03 5	0	F^+	For a performance which does not meet the minimum requirements for acceptance.
6	2	E	For a performance meeting only the minimum requirements for acceptance.
7	4	D	For a fair performance displaying some command of the relevant material but also some major weaknesses.
8 9	7	C	For a good performance displaying good command of the relevant material but also some weaknesses.
10	10	B	For a very good performance displaying a high level of command of most aspects of the relevant material, with only minor weaknesses.
11 13	12	A	For an excellent performance displaying a high level of command of all aspects of the relevant material, with no or only a few minor weaknesses.

Source: The Danish Ministry of Science, Innovation and Higher Education.
Notes: ECTS is the grading system defined by the European Commission. 6 (old) / 2 (new) is the passing threshold.

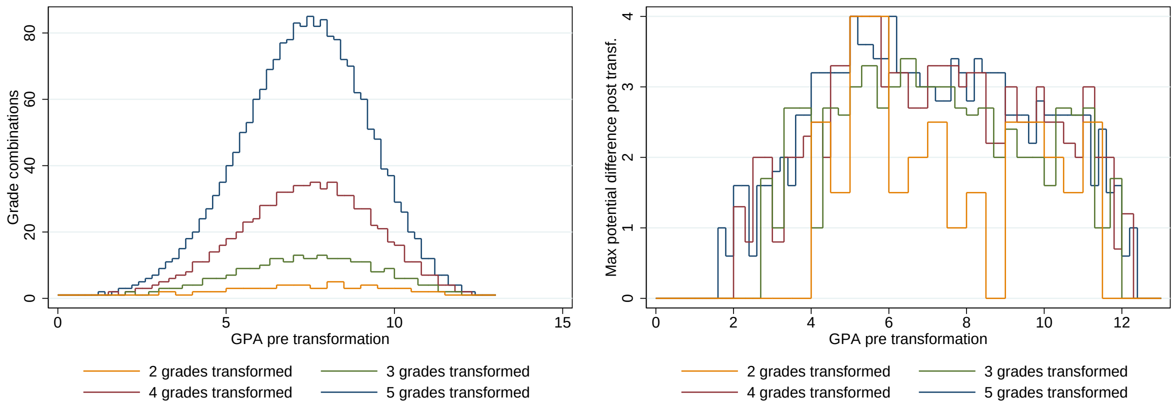
The first two columns of Table 1 present the transformation scheme from the old 13 scale to the new 7-point scale. There are two important sources of variation in the post-transformation GPA. First, as the new grading scale has fewer grades (seven compared to ten), three pairs of grades on the old scheme were collapsed to single new grades. For example, a student who only had 8s on the old scheme and another student who only had 9s on the old scale would have identical GPAs after the transformation. Second, the distance between the old and the new grade varies. Figure 1a shows the grade on the new scale against the grade on the old scale. The vertical distance between the dots and the gray 45-degree line shows the transformation penalty. Figure 1b plots these vertical distances. Some grades are closer to the 45-degree line than others, and one (grade 11) is even above it. Although most students were downgraded in absolute terms due to the differences in the two scales, two students with identical pre-transformation GPAs could have very different post-transformation GPAs because grades were punished differently. Consider, for example, a student with a 5 and a 6 on the old grading scale. These grades were punished heavily in the transformation process, and the student would go from a GPA of 5.5 to a GPA of 1. Another student with a 3 and an 8 on the old scale would go from a GPA of 5.5 to a GPA of 3.5. As a result, depending on the composition of grades, the students were either down- or upgraded relative to their peers.



(a) Mapping between grading scales **(b)** Grade distance

Figure 1: Mapping from 13 scale (old grading scale) to 7-point scale (new grading scale).

While more grades imply that more grade combinations can cause a specific pre-transformation GPA, as shown in figure 2a, the link between the number of grades given on the old scheme and the potential variation in post-transformation GPA is not trivial (see Figure 2b). For some pre-transformation GPAs, a lower number of grades given is associated with a greater potential post-transformation difference in GPA.



(a) Grade combinations **(b)** Maximum difference in post-transformation GPA

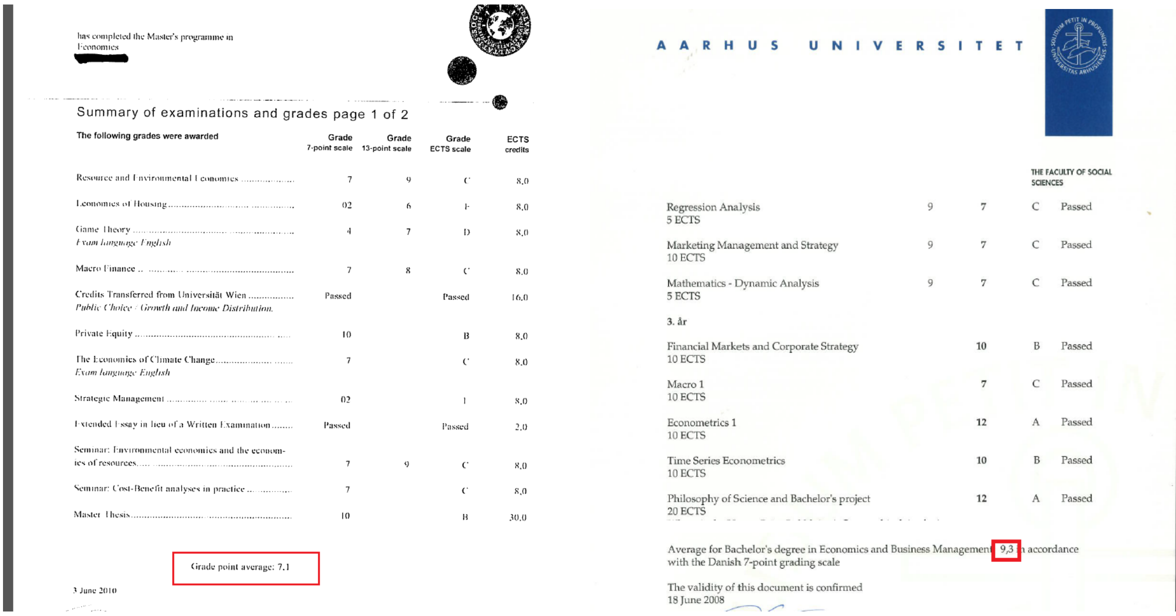
Figure 2: Combinations and maximum difference, given GPA and number of transformed grades.

2.2 The implementation of the grading reform in the Danish higher education system

After completing upper secondary education, students can apply for university programs in Denmark. All programs are free, and all students above the age of 18 receive a monthly grant to pay for their living costs. Enrollment in university programs depends almost exclusively on high school GPA.

Although Denmark has adopted a three-year bachelor and a two-year master structure for its university programs, most programs are still five-year programs in practice (some programs such as medical school are six-year programs), and more than 90 percent of the bachelor graduates progress to a master program within two years. As we focus on the importance of GPAs for sorting on the labor market, we focus on graduates from master programs. University modules are given an European Credit Transfer System (ECTS) weight according to their overall workload, and students are expected to complete 60 ECTS points in one year. A year is typically split into terms of 14–15 weeks (some programs have four terms of eight weeks), and most programs are finalized by a dissertation.

Students who were affected by the reform also completed exams after the recoding. The final GPA is, thus, a weighted average of the recoded GPA and the GPA for exams after the recoding. This is done in practice like the two example diplomas shown in Figure 3. For both universities, two columns of exam results are given. The first (second) column at the University of Copenhagen (Aarhus University) presents the grade on the new scale, and the second (first) column shows the grade on the old scale. The column with grades on the old scale is blank for grades given after the recoding. As explained above, it is not possible to reconvert the new grades to the old scale, as the new scale has fewer grades.


has completed the Master's programme in
Economics
[redacted]

Summary of examinations and grades page 1 of 2

The following grades were awarded	Grade 7-point scale	Grade 13-point scale	Grade ECTS scale	ECTS credits
Resource and Environmental Economics	7	9	C	8,0
Economics of Housing	02	6	E	8,0
Game Theory *Exam language English*	4	7	D	8,0
Macro Finance	7	8	C	8,0
Credits Transferred from Universität Wien *Public Choice / Growth and Income Distribution.*	Passed		Passed	16,0
Private Equity	10		B	8,0
The Economics of Climate Change *Exam language English*	7		C	8,0
Strategic Management	02		E	8,0
Extended Essay in lieu of a Written Examination	Passed		Passed	2,0
Seminar: Environmental economics and the economies of resources	7	9	C	8,0
Seminar: Cost-Benefit analyses in practice	7		C	8,0
Master Thesis	10		B	30,0

Grade point average: 7,1

3 June 2010

AARHUS UNIVERSITET

THE FACULTY OF SOCIAL SCIENCES

Regression Analysis 5 ECTS	9	7	C	Passed
Marketing Management and Strategy 10 ECTS	9	7	C	Passed
Mathematics - Dynamic Analysis 5 ECTS	9	7	C	Passed
3. år				
Financial Markets and Corporate Strategy 10 ECTS		10	B	Passed
Macro 1 10 ECTS		7	C	Passed
Econometrics 1 10 ECTS		12	A	Passed
Time Series Econometrics 10 ECTS		10	B	Passed
Philosophy of Science and Bachelor's project 20 ECTS		12	A	Passed

Average for Bachelor's degree in Economics and Business Management 9,3 in accordance with the Danish 7-point grading scale

The validity of this document is confirmed
18 June 2008

(a) University of Copenhagen

(b) Aarhus University

Figure 3: Examples of diplomas for treated individuals

2.3 How the quasi-experimental setting deviates from the ideal experiment

Before we describe how we conceptually exploit the variation in GPAs caused by the grading reform, we first briefly discuss how this setting deviates from the ideal experiment.

The reform-induced variation in GPA. The transformation of the individual pre-reform grades are deterministic; however, the variation in post-recoding GPAs depends on the combination of pre-recoding grades. It is not trivial that the reform-induced variation in GPA is completely random. However, as long as the recoding algorithm generates variation in GPAs that is unrelated to unobservables relevant to labor market outcomes, we can still use the variation to infer the causal effect of GPAs on labor market outcomes. To assess this, we conduct three types of tests. First, we assess whether the reform-induced variation in GPA is related to observables that predict labor market outcomes, such as high school GPA, parental income, and parental schooling. Second, we conduct a placebo test and assess whether the recoding algorithm would predict labor market outcomes in a non-treated cohort. Third, we assess whether the reform-induced variation in GPA is related to different student behavior.

Students could anticipate the grading-reform reform. As described above, the grading reform did not come as a surprise. Students could anticipate the reform and change their behavior accordingly. To get a sense of how anticipated the change was, Figure A.1 in Appendix A shows the Google search trend for the term ”the new grading scale” (in Danish: ”den nye karakterskala”). The search activity before the implementation was modest compared to the post-implementation period. However, we cannot rule out that students anticipated the reform. Instead, we consider how such anticipation would affect our setting.

