Grades and Employer Learning

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

[1]VIVE [2]University of Bristol [3]IZA

Version: December 5, 2022

Abstract

We identify the labor market returns to university grade point averages (GPA) by leveraging a nationwide change in the scaling of grades in Danish universities. Our results show that a reform-induced increase in GPA that is unrelated to ability causes higher earnings immediately after graduation, but the effect fades in subsequent years. The effect at labor market entry is largest for individuals with fewer alternative signals. Although employers initially screen candidates based on skill signals, our findings are consistent with a model in which employers rapidly learn about worker productivity.

Keywords: job market signaling, employer learning, higher education
JEL Classification: I20, J20, I26

[*]We thank Simon Burgess, Thomas S. Dee, Mette Ejrnæs, Fabian Lange, Christine Valente, and Miriam Wüst, as well as participants for their valuable comments at the 2017 ESPE meeting in Glasgow, the 2018 RES annual meeting in Sussex, the 2018 CEN meeting at Copenhagen Business School, the 2018 IWAEE in Catanzaro, the 2018 IZA workshop on the Economics of Education in Bonn, the 2018 DGPE meeting in Sønderborg, the 2019 SOLE annual meeting, and the seminar participants in Aarhus, Bath, Bristol, Copenhagen, Southampton, and Trondheim. We would further like to thank Marianne Toftegaard Hansen and Bjørn Bjørnsson Meyer for their assistance in obtaining access to the data. Sievertsen acknowledges receipt of financial support from the Danish Council for Independent Research through grant DFF: 4182-00200.
Hansen: ath@vive.dk
Hvidman: ulhv@vive.dk
Sievertsen: h.h.sievertsen@bristol.ac.uk

1 Introduction

A major challenge for employers is to identify the best-suited applicants for jobs. As productivity at labor market entry is imperfectly observed, employers must use signals of skills. While some evidence-based curriculum vitae designs suggest that employers value education credentials (Koedel and Tyhurst, 2012; Protsch and Solga, 2015; Piopiunik et al., 2020), little is known about the actual impact of such skill signals on labor market outcomes.

This paper examines the signaling value of the grade point average (GPA) of a graduate job applicant at labor market entry and how this effect persists over time in the labor market. To empirically differentiate the impact of the signals from actual skills, we exploit a grading reform in Denmark that mimics an ideal experiment by creating variation in university graduates' GPAs. Students who were enrolled in university during the implementation of the reform had their existing grades recoded to a new grading scale based on a scheme by the Ministry of Education. The recoding caused substantial variation in GPAs that is unrelated to ability, as two individuals with identical pre-reform GPAs could end up with more than a standard deviation difference in their GPAs due to the recoding. We use this reform-induced change in GPA to produce credible estimates of the effects of grade signals that are unrelated to achievement in a naturally occurring setting.

We use a data set containing GPAs for all students at the two largest universities in Denmark, Aarhus University and the University of Copenhagen, corresponding to around half of the total population of university students in Denmark.[1] Thus, we examine variation in salient grades among a large and diverse group of university students. Moreover, the detailed Danish administrative data allow us to study how the effect of the grade signal changes over time.

Our findings show that a reform-induced increase in GPA causes higher earnings in the initial years after graduation. However, this effect diminishes over time, and there is no detectable effect three years after graduation, which may suggest a rapid employer learning process. We assess the validity of the design and 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 three non-treated cohorts.

The effect at labor market entry is strongest among graduates with parents without a

[1] About one in ten of the Danish population aged 30–35 graduated from either the University of Copenhagen or Aarhus University and more than one in five people in Denmark with a yearly income of more than 60,000 euros graduated from one of the two universities (see Appendix Table A.1)

university degree, graduates with a short work history, and graduates from areas with a small student labor market. These findings suggest that the signaling value of the GPA is particularly important for graduates with limited alternative signals such as informal links (proxied by parental education) and work history. We also find that the effect is strongest for graduates from majors that are more likely to be employed in the private sector and in jobs with greater earnings dispersion. This indicates that the signaling effect is largest in less-regulated labor markets and in labor markets with high variation in earnings.

While there may be other explanations, the diminishing effect of the reform-induced variation on earnings over time is indicates an effect of an employer learning process. Looking into the wage adjustment process, we find no evidence of a link between reform-induced variation in GPAs and job changes in the first five years after graduation. Instead, we do find a slower earnings growth for individuals who experienced a positive reform-induced change in GPA in their second to third year after graduation. Although the adjustment is fastest among workers who change employer, the slower earnings growth is detectable for all workers, which suggests that the earnings adjustments occur both within and across employers.

Overall, these results demonstrate that 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. Moreover, our findings suggest that employer learning happens rapidly. An initially substantial earnings premium to variation in a signal of educational achievement that is unrelated to labor market productivity diminishes, and the adjustment occurs both within and across employers.

This study contributes to the literature on labor market sorting and employer learning (Arrow, 1973; Wise, 1975; Farber and Gibbons, 1996; Altonji and Pierret, 2001; Riley, 2001; Lange, 2007; Schönberg, 2007; Lang and Manove, 2011; Roy, 1951; Sattinger, 1975; Fredriksson, Hensvik and Skans, 2018). According to job market signaling theory, employers use completed schooling as a signal of labor market productivity to screen workers (Spence, 1973). However, as the graduate workforce has grown, these credentials constitute very crude signals and mask valuable information about the applicant's ability.[2] Consequently, employers often face a choice between applicants with similar levels of educational attainment (e.g., a university degree) and may therefore look for other signals

[2]Between 2000 and 2016, the proportion of the population 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 the UK in 2016, more than 50 percent of the population aged 25 to 34 had completed a tertiary education program.

of productivity, such as information on educational achievement such as the applicant's GPA.[3] While most existing research on signaling and sorting focuses on educational attainment, only few studies examine the signaling value of educational credentials among students with similar degrees. One group of studies applies experimental curriculum vitae (CV) designs and demonstrates that variation in signals at the intensive margin (i.e., how well a person completed their degree in terms of their GPA) is related to the likelihood of being invited to a job interview (Koedel and Tyhurst, 2012; Protsch and Solga, 2015; Piopiunik et al., 2020). A second group of studies documents returns to receiving an honors degree classification (Freier, Schumann and Siedler, 2015; Feng and Graetz, 2017; Khoo and Ost, 2018),[4] while a third group of studies exploits discontinuity around the passing cutoff for specific degrees (Tyler, Murnane and Willett, 2000; Clark and Martorell, 2014; Jepsen, Mueser and Troske, 2016; Graetz, 2021). Our study contributes to the literature by using a novel source of identifying variation to document the effect of the GPA on actual labor market outcomes over time for the population of students across university degree programs. Our findings reinforce the existing evidence that educational credentials are used to sort workers in the labor market and that employers learn rather quickly about actual productivity (Lange, 2007; Aryal, Bhuller and Lange, forthcoming).

2 Background

2.1 The 2007 Danish grading reform

On August 1, 2007, the Danish Government replaced the old '13 scale' with a 7-point grading scale in all educational programs from primary school to university. This was to harmonize grading across educational programs and to ease comparison of Danish grades with international grading systems. For students who had already graduated, the Ministry of Education provided a crosswalk from the old to the new GPA. However, for students already enrolled on degree programs on August 1, 2007, their prior degree program grades were recoded based on the scheme in Table 1. This recoding created the variation in GPAs we exploit in this study.

The first two columns of Table 1 show the mapping from the old 13 scale to the new 7-point scale. There are two sources of variation in the recoding of the grades. First, as

[3] Signals of educational achievement such as a GPA are common in many countries. For example, US colleges use letter grades (A through F), which are converted to a numerical GPA, whereas UK universities assign scores on a 100-point scale which are translated into a degree classification (e.g., first-class honors).

[4] An alternative signal of productivity is university prestige (Bostwick, 2016; MacLeod et al., 2017; Bordón and Braga, 2020)

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: The European Credit Transfer System (ECTS) is the grading system defined by the European Commission. 6 (old) / 2 (new) is the passing threshold.

the new scale has fewer grades (seven compared to ten), three pairs of grades in the old scheme were collapsed into single new grades. For example, a student who only had grades of 8 on the old scheme and another student who only had grades of 9 on the old scale would have identical GPAs after the recoding. Second, the distances between the old and the new grades vary.

Although most students were downgraded in absolute terms, two students with identical pre-reform GPAs could have very different post-reform GPAs because grades were recoded differently. Consider a student with grades 8 and 10 on the 13 scale and another student with two 9s. They both have a GPA of exactly 9.0. The recoding converts the 8 to a 7 and the 10 is unchanged, leading to a GPA of 8.5. The second student's 9s are transformed to a 7, leading to a GPA of 7.0. After the recoding, there is a 1.5 unit difference in the GPA across these two students with identical pre-recoding GPAs.

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 over the age of 18 receive a monthly stipend to pay for their living costs. Enrollment in university programs depends almost exclusively on high school GPA.

Denmark has adopted a three-year bachelor's and a two-year master's structure for most of its university programs.[5] As we focus on the importance of GPAs for labor market outcomes, we only consider graduates from master's programs. University modules are given an ECTS weight according to their overall workload, and students are expected to earn 60 ECTS points in each year. A year is typically split into terms of 14–15 weeks (some programs have four terms of eight weeks), and most programs end with a dissertation.

Students who had their pre-reform grades recoded 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. Appendix Figure A.2 shows two examples of diplomas for treated individuals. They illustrate what the employers observe when making hiring and wage decisions. The diplomas show the student's GPA based on the recoded grades and provide information about individual grades both before and after the recoding. It is not possible to reconvert the new grades to the old scale (as the new scale has fewer grades). Although it would require a significant effort, employers could, in principle, calculate each candidate's pre-recoding GPA or study their grades one by one. If such behavior is widespread, it would go against finding a labor market effect of the reform-induced variation in GPA.

3 Data

3.1 Data sources & sample selection

We consider all students who were enrolled in a master's program on August 1, 2007 at Aarhus University (including Aarhus School of Business) or at the University of Copenhagen. Students will be at different stages of the programs, as illustrated in Figure 1. As the treatment (i.e., the reform-induced variation in GPA) is caused by the recoding of their grades given up to August 1, 2007, we narrow our sample to those who are at the end of

[5] Most programs are five-year programs in practice. More than 90 percent of the bachelor's graduates progress to a master's program within two years of completing their bachelor's program (see Appendix Figure A.1). Some programs (e.g., medical school) are six-year programs.

their studies (i.e., the upper row of Figure 1). Specifically, we restrict the sample to the students who had at most 40 ECTS points remaining on the day the grades were recoded. [6] The 40 ECTS criterion is selected based on the fact that ECTS credits assigned to the dissertation vary between 30 and 60 in our sample.

To be able to follow the students in their first years on the labor market, we only include students who graduated before 2011. This sample restriction only requires students to have completed coursework that should take considerably less than a year within three years. As Appendix Figure A.3 shows, 23 percent of the students had graduated by the end of 2007 and 94 percent by the end of 2009. We show that our conclusions are not sensitive to changing the sample selection criteria. The final sample consists of 3,813 students.

We merge the student records with administrative registers from Statistics Denmark using the unique personal identifiers. The registers provide individual background information (i.e., age, gender, high school GPA, parental income, and education) and information about labor market outcomes.

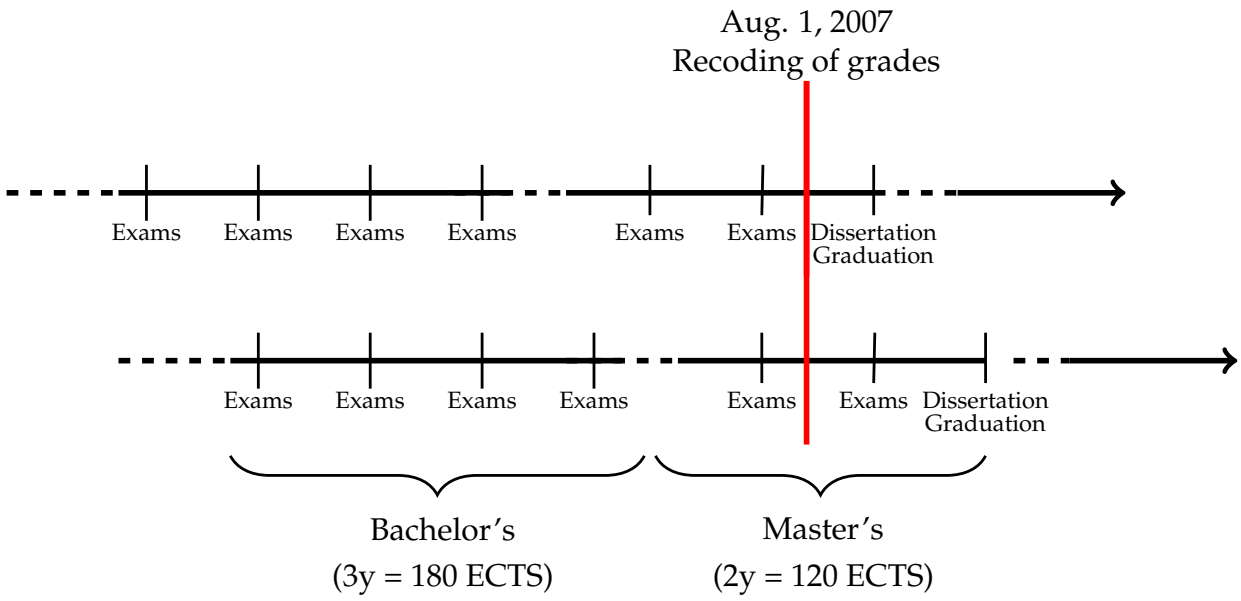


Figure 1: University students' exposure to the implementation of the new grading scheme

[6] Ideally, we would like to select students who were only waiting for their dissertation results. However, as university studies are very flexible in Denmark (meaning that students might complete some units after their dissertation), and because the credit load of the dissertation varies across years and programs, we cannot impose such a criterion. As we demonstrate, our results are robust to different bandwidths.

3.2 Variables

From the student records, we construct the final GPA, the GPA before the recoding, and the GPA after the recoding. We further record the number of credit points remaining at the time of the reform, the program studied, and the date of graduation.

Based on the register data, we create a set of student background variables including gender, age at the time of graduation, and high school GPA. For the students' parents, we generate variables on employment, total 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 the value is missing for an individual, we set the value to zero and include a dummy variable that equals one for all observations that have missing values.

Our main labor market outcome is log total gross earnings in each of the first five calendar years after graduation. In the model, individuals with zero earnings are excluded. However, to study whether the reform-induced variation in GPA affects the extensive margin, we separately estimate models using an indicator variable that equals one if the individual has positive earnings and zero otherwise.

3.3 Descriptive statistics

Table 2 shows the summary statistics for selected variables. The average age of the graduates is almost 31 years.[7] In our sample, 65 percent of the students are female and 26 percent of the sample have parents with a university degree. Forty-one percent are students from the University of Copenhagen. On average, students were close to graduation (23 ECTS remaining), and the grades given before the recoding accounted for about 70 percent of the overall GPA.

Ninety percent have positive earnings in the year after graduation. A university graduate in our sample earns on average 48 thousand USD (2015 level, gross) in the first calendar year after graduation, with 66 percent of them working in the public sector.

[7] In Appendix Figure A.4 we show that the median age at graduation is 29.1 years, and the median graduate spent 2.3 years on the labor market before they graduate with their masters degree, and 9.5 years in post compulsory schooling (which is scheduled to take 8 years). The unexplained 1.4 years might be spend on activities like travel, military service, and other time outside the labor force

Table 2: Summary statistics

	Mean	SD	P10	P50	P90
A. Background					
Age at graduation (years)	30.87	5.82	27.51	29.08	32.12
Female	0.65	0.48	0.00	1.00	1.00
Parental disposable income (1000 euros)	39.41	39.78	26.56	34.30	43.05
Parents with university degree	0.26	0.44	0.00	0.00	1.00
B. University status					
University of Copenhagen	0.41	0.49	0.00	0.00	1.00
ECTS remaining	23.21	13.38	14.00	30.00	30.00
Share of ECTS pre recoding	0.70	0.12	0.67	0.71	0.75
GPA	8.05	1.66	7.00	8.10	9.22
C. Labor-market status in year 1 after graduation					
Earnings > 0	0.90	0.30	1.00	1.00	1.00
Unemployment	0.08	0.17	0.00	0.00	0.05
Gross earnings (1000 euros)	42.88	21.25	31.90	45.28	56.59
Public sector	0.66	0.47	0.00	1.00	1.00
Observations	3813				

Notes: P10, P50, and P90 refer to the 10th pseudo-percentile, the 50th pseudo-percentile, and the 90th pseudo-percentile, respectively. Pseudo-percentiles are created by the average across the actual percentile and the two values above and below the percentile. Parental income is the average across observed parents, measured in the calendar year before the focal individual graduates from university. Parents with a university degree is an indicator variable that takes the value of one if at least one parent completed a university degree. All monetary values are adjusted to the 2015 price level using the consumer price index. GPA is the Grade Point average on the new scale.