First, if students thought they benefited from the recoding, they could delay graduation to ensure that they graduated after the implementation of the reform. If they thought otherwise, however, they could accelerate their studies and ensure that they graduated before the recoding. Assuming that this behavior was at play, we would expect the students who received a positive reform-induced change in their GPA to have behaved differently compared to those who received a negative shock (e.g., more forward-looking ones). We would, therefore, expect the reform-induced variation to be related to observables, such as high school GPA, which we test for.

Second, students could avoid grades that were penalized relatively more by the recoding. This would require that the students had perfect control over the grade or were more likely to complain about a given grade if it was penalized by the recoding. Figure A.2 in Appendix A shows the relative frequency of each given grade in the years up to the reform. On this aggregate level, we see no evidence of grades such as 6, 7, and 9, which were penalized relatively more, became less frequent in the period up to the recoding.

The original grades are observable to employers. The diplomas in Figure 3 show that employers can observe all individual grades if they take a more detailed look at the diplomas. What if employers use all this information instead of using the post-reform calculated GPA? If this is the case, we would not expect any return to the reform-induced variation on earnings. If employers use both the GPA and individual grades in the process of worker sorting, we would still expect an effect of the reform-induced variation on earnings.

The grading scale changed. The major grading reform required everyone to learn about the new grading scale. The employers had to get a sense of what a grade 7 on the new scale was compared to a grade 7 on the old scale. If employers did not understand the new grading scale, it is possible that they would use the individual grades on the diplomas reported both in terms of the new and the old scale. If employers use the individual pre-reform grades instead of the post-reform GPA, we would expect no effect of the reform-induced variation on earnings.

2.4 Empirical strategy

Our empirical strategy consists of two steps. First, we construct a variable capturing the variation in GPAs caused by the recoding of grades. Second, we use the variable created in the first step as an instrument in an instrumental variable (IV) approach.

2.4.1 Conceptual framework

To estimate the causal effect of university graduates' GPA on earnings, we consider the following linear relationship between GPA and earnings:

$$y_i = \alpha_0 + \alpha_1 GPA_i + \epsilon_i \tag{1}$$

where y_i is earnings for individual i with university GPA, GPA_i, and ϵ includes all other factors affecting the earnings, which could be both other signals of labor market productivity or factors directly related to labor market productivity (e.g., cognitive and non-cognitive skills).

An ordinary least squares regression (OLS) of earnings on GPA will not allow us to infer if a positive α_1 is due to GPA being used as a signaling device because we cannot identify the signaling effect from the overall GPA because (1) the GPA is correlated with other characteristics that are used as a signaling device or (2) employers directly observe and reward productivity, which is correlated with the GPA. To separately identify the signaling effect from (1) and (2), we require variation in GPA that is unrelated to other signals, which rules out explanation (1), and unrelated to observable and unobservable characteristics related to labor market productivity, which rules out explanation (2). To extract variation in GPA that satisfies this condition, we use the grading reform described above in an IV setting. We use variation in the graduates' GPA, generated by the recoding of grades as an instrument for the final GPA that appears on the diploma. We describe the link between the recoded GPA and the final GPA in the next section, which conceptualizes the first stage.

2.4.2 Constructing the instrument

As seen from the examples shown in Figure 3, the graduation diploma contains an overall GPA, which is the weighted average of the recoded grades, $GPA7_{i,pre}$, and the average for grades given after recoding, $GPA7_{i,post}$. The recoded GPA is a function of the GPA on the old scale, $GPA13$, and the reform-induced variation in GPA, e, generated by the recoding algorithm:

$$\begin{aligned} GPA7_i =& (1-\omega_i)GPA7_{i,post} + \omega_i GPA7_{i,pre} \\ =& (1-\omega_i)GPA7_{i,post} + \omega_i f(GPA13_i + e_i) \end{aligned} \tag{2}$$

where ω is the weight of the pre-recoding grades; in other words, the share of a student's total number of grades that are recoded from the old to the new grading scale.

If e_i is a valid instrument that we can identify empirically (i.e., it is additively separable from $GPA13_i$), the coefficient on e_i in a first stage regression should have the same size as the average of ω_i. This is because if e_i is a valid instrument, the impact on the final GPA is mechanically identical to the share of grades that are affected by the recoding. If the exclusion restriction is violated the first stage coefficient deviates from the average share of recoded grades. For example, if the students who receive a negative reform-induced

GPA change attempt to compensate for this shock by increasing their effort in subsequent assessments, the first stage coefficient would be smaller than the share of grades affected by the recoding because the effect is mitigated by the behavioral response of the students.

The only remaining empirical question is how to define $f()$ and thereby measure e_i. In our main specification, we use a third-order polynomial as an approximation of $f()$ and estimate the following equation using OLS:

$$GPA7_i = \eta_0 + \eta_1 GPA13_i + \eta_2 GPA13_i^2 + \eta_3 GPA13_i^3 + e_i \tag{3}$$

Through this approach, the fitted value corresponds to the best guess of the post-recoding GPA, and the reform-induced variation in GPA is the residual, $\hat{e}$.

As a sensitivity check, we estimate specifications with a second- and fourth–order polynomial as well as less parametric approaches where we non-parametrically approximate $f()$. In the latter approach, we construct e_i as the difference between the actual post-recoding GPA and the average or median post-recoding GPA for individuals with the same pre-reform GPA.[6]

2.4.3 The IV specification

Having identified the reform-induced variation in GPA, e, we estimate the following reduced form equation using OLS:

$$y_i = \tau_0 + \tau_1 \hat{e}_i + \epsilon_i \tag{4}$$

[6]In Appendix C we provide a Monte Carlo simulation of our empirical setting. Figure C.2 shows that with a linear specification, we fail to reject a true null-hypothesis of no relationship between GPA and earnings in nearly 100 percent of the cases with a linear specification, using a 5 percent cutoff. This is mainly because the linear approximation works poorly in the upper and lower end of the GPA distribution. Both the second and third polynomial specifications lead to rejection rates of the expected 5 percent. Interestingly, both the fourth-order polynomial and the non-parametric approaches perform slightly worse than the second- and third-order polynomials, which is our motivation for using the third-order polynomial as the main specification.

The coefficient τ_1 captures the reduced form relationship between the estimated reform-induced variation, $\hat{e}$, and the earnings, y.

To get a sense of the effect size, we need to scale the reduced form coefficient, τ_1. It is straightforward to scale this coefficient by the first stage coefficient using an IV approach where equation (2) corresponds to the first stage, and the expected first stage coefficient is equal to ω. In other words, the IV estimate of the causal effect of GPA on earnings is the reform-induced variation in GPA on earnings, scaled by the relative weight of the share of grades in the GPA, which was recoded due to the reform.

As we empirically measure e_i (i.e., the instrument), we need to take the additional uncertainty into account when estimating the standard errors. We do this by bootstrapping the full estimation procedure with all three estimation steps. To allow for arbitrary correlation within pre-recoding GPA cells, we cluster bootstrap on the GPA level.

Our identification strategy does not rely on any covariates. However, to reduce the residual variance in the outcome variables and obtain more precise estimates, we include a range of controls. First, we expect earnings to be related to cohort effects, the program studied, and the institution. We, therefore, include indicators for the program studied, the enrollment year, an indicator for institution (Copenhagen or Aarhus University), and interactions between enrollment year and institution. Second, earnings might also be related to individual characteristics and background. We, therefore, control for age at enrollment, an indicator for origin (non-Western or Western according to the definition by Statistics Denmark), parental income, parental unemployment, parental education (an indicator for university degree), and gender. Parental variables are created as the mean across observed parents (except schooling, which equals 1 if at least one parent has completed a university degree). We include indicators for the number of parents with in order of non-missing income, unemployment and education (i.e., 0, 1, or 2).

2.4.4 Identifying assumptions

Identifying assumption 1: Relevance. The reform-induced variation in the GPA only informs about the causal effect of GPA on earnings. The reform-induced variation in GPA should affect the overall GPA: $cov(GPA7_{i,pre}, e_i) \neq 0$. It is straightforward to test this assumption by regressing $GPA7$ on e, which, according to equation (2), should lead to a coefficient on e equivalent to ω (as long as e enters additively), which is simply the share of grades that is recoded by the reform.

Identifying assumption 2: Exogeneity. We require the reform-induced variation in GPA to be unrelated to individual characteristics that affect earnings: $cov(\epsilon_i, e_i) = 0$ (from equations 1). It could be the case that the combination of grades that leads to a large e is also related to higher earnings directly. To assess this assumption, we conduct several sensitivity checks. First, we show that the reform-induced variation in GPA is unrelated to observable individual characteristics, such as parental background, high school GPA, and undergrad GPA. Second, we conduct a placebo test whereby we implement the recoding scheme on a non-treated cohort. With this placebo test, we essentially test whether there is a link between the combination of grades leading to positive reform-induced variation in GPA and also being directly related to earnings without the recoding.

Identifying assumption 3: Exclusion restriction. We require that there are no other responses to the recoding of GPAs that could also affect earnings. In related work using the same reform, we show that high school students react to negative reform-induced variation in their GPA by increasing effort (Hvidman and Sievertsen, 2018). If the same scenario holds in this setting, the grades given after the recoding, which enters the overall GPA, might be affected through this behavioral response to the reform-induced variation in GPA. Moreover, it could be that the reform-induced variation affects the students' likelihood of graduation, time to graduation, and also course selection after the recoding. Therefore, the correlation between the reform-induced variation in GPA and earnings might be driven by these post-reform responses, and not because GPA acts as a sorting device. While we explicitly test for all these post-reform responses, it is worth noting that we do not expect a priori the same response as for the high school students for two reasons: First, for high school students, the GPA is high stakes as it determines access to higher education. Second, high school students have considerably more time to react on the grading reform, as we restrict our sample to students who are very close to graduation when the reform is implemented.

2.4.5 Measurement error interpretation.

Equation (2) along with the stated assumptions about exogeneity suggest that we apply a classical measurement interpretation where $GPA7_{i,pre}$ is the GPA measured with noise, $GPA13_i$ is the true GPA, and e_i is a classical measurement error that is assumed to be independent of $GPA13_i$. Using this interpretation we can think of the reform-induced variation in GPA as a measurement error and that employers over time will learn about this measurement error.

However, it is worth noting that we are not assuming that $GPA13$ is the true GPA measure without noise; we are only exploiting that the recoding from the old grading scale to the new grading scale creates variation in the GPA. Therefore, we do not treat the reform-induced variation as a measurement error but just as exogenous variation in the GPA.

3 Data

3.1 Sample selection

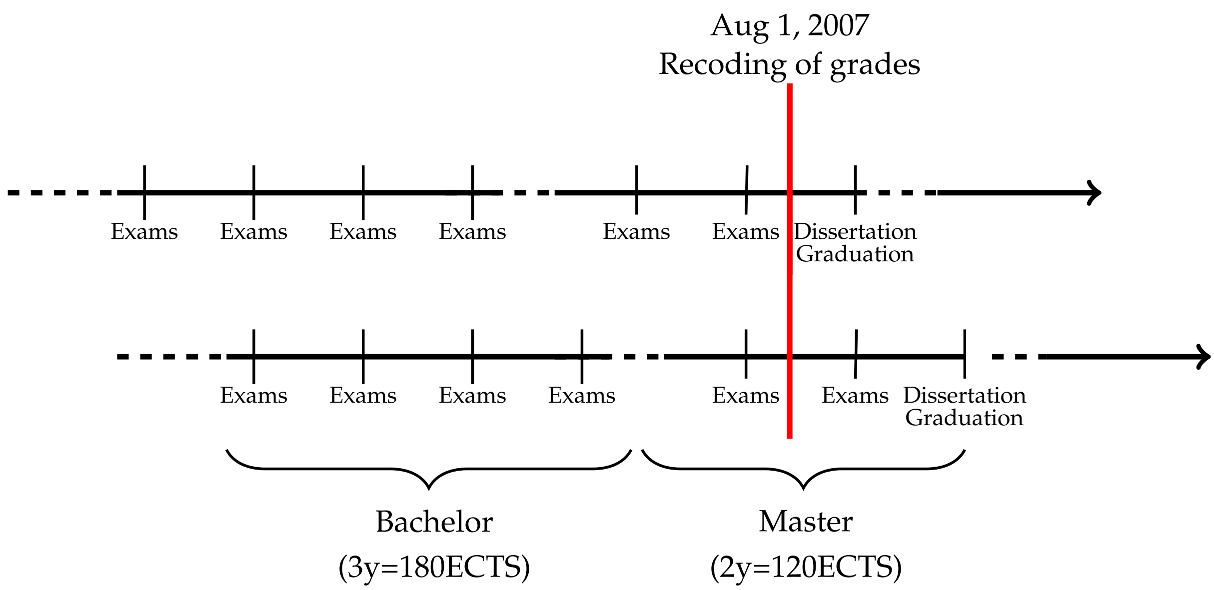


Figure 4: University students' exposure to the implementation of the new grading scheme

As a point of departure, we consider all students who were enrolled in a master program on 1 August 2007 at Aarhus University or University of Copenhagen. Students will be at different points of the programs at this point, as illustrated in Figure 4. As the treatment (i.e., the reform-induced variation in GPA) is caused by the recoding of given grades on 1 August 2007, we further narrow our sample to those who are at the end of their studies at this point (i.e., the upper row of Figure 4). Specifically, we restrict the sample to the students who had at most 40 ECTS points remaining of their program on the day the grades were recoded. Ideally, we would like to select students who were only waiting for their dissertation results. However, as university studies are very customizable in Denmark (which means that students might have to complete some modules after their dissertation) and because the credit load of the dissertation varies across years and programs, we cannot strictly impose such a criteria. The 40 ECTS criteria is selected based on the fact that ECTS assigned to the dissertation varies between 30 and 60 in our sample. We show in Figure 7 in section 4.3 that our results are not sensitive to this sample selection. No further sample restrictions are applied, and the final sample consists of 4,576 students.

We subsequently merge the student records with administrative registers from Statistics Denmark using a unique personal identifier. The registers from Statistics Denmark provide individual background information, including age, gender, high school GPA, parental income and education as well as information about labor market outcomes after university graduation.

3.2 Variables

From the student registers provided by Aarhus University and University of Copenhagen, we construct the overall GPA, the GPA before the recoding, and the GPA after the recoding. We further record the number of credit points, the type of units completed, the program studied, and the date of graduation.

From Statistics Denmark, we first create a variable on the students' age at the time of the reform, their high school GPA, parental employment, the total parental disposable income in the calendar year prior to the reform, and an indicator for at least one parent having a university degree. In cases where we do not observe a covariate, we set the covariate to zero and include a dummy variable that equals 1 for all observations that have missing values on this variable.

Our main labor market outcome is log total gross earnings in the first five calendar years after graduation. Individuals with zero earnings are excluded from earnings regressions; however we separately estimate a specification on whether the reform-induced variation in GPA is related to having positive earnings (i.e., the extensive margin). We further construct indicator variables for whether the individual had a job change in each following year after graduation, whether the individual has positive earnings, and whether the individual works in the public sector. Additionally, we also construct measures for accumulated earnings and earnings growth in the first five calendar years after graduation.

3.3 Descriptive statistics

Table 2 shows summary statistics for selected variables. In our sample, 65 percent of the students are female, and 4 percent are of non-Western origin. The average age of the master students is almost 27 years. Forty-one percent of the sample are students from University of Copenhagen, whereas the remaining 59 percent are students from Aarhus University. We observe that on average students were fairly close to graduation (23 ECTS remaining). On average, the grades given before the recoding accounted for about 70 percent of the overall GPA, suggesting that the reduced form estimates will be scaled by 0.7.

Ninety percent of our sample have positive earnings in the year after graduation. A university graduate in our sample earns on average 42,880 euros in the first calendar year after graduation (corresponding to around USD48,000), with 66 percent of them are working in the public sector. The unemployment rate among the graduates is 8 percent.

4 Results

4.1 Reform-induced variation in GPA

Figure 5 shows the actual variation in GPA caused by the recoding for the graduates in our sample. Each cross represents several observations. Cells with less than five observations are not shown in the graph but are included in the regressions. We observe substantial variation in post-recoding GPA over the entire distribution of passing grades (i.e. grades

Table 2: Summary statistics

	Mean	SD	P10	P50	P90
Female	0.65	0.48			
Non-western origin	0.04	0.19			
Age at entry	26.66	5.60	22.62	25.17	32.04
Parents' with university degree	0.27	0.44			
University of Copenhagen	0.41	0.49			
ECTS remaining	23.21	13.39	0.00	30.00	40.00
Pre reform grade share	0.70	0.12	0.50	0.71	0.82
Earnings>0 in year 1	0.90	0.30			
Earnings in year 1	42.88	21.25	10.04	45.27	66.61
Public sector in year 1	0.66	0.47			
Observations	3810				

Notes: P10, P50 and P90 refer to respectively the 10th pseudo-percentile, the 50th pseudo-percentile and the 90th pseudo-percentile. Pseudo-percentiles are created by the average across the actual percentile and the two values above and below the percentile. Age at entry refers to age at entry in the graduate program. Parental education is measured in the calendar year before the focal individual graduates from university.

above 6 on the old scale).

4.2 Returns to the reform-induced variation in GPA

Table 3 presents our main regression results of GPA on earnings in the first five years after graduation. Panel A of Table 3 shows the relationship between overall GPA and log earnings in the first five calendar years after graduation. The relationship is based on regressing log earnings on overall GPA using OLS. The coefficient suggests that a one-unit higher GPA is related to between 2 and 3 percent higher earnings across all five years. Scaling these coefficients by the standard deviation of the GPA implies that a one standard deviation higher GPA is related to between 3 and 4 percent higher earnings in the first five years after graduation.

Panel B of Table 3 shows the reduced form relationship between the reform-induced

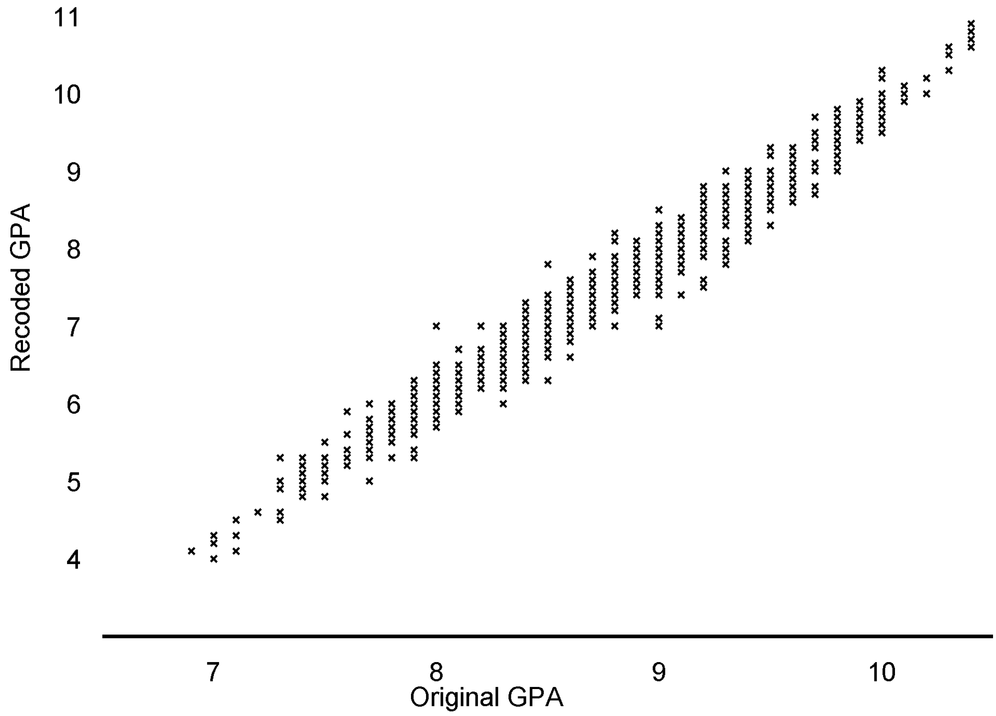


Figure 5: Pre- and post-recoding GPA for our sample.
Notes: Each cross represents a combination of pre and post recoding GPA. Only grade combinations with at least five observations are shown.

variation in GPA and log earnings by estimating equation (4) using OLS. The results show a positive and precisely estimated relationship in the first two years after which the coefficient goes to zero and becomes insignificant. This is consistent with the hypothesis of employer learning, whereby employers learn about workers' true productivity over time and, therefore, do not need to use a signal of productivity to form wages as explained in section 2.

To get a sense of the effect size of the causal effect of GPA on log earnings, panel C shows the estimates of the first stage relationship. The estimates predict overall GPA based on the reform-induced variation in GPA. Recall that if the exclusion restriction is valid, the first stage coefficients should be equal to average weight of pre-recoding GPA

Table 3: GPA and earnings regression results. Dependent variable: log earnings year one to five after graduation.