4 Identification and Estimation

4.1 Empirical strategy

In an ideal experiment, we would randomly assign different GPAs on graduates' diplomas and follow their labor market trajectory. In our empirical strategy, we exploit the grading reform's creation of a setting that very closely resembles this. The horizontal axis in Figure 2 shows the GPA before the recoding, and the vertical axis shows the GPA after the recoding. Consider the students who had a GPA of 8.0 before the reform. Among these students, some had a GPA of 5.3 and some a GPA of 7.1 after the recoding. This 1.8 unit difference for students with identical pre-reform GPAs is larger than the standard deviation of 1.66 on the final GPA (see Table 2). To test whether the GPA has a signaling value we compare the labor market outcomes of these students.

As Figure 2 shows, a cubic relationship between pre- and post-recoding GPA fits the variation well. In our main strategy, the deviation from this fitted line constitutes the

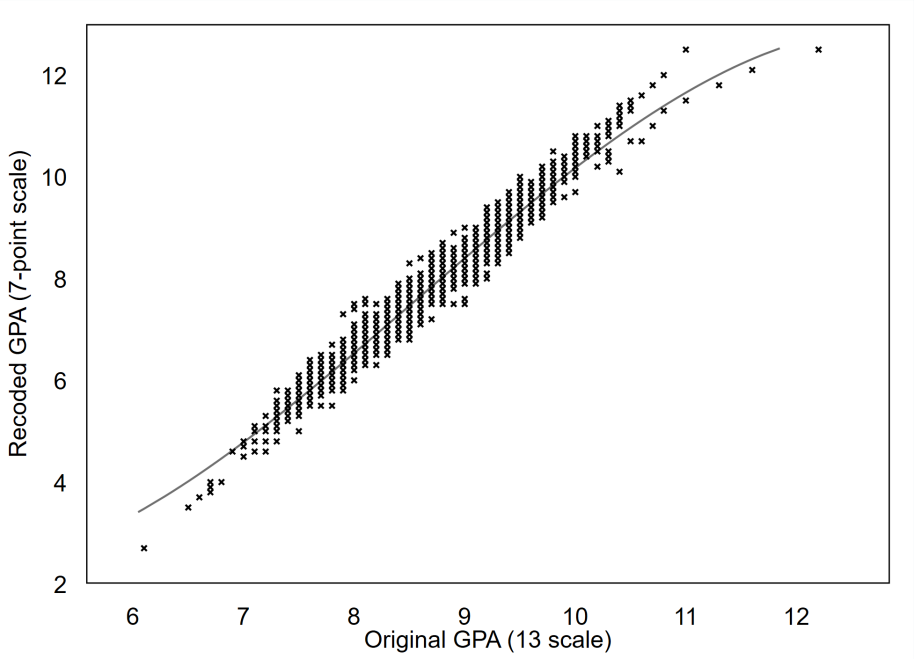


Figure 2: Pre- and post-recoding GPA

Notes: Each cross represents a combination of pre- and post-recoding GPA. Only grade combinations with at least three observations are shown.

treatment variable. We implement our empirical strategy by estimating the following equation using ordinary least squares (OLS):

$$Y_{it} = \beta_0 + \beta_1 GPA7_i + f(GPA13_i) + \boldsymbol{\lambda}'\boldsymbol{X}_i + e_{it} \tag{1}$$

where Y_{it} is log earnings in year t after graduation for individual i with a post-recoding GPA, $GPA7_i$. In our main specification, we control for pre-recoding GPA, $GPA13_i$, using a third-order polynomial, but we show that the results are not sensitive to changing the polynomial degree or using a less parametric approach.[8] As Table 2 shows that 70 percent of the grades are affected by the recoding, β_1 in (1) captures the reduced form estimate of both the signaling value of the GPA and potential behavioral responses by the students on labor market outcomes. As we discuss in the following subsection, several direct tests

[8]In Appendix B, we provide a Monte Carlo simulation of our empirical setting. Figure B.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 using a 5-percent cutoff. The linear approximation works poorly in the upper and lower end of the GPA distribution. Both the second- and third-order polynomial specifications lead to rejection rates of the expected 5 percent. 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.

suggest that there is no behavioral response.

Our strategy does not rely on any covariates.[9] However, to reduce the residual variance in the outcome variables and obtain more precise estimates, we include a range of controls in the vector $\boldsymbol{X_i}$. First, we expect earnings to be related to the program studied and the institution. We, therefore, include indicators for the program studied and for institution (University of Copenhagen or Aarhus University). Second, earnings may be related to individual characteristics and background. We, therefore, control for age, parental origin (non-Western or Western according to the definition by Statistics Denmark), parental income, parental unemployment, parental education, and gender. Parental variables are the mean across observed parents (except schooling, which equals one if at least one parent has completed a university degree). We include indicators for the number of parents with non-missing data (i.e., 0, 1, or 2). Third, we control for high school GPA. The error term e_i includes all other factors affecting the earnings, which could be both other signals of labor market productivity or factors directly related to productivity (e.g., cognitive and non-cognitive skills). To allow for arbitrary correlation within pre-recoding GPA cells, we cluster the standard errors on the pre-recoding GPA level as reported on the diploma with one decimal point.

4.2 Identifying assumptions

Identifying assumption 1 The causal interpretation of β_1 requires the variation in recoded GPA, given pre-recoding GPA, to be unrelated to individual characteristics that affect earnings. Such a correlation could arise for two reasons. First, the recoding algorithm in itself may capture characteristics that are valued on the labor market. To assess whether this is the case, we conduct placebo tests where we implement the recoding scheme on three non-treated cohorts. We apply the same sample selection on these cohorts and recode their GPA with exactly the same algorithm, and estimate equation (1) using the same outcome variables.

Second, the variation in GPA induced by the recoding might be correlated with individual characteristics. For example, if more forward-looking students anticipate the reform and either advance their studies (if they would be punished by the reform) or delay (if they would gain) their studies relative to the reform implementation date.[10] To assess whether individuals who benefited from the recoding are systematically different to those who suffered, we show that the reform-induced variation in GPA is unrelated to

[9] Estimating models without covariates give very similar results, as we show in Section 5.2.

[10] While the reform was announced in advance, we observe relatively little Google search activity before the actual implementation, as shown in Appendix Figure A.5.

observable individual characteristics, such as parental background, high school GPA, and undergraduate GPA.

Identifying assumption 2 Another potential explanation for a link between the recoded GPA and earnings could be that students reacted to the recoding by adjusting their study effort or selecting different elective units after the recoding. Importantly, if students adjust their study effort to compensate for the change in GPA—as shown for high school students in Hvidman and Sievertsen (2021)—it would go against finding labor market returns to a positive reform-induced GPA. In addition, the reform-induced variation could affect the students' likelihood of graduation, time to graduation, and course selection after the recoding. We test for these post-reform responses in terms of likelihood of graduation, time to graduation, study effort, and choice of elective units. Moreover, we show an instrumental variables specification in which we instrument the final GPA with the reform-induced variation in GPA. In absence of any behavioral responses, the first-stage coefficient should be 0.7, corresponding to the share of grades affected by the recoding. Two features, however, suggest that such behavior might be less pronounced in universities than in high schools. First, for high school students, the GPA is particularly high stakes as it determines access to higher education. Second, as we restrict our sample to university students close to graduation at the time of the reform, they have limited time to react to the recoding.

5 Results

5.1 Returns to the reform-induced variation in GPA

Panel A of Table 3 shows the results from regressing log earnings on the final GPA, conditional on the full set of covariates. A one-unit increase in the final GPA is associated with 2.3 percent higher earnings in the first year after graduation (corresponding to a one-standard-deviation increase in GPA being associated with $0.23 \times 1.66 = 3.8$ percent higher earnings). This relationship is relatively constant across the first five years after graduation.

Panel B of Table 3 shows that a one-unit increase in the recoded GPA is related to almost 8.7 percent higher earnings in the first year after graduation, and almost 8 percent higher earnings in the second year. The effect decreases slightly from year one to year two, and the coefficients are both smaller and not statistically different from zero in years three to five after graduation. Even though the decrease from year one to year two looks

considerably smaller than the the decrease from year two to year three, the wage adjustments are not statistically different, suggesting that employers learn and adjust wages over the first two to three years, whereas the effect has disappeared from years three to five.

Panel C of Table 3 confirms a first-stage coefficient of 0.7 corresponding to the share of grades affected by the recoding and indicating no average behavioral response. If students had compensated for the recoding by increasing their effort in response to a negative GPA recoding, the coefficient would have been smaller than 0.7. Finally, panel D of Table 3 shows that a one-unit increase in the final GPA induced by the reform causes 12 percent higher earnings in year one and 11 percent higher earnings in year two. While we would expect the "raw" specification to be downward biased if the GPA is positively correlated with unobserved characteristics that are also positively linked to earnings, this conjecture is less obvious at labor market entry where the wage setting very much depend on observable characteristics. Moreover, although the IV coefficients in panel D of Table 3 are larger than the raw estimates in panel A, they are also less precise. Thus, both the IV and the raw specification includes returns of around 0.045 in the confidence interval.