	Year after graduation				
	1	2	3	4	5
A. OLS specification					
GPA7	0.023**	0.018**	0.022**	0.022***	0.025***
	(0.012)	(0.009)	(0.009)	(0.008)	(0.010)
B. Reduced form					
$\widehat{e}$	0.090***	0.080**	0.008	0.003	-0.017
	(0.028)	(0.032)	(0.026)	(0.026)	(0.039)
C. First stage					
$\widehat{e}$	0.720***	0.699***	0.708***	0.714***	0.716***
	(0.058)	(0.059)	(0.057)	(0.057)	(0.062)
D. 2SLS					
GPA7	0.125***	0.115**	0.012	0.004	-0.023
	(0.040)	(0.046)	(0.037)	(0.036)	(0.054)
SD (GPA7)	1.65	1.65	1.65	1.65	1.65
F-stat (excl. instrument)	116.43	111.01	112.41	112.71	111.70
Observations	3443	3463	3420	3385	3363

Notes: GPA7 is the overall university GPA measured on the new grading scale from -3 (worst) to 12 (best). $\hat{e}$ is the residual from regressing post recoding GPA on a third order polynomial of the pre recoding GPA. All models are estimated with program fixed effects and the full set of covariates, which include enrollment year, an indicator for institution (Copenhagen or Aarhus) and interactions between enrollment year and institution, age at enrollment, an indicator for origin (non-western or western according to the definition by Statistics Denmark), parental income, parental unemployment, parental education (indicator for university degree) and gender. Parental variables are created as the mean across observed parents (except schooling, which is one if at least one parent has completed an university degree). We include indicators for the number of parents with respectively non-missing income, unemployment and education (i.e. 0, 1 or 2). Missing values are replaced with zeros, and an indicator for missing values is included. Bootstrap standard errors based on 500 iterations, clustered at the GPA level in parenthesis. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

of 0.7 shown in Table 2. Indeed, we find the the first stage coefficient is between 0.70 and 0.72, which support our identification of the model.

Panel D shows the IV coefficients that are equivalent to scaling the reduced form coefficients in panel B by the first-stage coefficients in panel D. The results show that a one-unit increase in GPA causes an 11-13 percent increase in earnings in the two first years after graduation. Scaling these coefficients by the GPA standard deviation implies that a one standard deviation increase in GPA causes a $12.5 \times 1.65 = 21$ percent increase in earnings in the first calendar year after graduation. The results is remarkably similar in the second year after graduation. Three years after graduation, the reform-induced GPA has no detectable effect on earnings. The point estimates for years three to five are smaller in magnitude and not statistically distinguishable from zero.

We can compare the expected "naive" OLS estimate from panel A with the IV estimate. Recall the two sources of bias in the OLS specification: First, the GPA is likely to be correlated with unobservables related to labor market outcomes. Second, the GPA is likely to be correlated with other signals used to screen workers. For the first source, we would clearly expect the OLS estimate to be upward biased and larger than the IV estimate because the omitted variables, such as ability or non-cognitive skills that affect earnings, are expected to be positively correlated with GPA. However, for the second source of bias, the direction is less obvious. Consider the case of gender, for example. If employers use gender to statistically discriminate workers and gender is correlated with GPA, the bias on the GPA coefficient could go in both directions.

In our case, the IV estimate is substantially larger than the OLS estimate. What could explain this? In an employer learning setting, we would expect the first source of bias not to be present in the initial periods after graduation. The employer has not observed worker productivity yet, and the earnings will not reflect true productivity but only the inferred productivity based on observed signals. The upward bias may, therefore, very well be limited in the initial period after graduation, and a potential downward bias from the second source would push the overall bias in the other direction.

4.3 Validity of the research design

Taken together, our findings support our predictions made in section 2. First, when workers enter the labor market, employers use workers' GPA as signals of productivity. Second, employers learn about workers' true productivity and the return to GPA in terms of earnings goes towards zero. We find a significant positive effect of the reform-induced variation in GPA on earnings in the two first years after graduation.

In this section, we empirically assess and discuss the identifying assumptions and the sensitivity of our findings. We first assess the exogeneity assumption, and then conduct several tests of the exclusion restriction and evaluate whether the results are sensitive to the choice of functional form.

The exogeneity assumption. For the exogeneity assumption to be met, the reform-induced variation in GPA needs to be unrelated to any individual characteristic that could affect earnings. Therefore, we estimate equation (4) using covariates as the dependent variables to test whether the reform-induced variation in GPA is related to any observable characteristics. Furthermore, we conduct a placebo test by implementing the recoding of grades for non-treated cohorts. We do this because we want to rule out the possibility of a combination of grades that is associated with higher earnings that also leads to a positive reform-induced variation in GPA.

Column (1) of Table 4 shows that the reform-induced variation in GPA is unrelated to individuals' gender, and columns (2) and (3) show that it is unrelated to high school GPA and undergraduate GPA. Columns (4), (5), and (6) reveal that the reform-induced variation is unrelated to parental income, employment, and education, respectively. Finally, in column (7), we construct a weighted average of all covariates by regressing log earnings in the first calendar year on all covariates and constructing predicted earnings based on the estimated coefficients. The coefficient in column (7) is both very small and not sta-

Table 4: Reform-induced variation GPA and observable characteristics. Dependent variables in column header.

	Female	High school GPA	Undergrad GPA	Parents' Income	Parents' Unempl.	Parents' uni degr.	Predicted earnings
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\hat{e}$	-0.016	-0.017	0.051	2.158	-0.003	0.014	-0.0024
	(0.018)	(0.048)	(0.057)	(1.976)	(0.004)	(0.020)	(0.018)
Observations	3808	3216	2046	3364	3810	3323	3810
Mean	0.65	0.75	-0.00	39.38	0.02	0.27	3.62

Notes: The table shows the coefficient from a reduced form regression using the variables denoted in the column headers as dependent variables. The mean refers to the mean of dependent variable. Parental variables are measured in the calendar year before graduation. Parental income is measured in 1000 EUR (2015 level). All models are estimated without covariates. Bootstrap standard errors based on 500 iterations and clustered on the GPA level in parenthesis. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

tistically different from zero, suggesting that the reform-induced variation is not related to a weighted average of all the observable characteristics. Together, the findings from Table 4 suggest that the reform-induced variation in GPA is not related to any observable characteristics of the students.

But even if the reform-induced variation in GPA is unrelated to observable characteristics, it might capture some unobserved qualifications rewarded on the labor market. To assess whether this is the case, we conduct a placebo test on students from a previous cohort that were not affected by the 2007 grading reform. More precisely, we implement the recoding of grades on 31 July 2004 for all students with less than 40 ECTS remaining in their master program. We then estimate the same models as in the main specification, to see whether the reform-induced variation in their GPA is related to their earnings. Recall that these students did not in reality, have their grades recoded. Any detectable effect would, therefore, be attributable to the underlying grade combination that caused this upgrade. As Table 5 shows, we find no effects of this placebo reform for the 2004 cohort in any of the first five years, suggesting that the reform-induced variation in GPA is not related to any underlying grade combination that affects earnings. Table 4 and 5 provide evidence of the reform-induced variation in GPA being exogenous.

Table 5: Placebo cohort. GPA and earnings regression results. Dependent variable: log earnings year one to five after graduation.

	Year after graduation				
	1	2	3	4	5
A. OLS specification					
GPA7	0.036***	0.031**	0.034***	0.029***	0.033***
	(0.014)	(0.012)	(0.009)	(0.009)	(0.009)
B. Reduced form					
$\widehat{e}$	-0.061	-0.011	-0.003	-0.022	0.013
	(0.061)	(0.050)	(0.039)	(0.046)	(0.038)
C. First stage					
$\widehat{e}$	0.503***	0.517***	0.509***	0.534***	0.535***
	(0.078)	(0.077)	(0.083)	(0.077)	(0.076)
D. 2SLS					
GPA7	-0.121	-0.022	-0.006	-0.042	0.024
	(0.135)	(0.099)	(0.079)	(0.086)	(0.070)
SD (GPA7)	1.67	1.68	1.67	1.67	1.68
F-stat (excl. instrument)	26.99	29.39	28.46	31.20	30.99
Observations	1900	1954	1955	1936	1924

Notes: This table resembles Table 3, but shows estimates based on a placebo sample. The placebo sample is obtained by implementing a placebo grading reform on July 31 in 2004 using the same recoding scheme, covariates and sample selection as in the main specification. See notes for Table 3. Bootstrap standard errors based on 500 iterations and clustered at the GPA level in parenthesis. * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

The exclusion restriction. We now turn to the second important assumption for the identification of our results. Even though the previous subsection indicate that the reform-induced variation in GPA is exogenous, the effect on subsequent earnings might still not be a result of GPA being used as a signal. The effect could still be driven by mediating factors. Recall that the students who experience the recoding are enrolled in a program and their subsequent student behavior could be affected by the recoding of their grades, and if this change in behavior is related to subsequent earnings, it could explain the observed relationship. For example, if students get their GPA downgraded by the reform, they could choose to work harder to get better grades or maybe choose courses where they think they will be able to get better grades. It could also have a discouraging effect, which could cause a student to drop out or increase his or her time to graduation.

To test this, we estimate the effect of the reform-induced variation in GPA on subsequent student behavior. First, we assess whether the recoding affected students likelihood of graduation. Second, we test whether the recoding affected performance in subsequent exams. Third, we examine whether the students select different optional units as a consequence of the recoding. Fourth, we test whether the recoding affected time to graduation.

To assess whether the students select different types of optional units, we calculate a measure of how demanding the subsequent units are. We do this by running a regression of exam results on unit fixed effects for the years 2000-2006. To capture selection into units, we further control for high school GPA. A larger unit fixed effect in this specification suggests that given high school GPA, a student receives a higher grade. In other words, the fixed effects capture unit difficulty (or grading generosity within a given unit). We then match these unit fixed effects to our treated cohort's units after the recoding of the grades, and use the fixed effects as the dependent variable.

Table 6 shows the effects of the reform-induced variation in GPA on subsequent student behavior. We find no behavioral responses in any of the four measures. The coefficients are very small compared to the variable means, and all coefficients are statistically indistinguishable from zero. The results support that the exclusion restriction is met.

Sensitivity of the model specification. We now assess whether our findings are sensitive to the empirical specification. Table 7 shows the reduced form coefficients for six different specifications. First, we show the estimates from our preferred specification us-

Table 6: Reform-induced variation GPA and observable characteristics. Dependent variables in column headers.

	(1) Graduated	(2) Time to graduation	(3) Unit FE	(4) Post GPA
$\hat{e}$	0.000	-0.004	-0.012	-0.057
	(0.000)	(0.035)	(0.030)	(0.163)
Observations	4045	4045	3913	4576
Mean	1.000	0.968	1.479	7.565

Notes: The table shows the coefficient from a reduced form regression using the variables denoted in the column headers as dependent variables. Graduated is an indicator for whether the focal individual graduated before 2011. Time to graduation is time from the recoding to graduation measured in years. Unit FE is average unit specific fixed effects of units completed after the recoding. The fixed effects are estimated based based on pre-reform cohorts, by regressing exam grade as the dependent variable on unit indicators capturing the fixed effects and high school GPA. A positive fixed effect suggest that conditional on high school GPA this unit has historically been graded more generously. Post GPA is the grade point average of all units completed after the recoding. The mean refers to the mean of dependent variable. All models are estimated with the full set of covariates (see notes for Table 3). Bootstrap standard errors based on 500 iterations and clustered on the GPA level in parenthesis. * p$<$ 0.1, ** p$<$ 0.05 and *** p$<$ 0.01.

ing a third-order polynomial. Second, we show in row 2 and 3 of Table 7 that coefficients are very similar to the baseline specification from row 1 when we use second- and fourth-order polynomials to identify the reform-induced variation in GPA.