The results in Table 3 suggest that employers learn about the noise in the signal in about two to three years, which is slightly faster than the five years found in Lange (2007). Appendix Table A.2 shows results from estimating a less flexible model of employer learning in line with Farber and Gibbons (1996); Altonji and Pierret (2001). We create a balanced pooled sample of the first five years after graduation and regress log earnings on the pre-reform GPA and the post-reform GPA, as well as both terms interacted with experience. Column (5) of Appendix Table A.2 corresponds to specification in column (2) of Table 1 in Altonji and Pierret (2001). The coefficient on the recoded GPA in column is similar to the results in Table 3: a unit higher GPA is associated with seven percent higher earnings in the first year. As expected, the coefficient on the interaction between the recoded GPA and experience is negative (corresponding to the interaction of education and experience in Altonji and Pierret (2001)), suggesting that the return to observable signals decreases over time. The coefficient on the pre-reform GPA, which is hard to observe for employers, is close to zero in year 0, but the interaction with experience (corresponding to the interaction between AFQT and experience in Altonji and Pierret (2001)) is positive. Using the coefficients from this specification, the results suggest that the impact of the recoded GPA, covering the noisy signal, is zero after (0.07/0.017) -1 =3.12 years, which is in line with the more flexible specification in Table 3.

Table 3: Regression results - The effect of grades on earnings in years 1 to 5 after graduation.

	Year after graduation				
	1	2	3	4	5
A. Raw specification - Dependent variable: log earnings					
Final GPA	0.023*	0.017*	0.022**	0.022***	0.025**
	(0.011)	(0.009)	(0.009)	(0.008)	(0.010)
B. Reduced form - Dependent variable: log earnings					
Recoded GPA	0.087***	0.078**	0.006	0.002	-0.019
	(0.030)	(0.033)	(0.026)	(0.027)	(0.038)
C. First stage - Dependent variable: final GPA					
Recoded GPA	0.713***	0.711***	0.715***	0.713***	0.712***
	(0.048)	(0.048)	(0.048)	(0.047)	(0.050)
D. IV specification - Dependent variable: log earnings					
Final GPA	0.122***	0.110**	0.008	0.003	-0.026
	(0.039)	(0.043)	(0.035)	(0.038)	(0.052)
Observations	3445	3465	3423	3388	3366
F-stat	221.21	219.83	226.27	232.60	201.75

Notes: Final GPA is the grade point average for the graduates after the recoding and including post-recoding assessment results. Final GPA is instrumented using the recoded grade point average as the instrumental variable, controlling for pre-recoding GPA using a third-order polynomial. All specifications are estimated with program fixed effects and the full set of covariates, which includes an indicator for institution (University of Copenhagen or Aarhus University), age, an indicator for parental origin (non-Western or Western, according to the definition by Statistics Denmark), parental income, parental unemployment, parental education (indicator for university degree), gender, and high school GPA. Parental variables are created as the mean across observed parents (except schooling, which is one if at least one parent has completed a university degree). We include indicators for the number of parents with 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. Standard errors clustered at the pre-recoding GPA level are in parentheses. The F-stat is for the excluded instrument in the first stage specification. Asterisks indicate significance at the following levels: * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

5.2 Validity of the research design

A. Placebo tests: Does the recoding algorithm predict labor market outcomes for non-treated cohorts? To assess whether the recoding algorithm captures aspects that are rewarded in the labor market, we conduct placebo tests on three cohorts that were unaffected by the reform. We implement the same sample selection criteria and the same recoding algorithm for cohorts in July 2002, 2003, and 2004 and conduct exactly the same analysis as shown in Table 3. We consider all students who had 40 ECTS remaining at most on July 31 in each of these years. We then recode all the grades given up to that point and define this GPA as *GPA7*, and estimate equation (1). Table 4 shows the results.

We find no detectable association between the recoded GPA and earnings in any of the years for any of the three cohorts.[11] This supplementary analysis suggests that the results in Table 3 do not simply reflect the algorithm capturing grade combinations that are appreciated on the labor market.

Table 4: Reduced-form regression results on placebo cohorts. Dependent variable: Log earnings in years one to five after graduation.

	Year after graduation				
	1	2	3	4	5
Recoded GPA	0.013	-0.010	-0.008	-0.001	-0.004
	(0.046)	(0.036)	(0.028)	(0.035)	(0.030)
Observations	3029	3095	3098	3102	3074
R-squared	0.16	0.17	0.17	0.15	0.17

Notes: This table resembles Table 3, but shows estimates based on pooling three placebo samples. The placebo samples are created by implementing a placebo recoding of grades on respectively July 31, 2002, July 31, 2003, and July 31, 2004 using the same recoding scheme, covariates, and sample selection as in the main specification. See notes for Table 3.

B. Covariate balance: Is the recoding related to background characteristics? To assess whether the reform-induced GPA variation is related to individual background, we estimate equation (1) using covariates as the dependent variables. Column (1) of Table 5 shows that the reform-induced variation in GPA is unrelated to an individual's 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 statistically 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 5 suggest that the reform-induced variation in GPA is not related to any observable characteristics.

[11] The placebo cohorts that we analyze are smaller than the treated cohort. The birth cohorts of the placebo cohorts are smaller. Furthermore, the student records from Aarhus University start in 2000. As some students spent more than two years completing their master's program, fewer students are included in the placebo analysis (particularly for the 2002 cohort) than in the main analysis. Appendix Table A.3 shows that estimating placebo specifications separately for each cohort confirms the conclusion from the pooled regression.

Table 5: Reduced-form regression results of covariate balance. Dependent variables in column header.

	Female	HS GPA	BSc/BA GPA	— Parents' — Income	Unempl.	Uni. degr.	Predict. earn.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Recoded GPA	-0.018	-0.001	0.048	2.766	-0.004	0.020	-0.000
	(0.017)	(0.031)	(0.053)	(2.112)	(0.004)	(0.020)	(0.004)
Observations	3811	3218	2046	3363	3813	3322	3813
R-squared	0.14	0.31	0.31	0.04	0.02	0.07	0.94
Mean dep. var	0.65	0.75	-0.00	39.41	0.02	0.26	3.62

Notes: The table shows the coefficients from estimating equation (1) using the variables denoted in the column headers as dependent variables. Parental variables are measured in the calendar year before graduation. Parental income is the average disposable income across the observed parents, measured in 1000 EUR (2015 level). Unemployment is the average annual unemployment of the observed parents. University degree is an indicator for whether at least one parent has completed a university degree. All models are estimated without covariates but with program fixed effects. Predicted earnings is the predicted earnings based on coefficients from a regression of log earnings in the first year after graduation on all covariates. Standard errors clustered at the pre-recoding GPA level are in parentheses. Asterisks indicate significance at the following levels: * p< 0.1, ** p < 0.05, and *** p< 0.01.

C. Did students change their behavior in response to the reform? We also test whether the recoding of grades affected subsequent student behavior after the recoding. We test whether the recoding affected students' likelihood of graduation, their performance in subsequent exams, selection of different elective units, and their time to graduation. Table 6 shows the results of these estimations.

We find no evidence of any behavioral responses. Columns (1) and (2) show that the reform-induced variation in GPA does not affect the likelihood of graduation, nor the time to graduation. Columns (3) and (4) show that the reform-induced variation is unrelated to unit difficulty[12] and performance in subsequent assessments.

D. Are the results sensitive to model specification? To assess whether our findings are sensitive to the empirical specification, Figure 3 shows point estimates and confidence intervals for β_1 based on 72 different specifications using log earnings in the first calendar year after graduation as the dependent variable. The empirical specification is indicated by the markers below the chart. All 72 specifications lead to positive and significant point

[12]We measure unit difficulty as follows. First, we consider all individual exam results for the units in the six years prior to the reform (2000 to 2006). Second, we regress these individual exam results on unit fixed effects and students' high school GPA. A larger unit fixed effect suggests that given their high school GPA, a student receives a higher grade in that subject. 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 attended units after the recoding of the grades, and use the fixed effects as the dependent variable.

Table 6: Reduced-form regression results. Behavioral responses. Dependent variables in column headers.

	(1) Graduated	(2) Time to graduation	(3) Unit FE	(4) Post GPA
Recoded GPA	-0.013	-0.019	0.000	0.049
	(0.012)	(0.039)	(0.000)	(0.122)
Observations	4579	4048	3696	4048
R-squared	0.21	0.15	0.74	0.24
Mean dep. var	0.88	0.97	0.00	8.18

Notes: The table shows the coefficients from estimating equation (1) 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 the time from recoding to graduation, measured in years. Unit FE is the average unit-specific fixed effects of units completed after the recoding. The fixed effects are estimated 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 suggests 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. All models are estimated with the full set of covariates (see notes for Table 3). Standard errors clustered at the pre-recoding GPA level are in parentheses. Asterisks indicate significance at the following levels: * $p < 0.1$, ** $p < 0.05$, and *** $p< 0.01$.

estimates in the range 0.05 and 0.10. The chart also shows that our main specification (indicated with a blue marker) is not an outlier.