In rows 3 and 4, we show that we also obtain very similar coefficients when using a less parametric approach to identify the reform-induced variation in GPA in terms of within pre-recoding GPA comparisons. The reform-induced variation in GPA is obtained by calculating the difference in the individual's recoded GPA to the mean (row 4) or the median (row 5) recoded GPA of all other students with the same GPA prior to the recoding.

Finally, row 6 of Table 7 shows that we obtain similar coefficients when estimating a specification without covariates. Estimations using subsets of the covariates (not shown) reveal that the main difference between the model without covariates and the main spec-

Table 7: GPA and earnings regression results using alternative specifications. Dependent variable: log earnings year one to five after graduation.

	Year after graduation				
	1	2	3	4	5
3rd order pol.	0.090***	0.080**	0.008	0.003	-0.017
	(0.028)	(0.032)	(0.026)	(0.026)	(0.039)
2nd order pol.	0.083***	0.074**	0.007	-0.002	-0.014
	(0.027)	(0.032)	(0.026)	(0.026)	(0.037)
4th order pol.	0.089***	0.077**	0.010	0.003	-0.014
	(0.028)	(0.033)	(0.027)	(0.027)	(0.039)
Deviation from mean	0.081**	0.084**	0.011	-0.001	-0.029
	(0.032)	(0.036)	(0.028)	(0.030)	(0.037)
Deviation from median	0.074**	0.067**	0.010	-0.002	-0.029
	(0.032)	(0.033)	(0.026)	(0.028)	(0.036)
No covars	0.087***	0.075**	-0.000	-0.012	-0.028
	(0.030)	(0.035)	(0.029)	(0.026)	(0.041)
Observations	3443	3463	3420	3385	3363

Notes: The table shows the reduced form coefficients for the years 1 to 5 using alternative specifications. Rows one and two show the coefficients from using respectively a second and fourth order polynomial in the regression of post recoding GPA on pre recoding GPA. Row three (four) shows coefficients from a specification, where we identify the reform-induced variation in GPA, by comparing each individuals post-recoding GPA to the mean (median) of the recoded GPA among everyone else with the same pre-recoding GPA. Row five shows the coefficients from estimating a model without any covariates. All models, with the exception of the no-covariates specification, are estimated with the full set of covariates (see notes for Table 3). Bootstrap standard errors based on 500 iterations and clustered at the GPA level in parenthesis. * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

ification comes from the exclusion of program fixed effects.

Taken together, all the models presented in Table 7 suggest that our results are robust in all model specifications.

Is the relationship symmetric? So far, we have been assuming a linear relationship between the reform-induced variation in GPA and log earnings. However, there are reasons to expect a non-linear relationship. For example, job candidates who receive a negative recoding shock ($\hat{e} < 0$) have an incentive to inform employers about this, while job candidates who receive a positive recoding shock ($\hat{e} > 0$) have no incentive to do so. It might, therefore, be the case that the relationship is driven by the positive shocks.

Figure 6 shows the relationship between residualized earnings and residualized reform-induced GPA. We estimate the relationship using a natural cubic spline with three knots. The more flexible relationship (compared to the OLS relationship) shows a positive relationship throughout and is always within the 95-percent confidence interval of the linear OLS relationship. Our results show that the returns to the reform-induced variation appear to be fairly linear across the entire scale of the reform-induced variation. In other words, earnings in year one are increasing linearly in the reform-induced variation.

Sample selection sensitivity. Recall that we restrict our sample by the students' progress in the master program at the time of the grading reform. Ideally, we would restrict the sample to the students who do not have any course work left before graduation. However, this is not easily done in a Danish setting where the programs are very flexible. Instead, we have restricted the sample to students who have 40 credit points or less left at the time of the reform. In Figure 7 we assess whether our findings are sensitive to the sample selection. We show the reduced form coefficient for the relationship between reform-induced variation in GPA and earnings in the first calendar year after graduation using different bandwidths. The values on the horizontal axis show the remaining credit points. The smaller the number is, the closer the students are to graduation at the time of the recoding. We note that the point-estimate increases slightly (but not significantly) as we narrow the sample compared to our main specification of 40 credit points; however, as expected, the standard errors increase (as the increased 95-percent confidence intervals show).

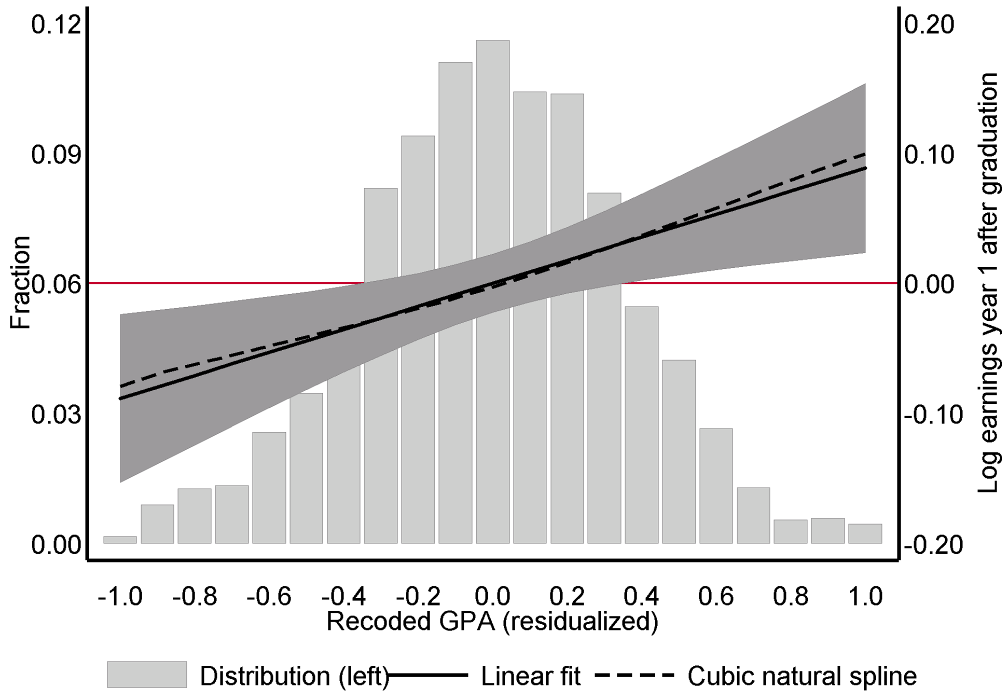


Figure 6: The non-parametric relationship between log earnings in the calendar year after graduation and reform induced variation in GPA.

Notes: The solid line shows the linear relationship estimated in our main specification. The shaded area shows the 95-percent confidence interval obtained through 500 bootstrap iterations clustered on the pre reform GPA level. The dashed line shows the natural cubic spline using three knots. The reform-induced variation in GPA and the log earnings are residualized using all covariates in the main specification and program fixed effects.

4.4 Mechanisms and heterogeneity

To delve into how the effect of the reform-induced variation in GPA affects labor market outcomes of recent university graduates this section will present different results of both mechanisms that could drive the effect and if subgroups of workers are affected differently. It could, for example, be that the effect that the effect is primarily driven by workers in the private sector, or that the disappearance of the effect after year two is caused by workers changing jobs; or perhaps it is the case that the reform-induced variation only affect specific subgroups of graduates. This section will take a closer look into the mecha-

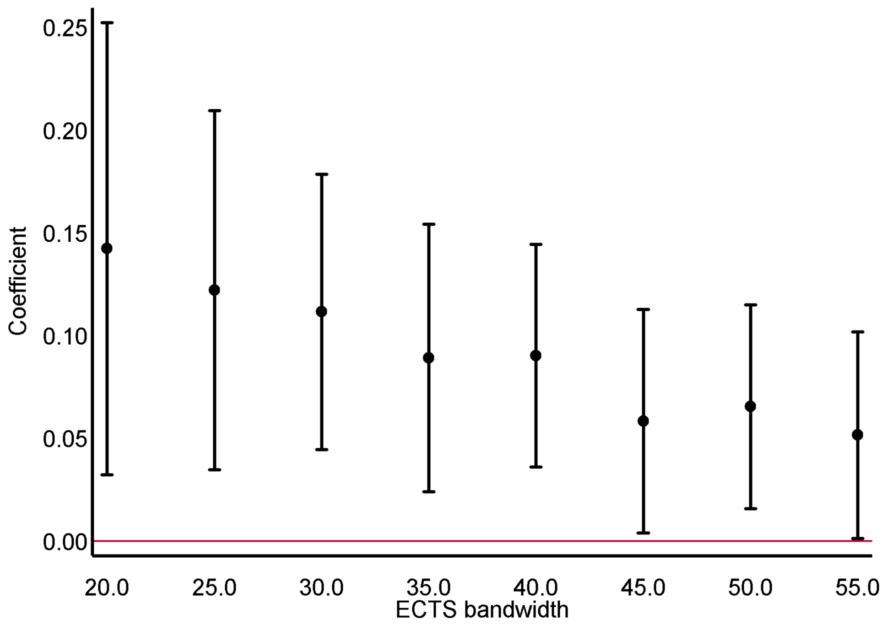


Figure 7: Bandwidth sensitivity.
Notes: The graph shows the reduced form point-estimate and 95-percent confidence intervals using the sample selection shown on the horizontal axis. The sample selection refers to the number of ECTS remaining at the time of the grading reform.

nism and hererogeneity of the effect of GPA on earnings.

Mechanisms. In Table 8, we investigate the effect of the reform-induced variation in GPA on other labor market outcomes over the first five years after graduation. We estimate equation (4) whereby, instead of using log earnings as dependent variable, we estimate the model with different labor market outcomes. In panel A we estimate the model with an indicator for positive earnings as the dependent variable.In panel B we accumulate income over the first five years after graduation to investigate the persistence of the effect of the reform-induced variation on earnings. For each year, the previous years' earnings are accumulated, which means that the earnings estimate in year five show the effect of the reform-induced variation on total earnings over the first five years after graduation. In panel C, we look at whether the reform-induced variation in GPA has an effect on whether the workers work in the public or private sector. Panel D presents the results on job change, and panel E considers earnings growth and earnings growth conditional on remaining in the same job.

The results in panel A reveals no effect of the reform-induced variation on employment (earnings > 0). Using accumulated earnings as an outcome variable in panel B, we find that the effect of reform-induced variation on earnings is persistent through all five years on the labor market. Scaling the coefficient with the first stage estimate (0.7) and the standard deviation of GPAs (1.65), the results show that one standard deviation increase in GPA causes an $1.65 \times 0.035/0.7 = 0.0825$SD increase in earnings.

We do not find any evidence of the reform-induced variation in GPA having an effect on whether the worker works in the public or private sector (panel C).

The results in panel D show that the reform-induced variation does not affect job changes. In panel E, we closely examine the earnings growth both in total and for workers staying in the same job, and as expected, we find that a positive reform-induced variation in GPA is related to lower wage growth from year two to three; we also find this both across and within employers. While the point-estimate is considerably larger for individuals changing jobs, the group is comparably small and we lack power to identify the precise effects.

Heterogeneity In Table 9, we assess whether effects vary by subgroups. First, in panel A, we find some evidence of stronger effects for men compared to women, although the difference is not significant. Second, the results in panel B suggest that effects are largest for children of parents with no university degree. One might expect that graduates with better networks on the relevant job market are less reliant on the GPA as a signal. In this case, having parents with a university degree could act as a proxy for network and

Table 8: GPA and labor market outcomes - regression results.

	Year after graduation				
	1	2	3	4	5
A. The extensive margin					
Earnings>0	-0.007	0.005	0.000	-0.010	-0.007
	(0.012)	(0.013)	(0.013)	(0.011)	(0.013)
B. Accumulated earnings					
Log accumulated earnings	0.090***	0.054**	0.050***	0.044***	0.035**
	(0.028)	(0.021)	(0.018)	(0.017)	(0.017)
C. Sector choice					
Public sector	0.010	0.001	-0.003	0.006	0.034
	(0.017)	(0.019)	(0.020)	(0.023)	(0.052)
D. Job change					
Job change		-0.018	0.021	-0.016	0.013
		(0.017)	(0.020)	(0.023)	(0.020)
E. Earnings growth					
Earnings growth		-0.008	-0.069***	-0.024	-0.016
		(0.031)	(0.026)	(0.021)	(0.031)
Earnings growth, cond. on same job		0.030	-0.046*	-0.029*	-0.004
		(0.029)	(0.024)	(0.018)	(0.023)
Earnings growth, cond. on job change		-0.120	-0.136	-0.021	-0.035
		(0.090)	(0.088)	(0.069)	(0.109)

Notes: The table shows the reduced form coefficients for the relationship between reform-induced variation in GPA and the outcomes shown in the row titles across the five years shown in the column header. Data on unemployment and disposable income is not available after 2013 All models are estimated with the full set of covariates (see notes for Table 3). Bootstrap standard errors based on 500 iterations and clustered at the GPA level in parenthesis. * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

confirm this hypothesis.

Third, in panel C of Table 9 we assess whether the effects are larger at the top or the bottom of the GPA distribution. The effects appear to be larger at the bottom, although no difference is statistically significant. The students at the bottom of the GPA distribution might benefit from an upgrade of their GPA more than the students with already high GPAs if employers use the GPA for initial screening of job applications. One could imagine that a positive reform-induced variation in GPA, with everything else equal, increases

Table 9: GPA and earnings regression results across subgroups. Dependent variable: log earnings year one to five after graduation.

	Year after graduation				
	1	2	3	4	5
A. By gender					
Male	0.152**	0.105	0.037	0.018	-0.059
	(0.062)	(0.064)	(0.058)	(0.053)	(0.069)
Female	0.068	0.065	-0.010	-0.006	0.004
	(0.043)	(0.043)	(0.033)	(0.030)	(0.038)
Difference	-0.084	-0.040	-0.047	-0.024	0.064
	(0.085)	(0.087)	(0.071)	(0.058)	(0.072)
B. By parental education					
No university degree	0.099***	0.123***	0.023	0.016	0.009
	(0.037)	(0.042)	(0.037)	(0.035)	(0.046)
University degree	0.044	-0.044	-0.019	-0.037	-0.114*
	(0.050)	(0.058)	(0.046)	(0.052)	(0.062)
Difference	-0.055	-0.167**	-0.042	-0.053	-0.124*
	(0.067)	(0.074)	(0.068)	(0.067)	(0.068)
C. By GPA					
Below median	0.112**	0.100*	0.015	0.005	-0.042
	(0.045)	(0.054)	(0.046)	(0.033)	(0.056)
Above median	0.055	0.041	-0.009	-0.047	-0.024
	(0.046)	(0.053)	(0.043)	(0.045)	(0.060)
Difference	-0.057	-0.059	-0.024	-0.052	0.018
	(0.068)	(0.077)	(0.063)	(0.057)	(0.081)
D. By educational program wage dispersion					
Below median	0.052	0.073**	-0.014	-0.060**	-0.072
	(0.038)	(0.035)	(0.031)	(0.029)	(0.048)
Above median	0.149***	0.083	0.035	0.090*	0.067
	(0.050)	(0.071)	(0.054)	(0.047)	(0.059)
Difference	0.096	0.010	0.049	0.150***	0.139*
	(0.064)	(0.084)	(0.066)	(0.050)	(0.074)
E. By educational program public sector employment					
Below median	0.107	0.202***	0.038	0.034	0.023
	(0.066)	(0.071)	(0.065)	(0.064)	(0.079)
Above median	0.081***	0.035	-0.004	-0.002	-0.036
	(0.030)	(0.038)	(0.033)	(0.035)	(0.045)
	(0.012)	(0.010)	(0.012)	(0.008)	(0.009)
Difference	-0.025	-0.166**	-0.042	-0.035	-0.060
	(0.074)	(0.083)	(0.077)	(0.080)	(0.093)

Notes: The table shows the reduced form coefficients the relationship between reform induced variation in GPA and log-earnings for the calendar years one to five after graduation across subgroups. All models are estimated with the full set of covariates (see notes for Table 3). Bootstrap standard errors based on 500 iterations and clustered at the GPA level in parenthesis. * p< 0.1, ** p< 0.05 and *** p< 0.01.

the probability of crossing a screening GPA threshold in a hiring process for graduates with lower GPAs compared to graduates with high GPAs because, the high-performing graduates will most likely already be above the threshold.

Panels D and E show labor market differences. First, we assess whether the reform-induced variation in GPA matters more for degrees with greater earnings dispersion. We find suggestive evidence for this hypothesis, although the differences are not significant. Second, we assess if the effect of the reform-induced variation differs between degrees with high or low public sector employment. We find no clear pattern on the effect between these types of degrees.

5 Conclusion

Using variation in university students' GPAs caused by a grading reform, we document a signaling value of university GPAs, and that employers rapidly learn about variation in GPA that is unrelated to labor market productivity. We identify an exogenous variation in GPA uncorrelated with labor market productivity. We find no evidence on extensive margin labor market outcome effects and job changes, but evidence that the earnings adjustment over time occurs both within and across firms. We show that the signaling value of GPAs is strongest for men and children of parents with no university degree. The latter result could suggest that signals are more relevant to workers with no informal connections to the relevant labor market. Finally, we find suggestive evidence that the signaling effect is strongest for majors that are related to larger wage dispersion and strongly connected to the private sector.

Our finding that university grades explain sorting for labor market entrants also has policy implications. First, it suggests that the grading system influences the labor market matching process. Grade inflation, for example, makes it harder for employers to identify the best applicant. Moreover, systems that focus on some parts of the achievement distribution (e.g., through honor degrees) might involve lower matching efficiency at the lower end of the distribution.

Second, our results illustrate the importance of developing systems that produce accurate and reliable skill signals. A range of factors can affect assessments, including pollution (Ebenstein et al., 2016), weather (Goodman et al., 2018), time of the day (Sievertsen et al., 2016), and teacher manipulation (Dee et al., forthcoming; Diamond and Persson, 2016). Our finding that grades are used as signaling device suggests that such external factors could have implications for labor market outcomes.[7]

Finally, we should mention that we use a unique setting to identify the signaling value of GPAs, and the overall circumstances of a grading reform are not uncommon. Grading reforms are relatively widespread, and although the implementations might vary from reform to reform, they will typically generate some noise in the signaling process. However, our results can be more generalizable. Given that we find a signaling effect in a setting where employers can observe almost precise information on educational achievement, our results likely constitute a lower bound in terms of the employer learning process. Future research on signaling and employer learning based on educational achievement could provide more insights into this learning process exploiting other sources of variation in signals.

[7]It is worth noting that none of the studies above are concerned with assessment in a university setting. However, there is substantial anecdotal evidence for errors in grading in higher education (see e.g., Nightingale (2017)). This present paper provides evidence of one source of "noise" in university graduates signal of educational achievement. While the grading reform we use is unique, grading reforms are widespread, and the transition to a new scale might involve substantial noise.

References

Altonji, J. G. and C. R. Pierret (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics 116*(1), 313–350.

Arcidiacono, P., P. Bayer, and A. Hizmo (2010, October). Beyond signaling and human capital: Education and the revelation of ability. *American Economic Journal: Applied Economics 2*(4), 76–104.

Arrow, K. J. (1973). Higher education as a filter. *Journal of Public Economics 2*(3), 193–216.

Aryal, G., M. Bhuller, and F. Lange (2019). Signaling and employer learning with instruments. Technical report, National Bureau of Economic Research.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy 70*(5, Part 2), 9–49.

Bordón, P. and B. Braga (2017). Employer learning, statistical discrimination and university prestige. *mimeo*.

Christensen, J. J. and N. Westergard-Nielsen (1999). Wages and human capital: the danish evidence. *Returns to Human Capital in Europe*, 29–42.

Clark, D. and P. Martorell (2014). The signaling value of a high school diploma. *Journal of Political Economy 122*(2), 282–318.

Dee, T. S., W. Dobbie, B. A. Jacob, and J. Rockoff (forthcoming). The causes and consequences of test score manipulation: Evidence from the new york regents examinations. *American Economic Journal: Applied Economics*. forthcoming.

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

Diamond, R. and P. Persson (2016). The long-term consequences of teacher discretion in grading of high-stakes tests. *National Bureau of Economic Research* (22207).

Ebenstein, A., V. Lavy, and S. Roth (2016). The long-run economic consequences of high-stakes examinations: evidence from transitory variation in pollution. *American Economic Journal: Applied Economics 8*(4), 36–65.

Farber, H. S. and R. Gibbons (1996). Learning and wage dynamics. *The Quarterly Journal of Economics 111*(4), 1007–1047.

Fredriksson, P., L. Hensvik, and O. N. Skans (2018). Mismatch of talent: Evidence on match quality, entry wages, and job mobility. *American Economic Review*.

Freier, R., M. Schumann, and T. Siedler (2015). The earnings returns to graduating with honors — evidence from law graduates. *Labour Economics 34*, 39 – 50. European Association of Labour Economists 26th Annual Conference.

Goodman, J., M. Hurwitz, J. Park, and J. Smith (2018). Heat and learning. *National Bureau of Economic Research* (24639).

Huntington-Klein, N. (2018). Human capital vs. signaling is empirically unresolvable. *California State University Fullerton - Department of Economics Working Paper Series* (2018/010).

Hvidman, U. and H. H. Sievertsen (2018). Grades and student behavior. *University of Bristol Discussion Paper 18/698*.

Jepsen, C., P. Mueser, and K. Troske (2016). Labor market returns to the ged using regression discontinuity analysis. *Journal of Political Economy 124*(3), 621–649.