E. Is the relationship symmetric? So far, we have assumed 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 reform-induced change (relatively speaking) to their GPA have an incentive to inform employers about this, while job candidates who receive a positive shock have no incentive to do so. Thus, it could be the case that the relationship is driven by the positive shocks.

Figure 4 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.

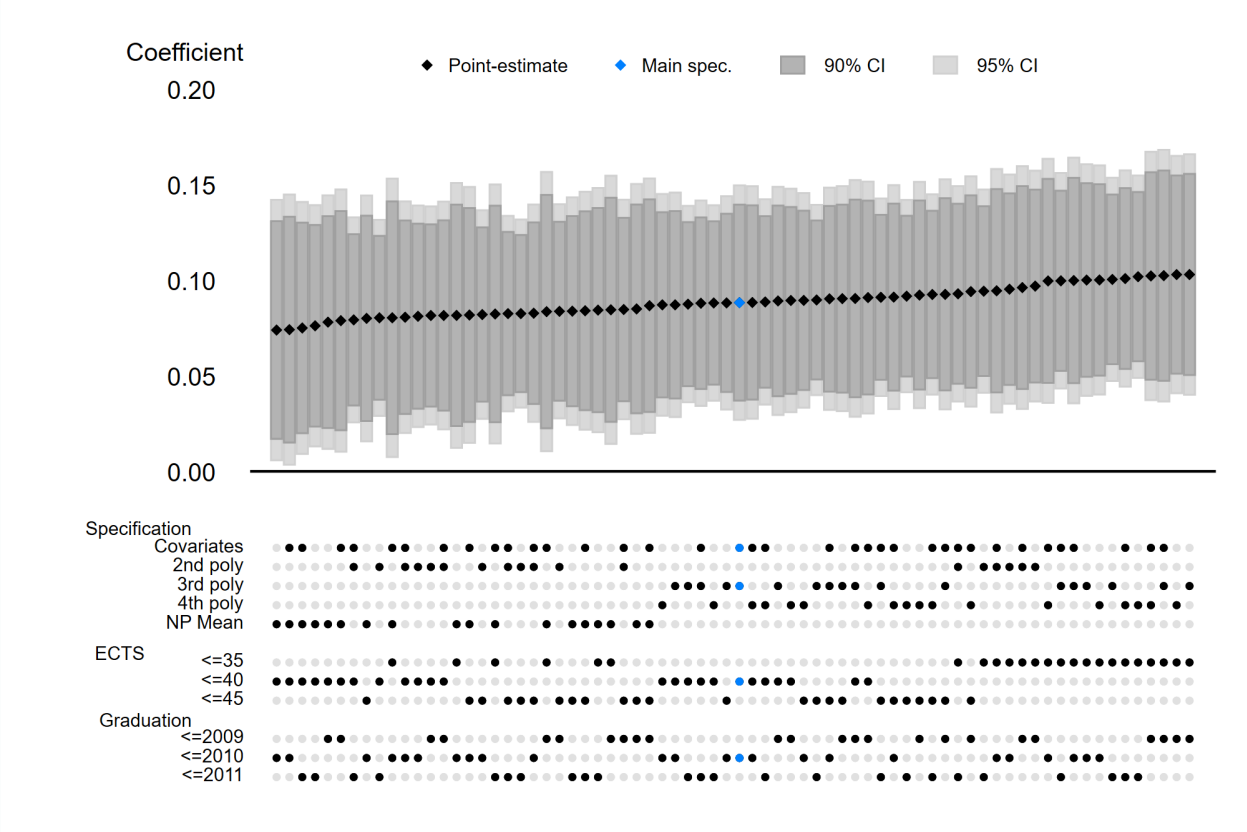


Figure 3: Specification curve.

Notes: The chart shows the reduced-form point estimates and confidence intervals using the specification indicated by the markers below the chart. The first row of markers indicates whether the specification is estimated with or without any control variables. Specifications with a black marker include the full set of controls. Specifications with a gray marker only include major fixed effects. The second to fourth rows of markers capture the functional form of $f(GPA13_i)$, where a black marker indicates that the specification named on the left is applied. The fifth row, NP Mean, shows results from a less parametric approach where we compare an individual's recoded GPA to the mean recoded GPA among all other students with the same pre-recoding GPA. This specification also includes pre-recoding GPA fixed effects. In the two lower panels, ECTS and Graduation, we alter the sample selection criteria. ECTS relates to the number of ECTS credit points remaining at the time of the reform, and graduation refers to the time limit for graduation that was imposed.

5.3 Mechanisms and heterogeneity

This section explores mechanisms that could drive the effect of the reform-induced variation in GPA and subgroup analyses.

Mechanisms: In Table 7, we investigate the effect of the reform-induced variation in GPA on alternative labor market outcomes over the first five years after graduation. We estimate equation (1) with different labor market outcomes as the dependent variable.

The first row of Table 7 shows there is no effect on the extensive margin of earnings.

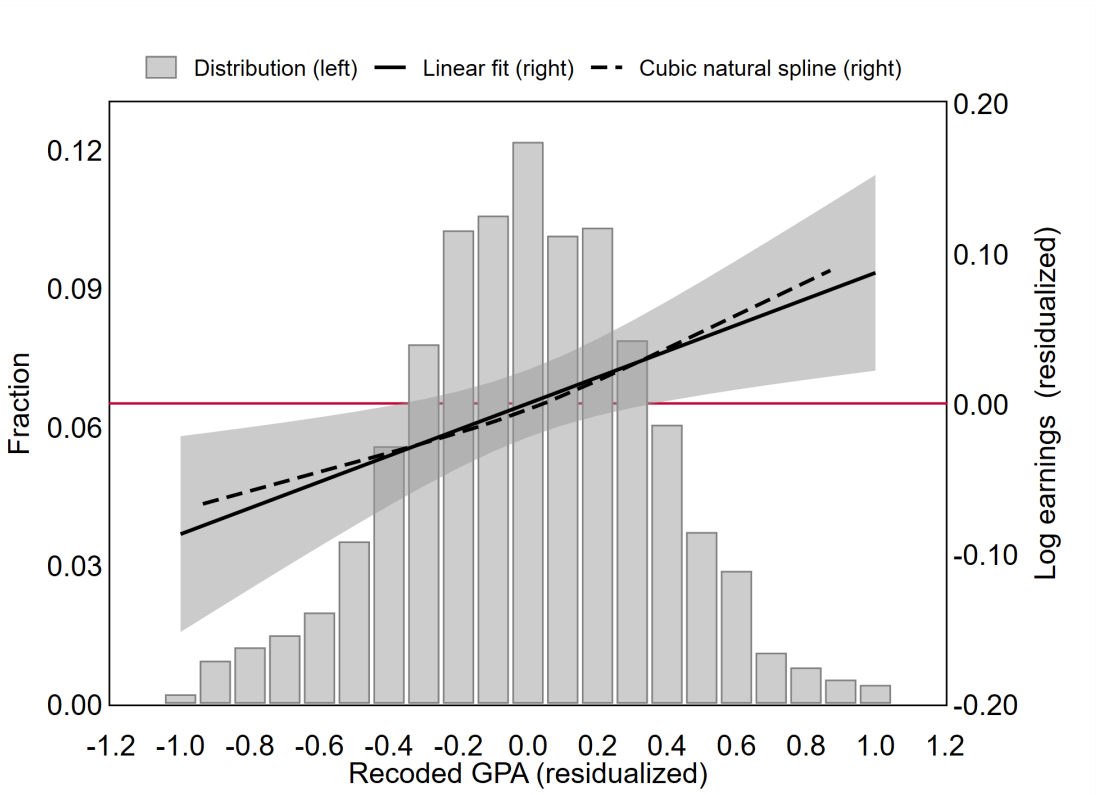


Figure 4: The relationship between log earnings in the calendar year after graduation and reform-induced variation in GPA.

Notes: The solid line shows the linear relationship estimated in our main specification. The shaded area shows the 95 percent confidence interval. 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.

The second and third rows show the effects on log disposable income (after transfers and taxes). Without conditioning on actually having earnings (second row), the point estimates in years one and two are somewhat noisy, but after conditioning on positive earnings (third row), we find that a one-unit increase in reform-induced GPA leads to an increase in disposable income of respectively 4.6 percent and 3.7 percent in the first two calendar years after graduation.

The results in rows four to seven suggest that there is no link between reform-induced GPA and unemployment, public sector employment, or job changes within the first five years after graduation. Finally, row eight shows that a higher reform-induced GPA leads to lower earnings growth in year three after graduation, and as row nine shows, this is also evident for employees who stay with the same employer.