Khoo, P. and B. Ost (2018). The effect of graduating with honors on earnings. *Labour Economics 55*, 149 – 162.

Koedel, C. and E. Tyhurst (2012). Math skills and labor-market outcomes: Evidence from a resume-based field experiment. *Economics of Education Review 31*(1), 131–140.

Lacetera, N., D. G. Pope, and J. R. Sydnor (2012). Heuristic thinking and limited attention in the car market. *American Economic Review 102*(5), 2206–36.

Lange, F. (2007). The speed of employer learning. *Journal of Labor Economics 25*(1), 1–35.

National Association of Colleges and Employers (2016). Job Outlook Report 2017. *NACE*.

Nightingale, V. (2017). University grading errors are more common than you might think. *Times Higher Education*.

Piopiunik, M., G. Schwerdt, L. Simon, and L. Woessmann (2018, January). Skills, Signals, and Employability: An Experimental Investigation. IZA Discussion Papers 11283, Institute for the Study of Labor (IZA).

Protsch, P. and H. Solga (2015). How employers use signals of cognitive and noncognitive skills at labour market entry: insights from field experiments. *European Sociological Review 31*(5), 521–532.

Riley, J. G. (2001). Silver signals: Twenty-five years of screening and signaling. *Journal of Economic literature 39*(2), 432–478.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers 3*(2), 135–146.

Sattinger, M. (1975). Comparative advantage and the distributions of earnings and abilities. *Econometrica (pre-1986) 43*(3), 455.

Schönberg, U. (2007). Testing for asymmetric employer learning. *Journal of Labor Economics 25*(4), 651–691.

Schultz, T. W. (1961). Investment in human capital. *The American Economic Review*, 1–17.

Sievertsen, H. H., F. Gino, and M. Piovesan (2016). Cognitive fatigue influences students' performance on standardized tests. *Proceedings of the National Academy of Sciences 113*(10), 2621–2624.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 355–374.

Wise, D. A. (1975). Academic achievement and job performance. *The American Economic Review*, 350–366.

A Additional material

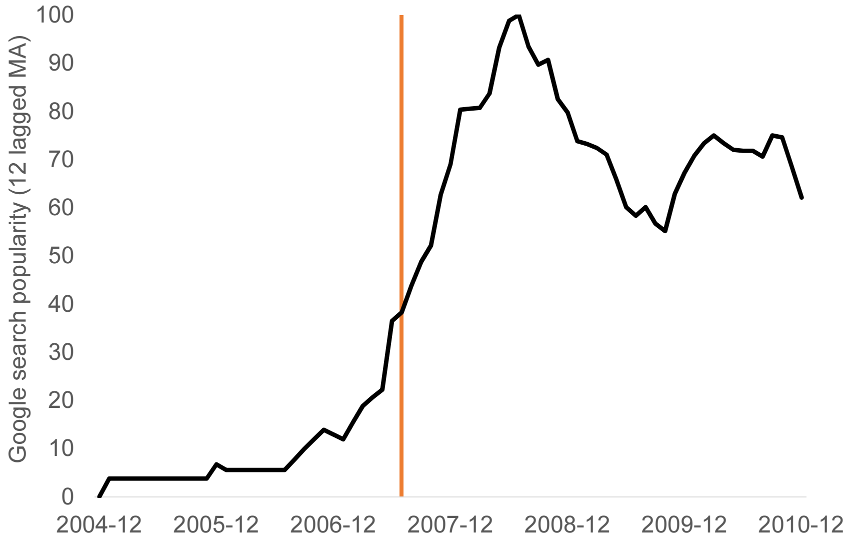


Figure A.1: Google search trend for ”den nye karakterskala” (English: ”the new grading scale”).

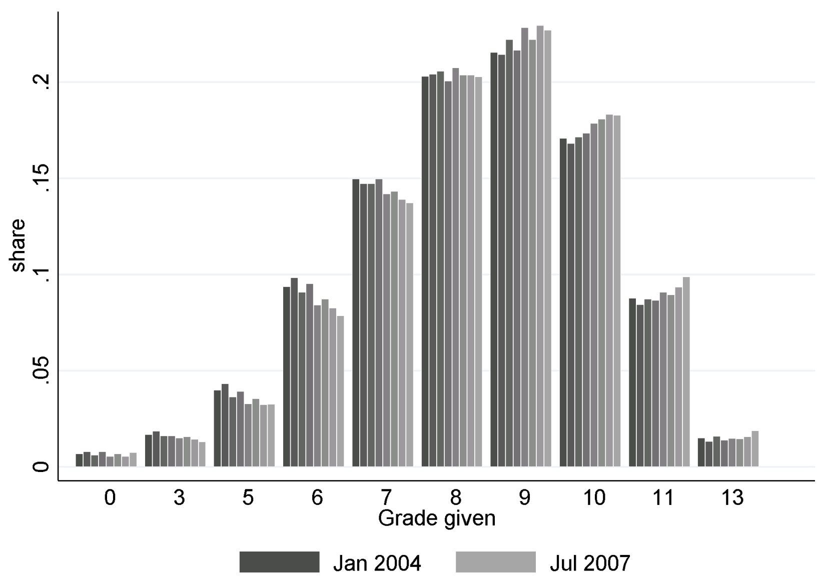


Figure A.2: Relative grade frequency

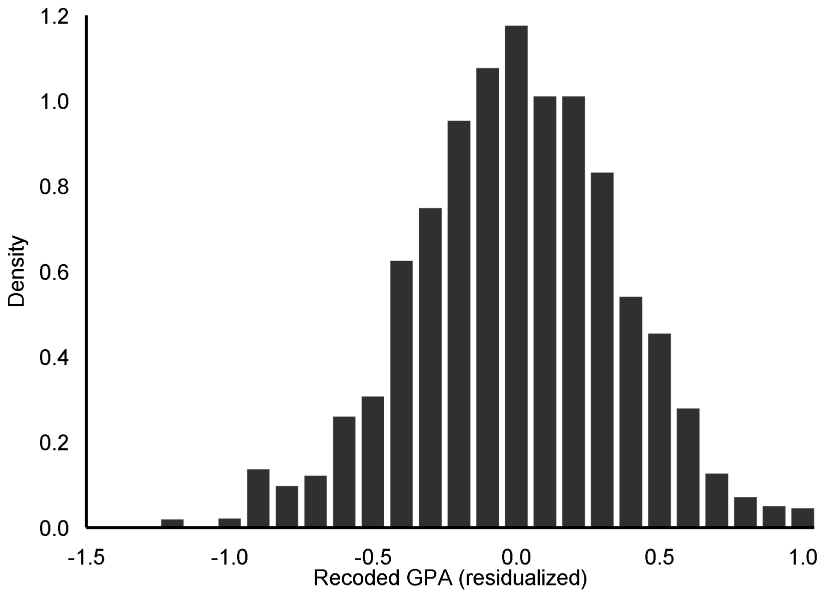


Figure A.3: Histogram treatment variable

B University grade point average and earnings

While the correlation between years of schooling and earnings is well-established, considerably less is known about the correlation between grade point averages and earnings. In this section we document that a higher GPA is associated with higher earnings.

Before we turn to the association between GPA and earnings Figure B.1 shows the overall variation in the GPA. The average GPA in the covered cohorts was 9.27 on scale from 6 to 13 (6 is the lowest passing grade), and the standard deviation is one. The distribution is fairly symmetric, with a median almost identical to the mean of 9.25.

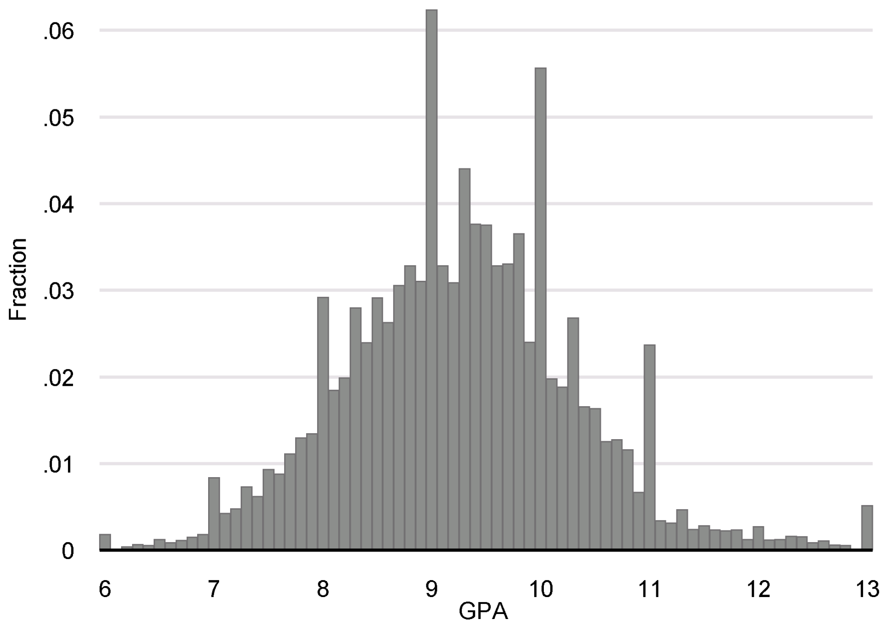


Figure B.1: The GPA distribution for university graduates.

Notes: The sample consists of all students who commenced and finished their postgraduate studies in the years 2000-2007 and who received their final GPA on the 13-scale, at the University of Copenhagen or Aarhus University.

Table B.1 shows the correlation between university GPA and earnings. In the first calendar year after graduation a one standard deviation higher GPA is associated with seven percent higher earnings. In the fifth calendar year after graduation a one standard deviation higher GPA is associated with four percent higher earnings. As a comparison, Christensen and Westergard-Nielsen (1999) finds that one additional year of schooling is associated with 4.5 percent higher earnings in Denmark.

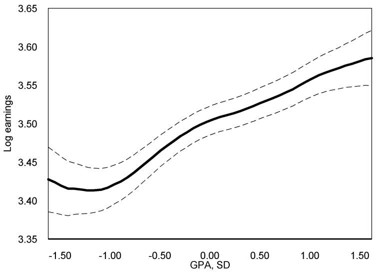
Table B.1: Regression results. Dependent variable: log earnings

	Year after graduation				
	1	2	3	4	5
GPA	0.072^{***}	0.048^{***}	0.056^{***}	0.047^{***}	0.037^{***}
	(0.022)	(0.013)	(0.011)	(0.013)	(0.013)
Observations	14,337	14,560	14,583	14,493	14,384

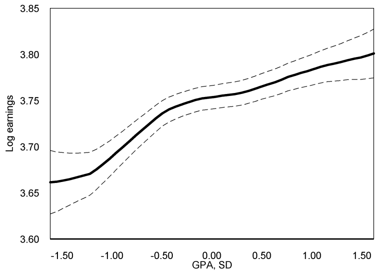
Notes: The table shows the coefficients from regressing log earnings in year 1 to 5 after graduation on university GPA. The university GPA is standardized to a mean of zero and standard deviation of one within university major, grading scale and graduation year. The sample consists of all students who commenced and finished their postgraduate studies in the years 2000-2007 at the University of Copenhagen or Aarhus University. Standard errors clustered on the major level (59 levels) in parenthesis parenthesis. Asterisks indicate significance at the following levels: * $p< 0.1$, ** $p< 0.05$ and *** $p< 0.01$.