Overall, the results in Table 7 show that the reform-induced variation in GPA is unre-

lated to the extensive margin adjustments: there is no effect on labor market participation, sector, or job changes. However, we see evidence of earnings adjustments in years two and three, and this also occurs within the firm.

Table 7: Reduced-form regression results. Other labor market outcomes. Dependent variables in row titles.

	Year after graduation				
	1	2	3	4	5
Earnings > 0	-0.005	0.008	0.002	-0.008	-0.005
	(0.012)	(0.013)	(0.013)	(0.011)	(0.013)
Log disp. income	0.052	0.082***	-0.009	-0.002	0.005
	(0.039)	(0.031)	(0.029)	(0.022)	(0.025)
Log disp. income | earnings > 0	0.046***	0.037**	-0.006	-0.012	-0.010
	(0.015)	(0.014)	(0.015)	(0.015)	(0.022)
Unemployment	-0.006	-0.005	-0.006	0.000	0.012
	(0.007)	(0.006)	(0.005)	(0.004)	(0.007)
Public sector	0.008	0.001	-0.006	0.005	0.035
	(0.017)	(0.018)	(0.019)	(0.023)	(0.046)
Job change		-0.019	0.020	-0.016	0.015
		(0.016)	(0.020)	(0.021)	(0.020)
Job change with earnings growth		0.006	-0.061	0.003	0.052
		(0.047)	(0.047)	(0.053)	(0.071)
Earnings growth year		-0.006	-0.069***	-0.022	-0.015
		(0.030)	(0.025)	(0.020)	(0.031)
Earnings growth | same employer		0.032	-0.046*	-0.028*	-0.002
		(0.027)	(0.024)	(0.017)	(0.022)

Notes: The table shows the coefficients from estimating equation (1) using the variables denoted in the first column as dependent variables. All models are estimated with the full set of covariates (see notes for Table 3). Standard errors clustered at the pre-recoding GPA level are in parentheses. Asterisks indicate significance at the following levels: * $p < 0.1$, ** $p < 0.05$, and *** $p< 0.01$.

Heterogeneity We consider two sources of heterogeneity: the potential for individual wage setting and the importance of educational credentials as signals.

The reform-induced GPA could play a bigger role in less regulated labor markets with greater earnings dispersion. Panels A to C of Table 8 show point estimates for subsamples that vary in their structure. First, the earnings distribution for women tends to be more compressed than for men. We find the largest effects for men, as shown in panel A. Furthermore, based on cohorts graduating from 2002 to 2006, we split the sample according to the pre-reform wage dispersion (panel B) and public employment share (panel C)

of the university programs. We find larger effects for majors with higher wage dispersion and lower public employment shares. Nevertheless, splitting the samples leads to small sample sizes, and none of the sub-group differences are statistically significant.

The second set of heterogeneity analyses investigates whether students with a strong network or extensive labor market experience are less dependent on skill signals (e.g., GPA). First, we split the sample by parental education. Graduates with better networks on the relevant job market—as proxied by parental education— could be less reliant on the GPA as a signal. In line with this hypothesis, we find stronger effects for children of parents without a degree (panel D). Relevant labor market experience might also lower the importance of educational credentials in the job search. Danish university students often have part-time jobs that are relevant to their studies and later careers. Panels E and F show that effects are stronger for graduates with low earnings while studying (a proxy for little relevant labor market experience) and for graduates from Aarhus University. Because the student labor market is considerably larger in Copenhagen, the findings could suggest that graduates rely less on the GPA as a signal to employers if they have had better access to the labor market during their studies.[13]

[13]There is no clear institution level reputation difference between Aarhus University and the University of Copenhagen. To test this empirically, we estimated a regression of log earnings on GPA, major fixed effects and institution fixed effect. The coefficient on the institution fixed effect is not statistically significantly differently from zero.

Table 8: Reduced-form regression results for subgroups. Dependent variable: log earnings year one to five after graduation.

	Year after graduation				
	1	2	3	4	5
A. By gender					
Female = 0	0.154**	0.108	0.031	0.014	-0.067
	(0.057)	(0.061)	(0.054)	(0.050)	(0.065)
Female = 1	0.056	0.051	-0.015	-0.015	0.004
	(0.047)	(0.043)	(0.030)	(0.032)	(0.036)
P-val	0.24	0.48	0.47	0.58	0.27
B. Major wage dispersion					
$<$ p(50)	0.049	0.060	-0.009	-0.063*	-0.075
	(0.038)	(0.032)	(0.030)	(0.029)	(0.044)
$>$ p(50)	0.130***	0.083	0.011	0.071	0.052
	(0.048)	(0.069)	(0.049)	(0.042)	(0.055)
P-val	0.16	0.77	0.73	0.00	0.06
C. Major public sector share					
<p(50)	0.115*	0.207**	0.050	0.039	0.049
	(0.056)	(0.071)	(0.057)	(0.056)	(0.071)
>p(50)	0.073*	0.018	-0.017	-0.016	-0.048
	(0.032)	(0.035)	(0.029)	(0.037)	(0.041)
P-val	0.49	0.02	0.29	0.43	0.23
D. Parents with university degree					
No	0.093*	0.111*	0.014	-0.001	-0.000
	(0.039)	(0.043)	(0.036)	(0.036)	(0.043)
Yes	0.028	-0.035	-0.024	-0.033	-0.106
	(0.045)	(0.057)	(0.045)	(0.050)	(0.057)
P-val	0.30	0.05	0.55	0.61	0.09
E. Earnings while studying					
$<$ p(50)	0.135*	0.074	-0.026	0.107*	-0.007
	(0.058)	(0.068)	(0.049)	(0.045)	(0.061)
$>$ p(50)	0.063	0.077**	0.019	-0.063*	0.027
	(0.038)	(0.029)	(0.031)	(0.030)	(0.045)
P-val	0.31	0.96	0.45	0.00	0.78
F. University					
Aarhus	0.137**	0.105	0.016	0.037	-0.041
	(0.048)	(0.054)	(0.039)	(0.037)	(0.051)
Copenhagen	0.032	0.039	-0.015	-0.050	0.010
	(0.034)	(0.042)	(0.035)	(0.041)	(0.048)
P-val	0.06	0.33	0.55	0.13	0.44

Notes: The coefficients shown are for the sub-group denoted on the left. The p-value rows show the p-value for the test of the null hypothesis that the estimate of β_1 is the same in both sub samples. See notes for Table 3).

6 Conclusion

Using variation in university students' GPAs that is unrelated to labor market productivity, we document a signaling value of university GPAs at labor market entry. Studying outcomes over time, we find evidence that employers rapidly learn about productivity and that earnings adjustments occur both within and across firms.

We find no effects on the likelihood of employment just after graduation or on job changes within the first five years on the job market. Additionally, we find that the signaling value of GPAs is strongest for men, and children of parents with no university degree. The latter result may suggest that signals are more relevant to workers with no informal connections to the labor market. Finally, our findings indicate the effect is strongest for majors that are related to larger wage dispersion and strongly connected to the private sector.

Our findings have important policy implications. First, it suggests that the grading system influences the labor market matching process. Grade inflation and bunching, for example, could make it harder for employers to identify the best applicant. Moreover, systems that focus on parts of the achievement distribution (e.g., through honors 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 skill signals. Several factors can affect assessments, including pollution (Ebenstein, Lavy and Roth, 2016), weather (Park et al., 2020), time of the day (Sievertsen, Gino and Piovesan, 2016), and teacher manipulation (Dee et al., 2019; Diamond and Persson, 2016). Our findings indicate that such external factors have implications for labor market outcomes.[14] Our finding of a large initial signaling value of educational credentials to some degree justifies the students' focus on grades. However, we find no evidence that the grades are important for securing a first job that brings the graduates on a different earnings trajectory that impossible to catch-up on. Third, Bar, Kadiyali and Zussman (2009) show that students select into elective courses that are more leniently marked. Our short-run results justify this behavior, but the rapid employer learning suggests that the benefits of prioritizing the signal (a higher grade) over human capital (selecting units based on content) might be temporary. Finally, while the setting that we study is unique, grading reforms are relatively widespread. Although the implementations might vary from reform to reform, they will typically generate some noise in the signaling process that could have important implications.

[14]Note that none of the studies above are in a university setting. However, there is substantial anecdotal evidence for errors in grading in higher education (see e.g., Nightingale (2017)).

More generally, our findings provide evidence of the importance of skill signals in the labor market and key insights to the employer learning process. Future research on signaling and employer learning based on educational achievement could provide an even deeper understanding of this learning process by exploiting alternative sources of variation in signals.

References

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

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

Aryal, Gaurab, Manudeep Bhuller, and Fabian Lange. forthcoming. "Signaling and Employer Learning with Instruments." *American Economic Review*.

Bar, Talia, Vrinda Kadiyali, and Asaf Zussman. 2009. "Grade information and grade inflation: The Cornell experiment." *Journal of Economic Perspectives*, 23(3): 93–108.

Bordón, Paola, and Breno Braga. 2020. "Employer learning, statistical discrimination and university prestige." *Economics of Education Review*, 77: 101995.

Bostwick, Valerie. 2016. "Signaling in higher education: The effect of access to elite colleges on choice of major." *Economic Inquiry*, 54(3): 1383–1401.

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

Dee, Thomas S, Will Dobbie, Brian A Jacob, and Jonah Rockoff. 2019. "The Causes and Consequences of Test Score Manipulation: Evidence from the New York Regents Examinations." *American Economic Journal: Applied Economics*, 11(3).

Diamond, Rebecca, and Petra Persson. 2016. "The long-term consequences of teacher discretion in grading of high-stakes tests." National Bureau of Economic Research Working Paper 22207.

Ebenstein, Avraham, Victor Lavy, and Sefi 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, Henry S, and Robert Gibbons. 1996. "Learning and wage dynamics." *The Quarterly Journal of Economics*, 111(4): 1007–1047.

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

Fredriksson, Peter, Lena Hensvik, and Oskar Nordstrom Skans. 2018. "Mismatch of Talent: Evidence on Match Quality, Entry Wages, and Job Mobility." *American Economic Review*.

Freier, Ronny, Mathias Schumann, and Thomas Siedler. 2015. "The earnings returns to graduating with honors—Evidence from law graduates." *Labour Economics*, 34: 39–50.

Graetz, Georg. 2021. "On the interpretation of diploma wage effects estimated by regression discontinuity designs." *Canadian Journal of Economics/Revue canadienne d'économique*.

Hvidman, Ulrik, and Hans Henrik Sievertsen. 2021. "High-Stakes Grades and Student Behavior." *Journal of Human Resources*, 56(3): 821–849.

Jepsen, Christopher, Peter Mueser, and Kenneth Troske. 2016. "Labor market returns to the GED using regression discontinuity analysis." *Journal of Political Economy*, 124(3): 621–649.

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

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

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

Lang, Kevin, and Michael Manove. 2011. "Education and Labor Market Discrimination." *American Economic Review*, 101(4): 1467–96.

MacLeod, W Bentley, Evan Riehl, Juan E Saavedra, and Miguel Urquiola. 2017. "The big sort: College reputation and labor market outcomes." *American Economic Journal: Applied Economics*, 9(3): 223–61.

Nightingale, Victoria. 2017. "University grading errors are more common than you might think." *Times Higher Education*.

Park, R Jisung, Joshua Goodman, Michael Hurwitz, and Jonathan Smith. 2020. "Heat and learning." *American Economic Journal: Economic Policy*, 12(2): 306–39.

Piopiunik, Marc, Guido Schwerdt, Lisa Simon, and Ludger Woessmann. 2020. "Skills, signals, and employability: An experimental investigation." *European Economic Review*, 123: 103374.

Protsch, Paula, and Heike 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, John G. 2001. "Silver signals: Twenty-five years of screening and signaling." *Journal of Economic literature*, 39(2): 432–478.

Roy, Andrew Donald. 1951. "Some thoughts on the distribution of earnings." *Oxford economic papers*, 3(2): 135–146.

Sattinger, Michael. 1975. "Comparative advantage and the distributions of earnings and abilities." *Econometrica*, 43(3): 455.

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

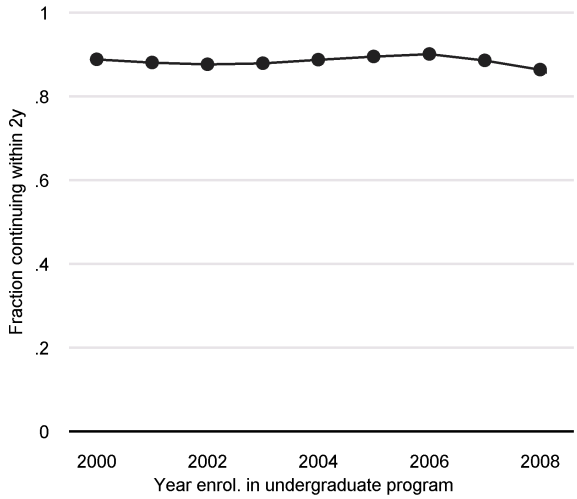
Sievertsen, Hans Henrik, Francesca Gino, and Marco Piovesan. 2016. "Cognitive fatigue influences students' performance on standardized tests." *Proceedings of the National Academy of Sciences*, 113(10): 2621–2624.

Spence, Michael. 1973. "Job market signaling." *The Quarterly Journal of Economics*, 355–374.

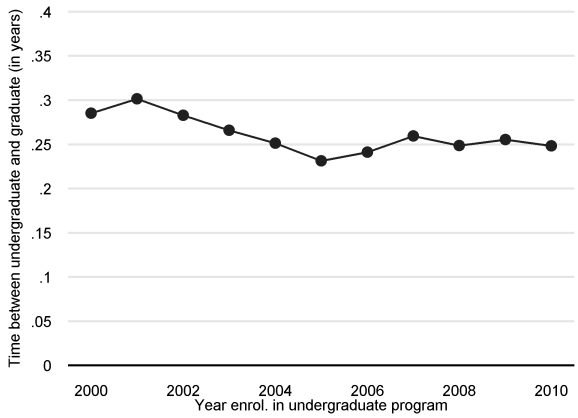
Tyler, John H, Richard J Murnane, and John B Willett. 2000. "Estimating the labor market signaling value of the GED." *The Quarterly Journal of Economics*, 115(2): 431–468.

Wise, David A. 1975. "Academic achievement and job performance." *The American Economic Review*, 350–366.

A Appendix



(a) Share continuing within 2y



(b) Time btw degrees

Figure A.1: Students enrolling in a master's program after their undergraduate degree.

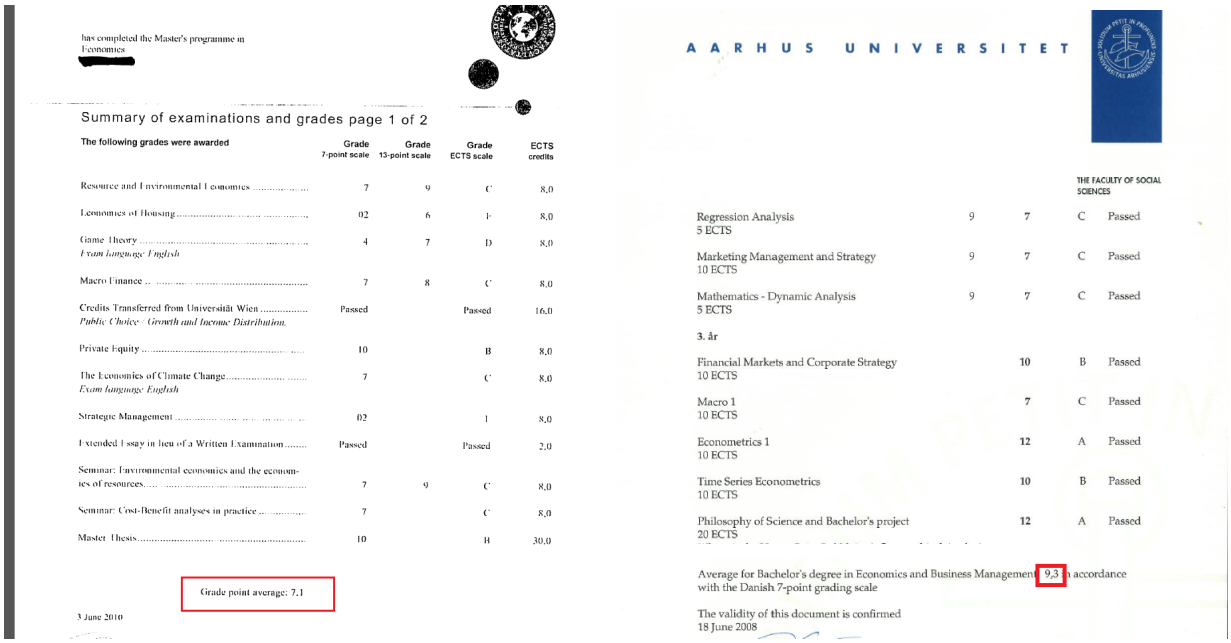
Table A.1: Educational background for the population aged 30-35, 2008-2012