Figure B.2 shows the relationship between earnings and university GPA in a less parametric way using local regressions. While The relationship seems to flatten of a bit in the lower part of the distribution, the overall relationship appears fairly linear.

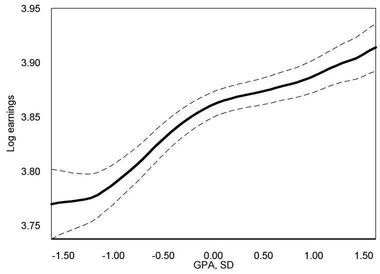
While the results in Table B.1 and Figure B.2 simply confirm an association between GPA and earnings, Figure B.3 provides one type of test for signaling inspired by the evidence on left-digit bias (Lacetera et al., 2012). If employers use GPAs in their screening process and focus on the left-most digit, we would expect a discontinuity in the relationship between earnings and GPA across integer thresholds. We therefore computed the distance to the nearest integer and show the relationship between earnings and this distance in Figure B.3. We observe no sign of a discontinuity in earnings at the integer thresholds.



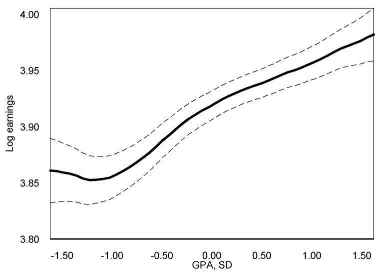
(a) Year 1



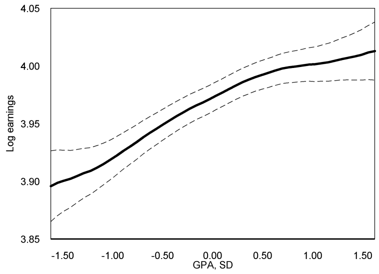
(b) Year 2



(c) Year 3



(d) Year 4



(e) Year 5

Figure B.2: Log earnings in the first five calendar years after graduation and university grade point averages.

Notes: The figures show local regressions of log earnings in the calendar years 1 to 5 after graduation on university GPA. The local regression is computed using a triangular kernel, a bandwidth of 0.75 and a degree of zero. The sample consists of all students who commenced and finished their postgraduate studies in the years 2000-2007 at the University of Copenhagen or Aarhus University. The dashed lines show the 95 percent confidence intervals.

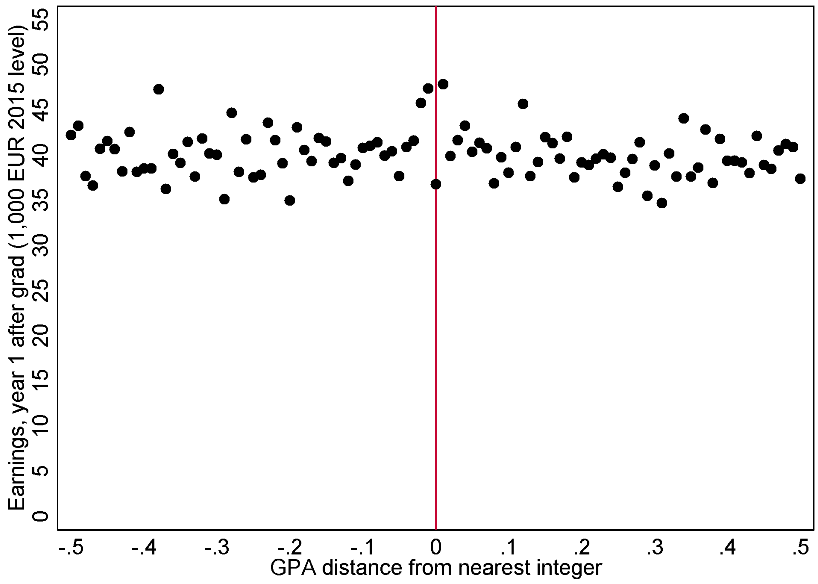


Figure B.3: Log earnings in the first calendar year after graduation and GPA distance to nearest integer.

Notes: The sample consists of all students who commenced and finished their postgraduate studies in the years 2000-2007 at the University of Copenhagen or Aarhus University.

C Monte Carlo simulations

In this section we simulate the grading reform to assess the validity of our research design. Specifically, we investigate whether our method to measure the reform-induced variation in GPA leads to the expected hypotheses rejection behavior when the GPA is respectively correlated and uncorrelated with earnings.

The Data Generating Process

- N individuals
- with unobserved ability $a \sim U(0, 100)$,
- they attend an exam and score $e \sim \mathcal{N}(a, 25)$
- exam scores are translated into grades based on the observed distribution.
- each student receives 5 grades.
- each grade is transformed to the 7-point scale, and then GPA13 and GPA7 are computed as the simple average of all grades.
- earnings (y) are a function of grades and ability: $y = 10 + 0.3a + \gamma GPA7 + \varepsilon$

  (where $\varepsilon \sim \mathcal{N}(0, 1)$)

Rejection rates.

- We let γ be between 0 (grades should have no effect, given ability) to 0.5.
- We estimate the relationship between earnings and the recoding "noise" using five specifications.

An Illustration of the Variation in GPA

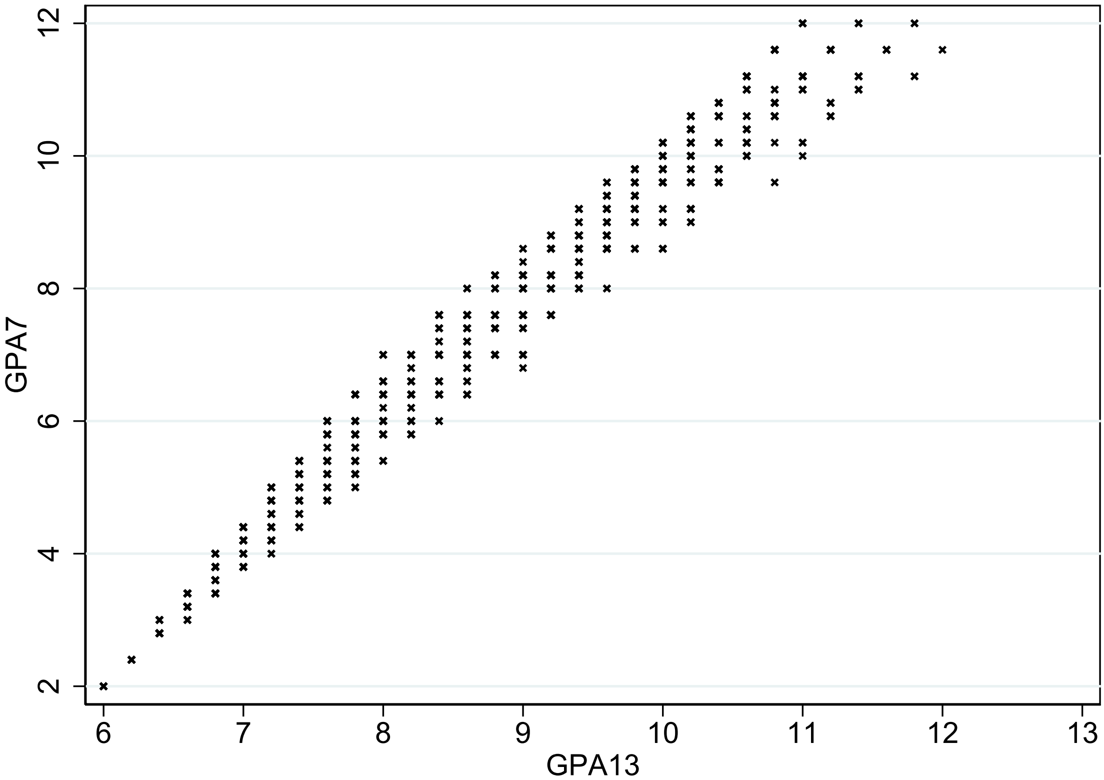


Figure C.1: A simulated example

– Spec 1-4: $log(y) = \alpha_0 + \alpha_1 GPA7 + f(GPA13) + u$,

  where $f()$ is respectively a 1st, 2nd, 3rd, and 4th order polynomial.

– Spec 5: Median deviation (see main text). Deviation between recoded GPA and the median recoded GPA among everyone with the same original GPA.

- We run 10,000 replications with $N = 5000$.

- We then check how often we reject $H0 : \alpha_1 = 0$ on a 5 percent level.

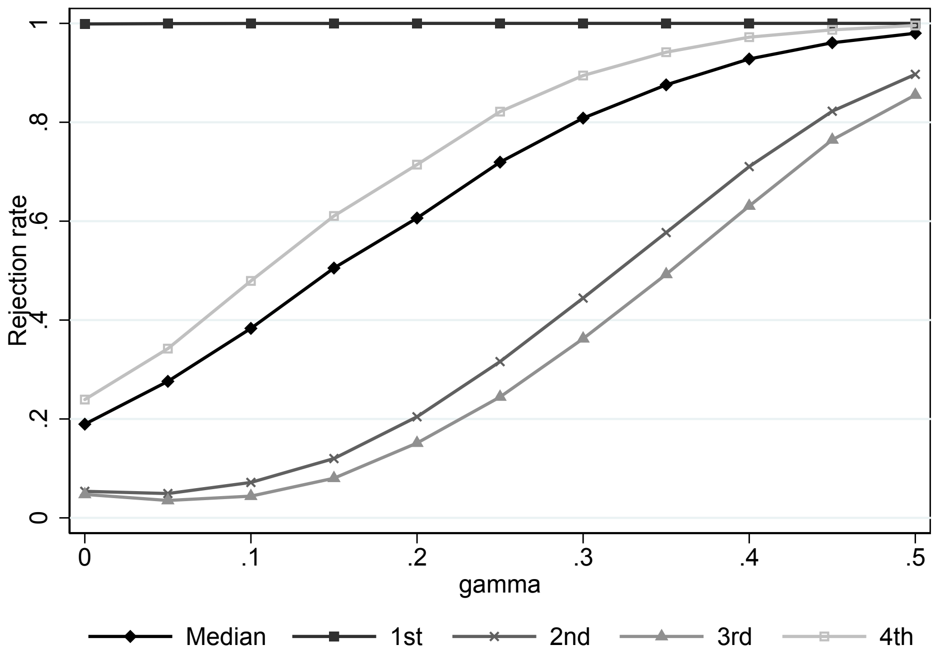


Figure C.2: Rejection rates