	Earnings (1000 euros, 2015 level)						
	All	>0	0	0-20	20-40	40-60	>60
Compulsory educ.	9.98	7.34	28.48	16.08	8.96	4.23	2.2
Upper second. educ.	14.43	12.89	25.21	22.09	14.11	9.23	10.12
Vocational training	36.96	37.92	30.19	32.9	44.53	37.59	25.93
Short prof. programs	5.76	6.13	3.14	4.08	4.59	7.55	8.47
College programs	15.89	17.46	4.87	10.26	18.93	21.07	10.15
University	16.98	18.25	8.06	14.6	8.89	20.34	43.13
Total	100	100	100	100	100	100	100
Of university graduates:							
Aarhus University	3.81	4.14	1.5	2.93	2.15	4.67	9.7
University of Copenhagen	4.68	5.03	2.23	4.52	2.46	5.56	11.4
Total	8.49	9.17	3.73	7.45	4.61	10.23	21.1

Table A.2: Regression results. Employer Learning Specification. Dependent variable: Log earnings

	(1)	(2)	(3)	(4)	(5)
Recoded GPA	0.011**		0.002		0.070***
	(0.005)		(0.007)		(0.024)
GPA13		0.018*		-0.004	-0.126***
		(0.009)		(0.012)	(0.045)
t			0.024	-0.033	-0.169***
			(0.015)	(0.031)	(0.064)
Recoded GPA × t			0.005**		-0.017**
			(0.003)		(0.007)
GPA13 × t				0.011***	0.041***
				(0.004)	(0.013)
Observations	15,560	15,560	15,560	15,560	15,560
R-squared	0.17	0.17	0.19	0.19	0.19

Notes: The table shows results from estimating a regression with log earnings as the dependent variable, and the variables listed as the independent variables and controlling for major fixed effects on the pooled sample of earnings in years 1 to five after graduation. The sample is is restricted to be balanced such that all individuals are observed with non-missing earnings in all years (dropping 18% of the observations). Standard errors clustered on the major level (59 levels) in parentheses. Asterisks indicate significance at the following levels: * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.



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	F	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		F	8,0
Extended Essay in lieu of a Written Examination	Passed		Passed	2,0
Seminar: Environmental economics and the economics 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 A.2: Examples of diplomas for treated individuals

Notes: The diplomas show the student's GPA based on the recoded grades (marked with red) and provide information about individual grades both before and after the recoding. For exam grades that are recoded, 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 exams taking place after the recoding.

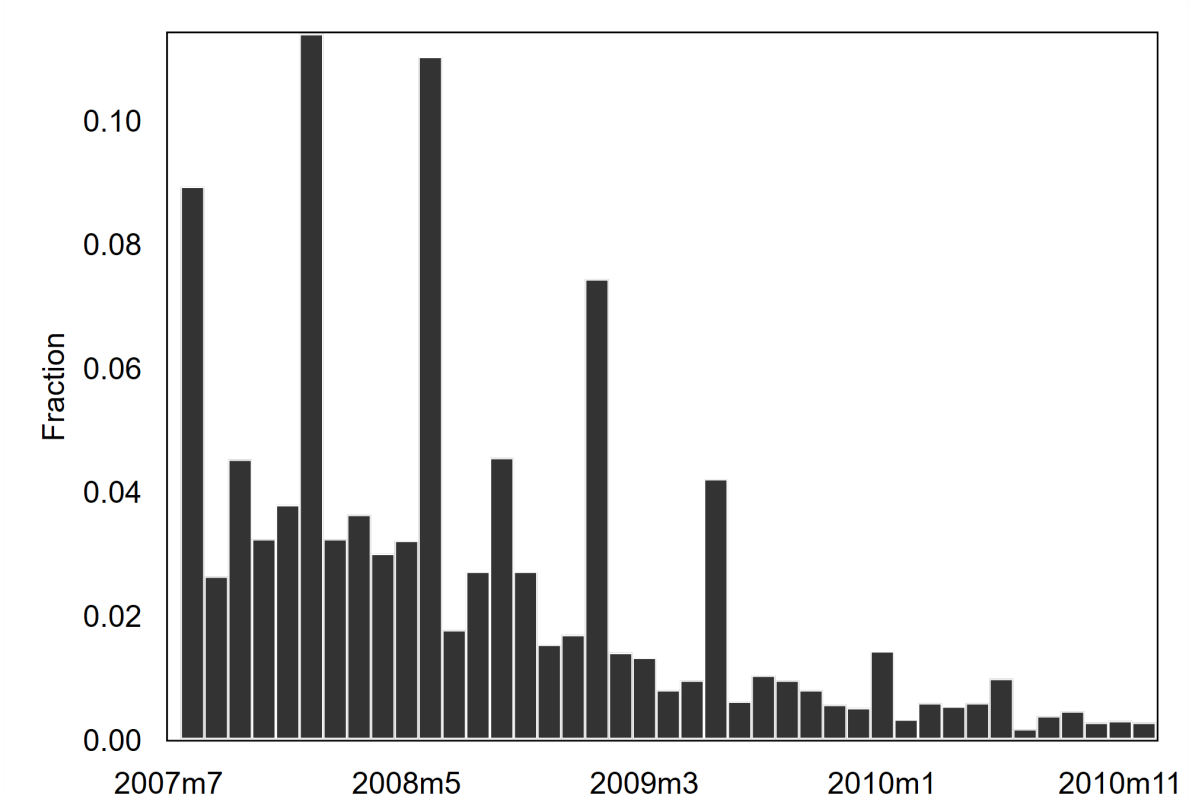


Figure A.3: Time of graduation for treated students

Table A.3: Separate regression results on placebo cohorts. Dependent variable: Log earnings in years one to five after graduation.

	Year after graduation				
	1	2	3	4	5
A. Placebo Cohort 2002					
Recoded GPA	0.001	-0.017	-0.049	-0.008	-0.073
	(0.082)	(0.053)	(0.045)	(0.052)	(0.052)
B. Placebo Cohort 2003					
Recoded GPA	0.071	0.021	0.038	0.041	-0.023
	(0.066)	(0.051)	(0.040)	(0.045)	(0.039)
C. Placebo Cohort 2004					
Recoded GPA	-0.098	-0.009	-0.010	-0.028	-0.002
	(0.061)	(0.054)	(0.037)	(0.051)	(0.039)

Notes: This table resembles Table 3, but shows estimates based on placebo samples. The placebo samples are obtained by implementing a placebo recoding of grades on July 31, 2002, July 31, 2003, and July 31, 2004 using the same recoding scheme, covariates, and sample selection as in the main specification. See notes for Table 3.

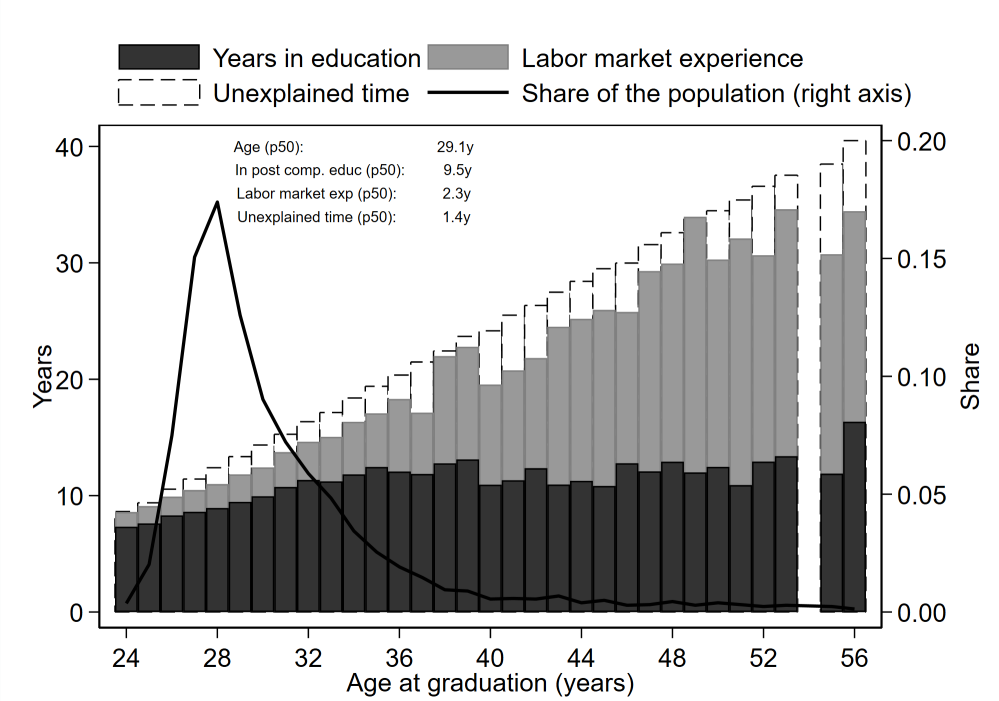


Figure A.4: Age profile of graduates in the analysis sample

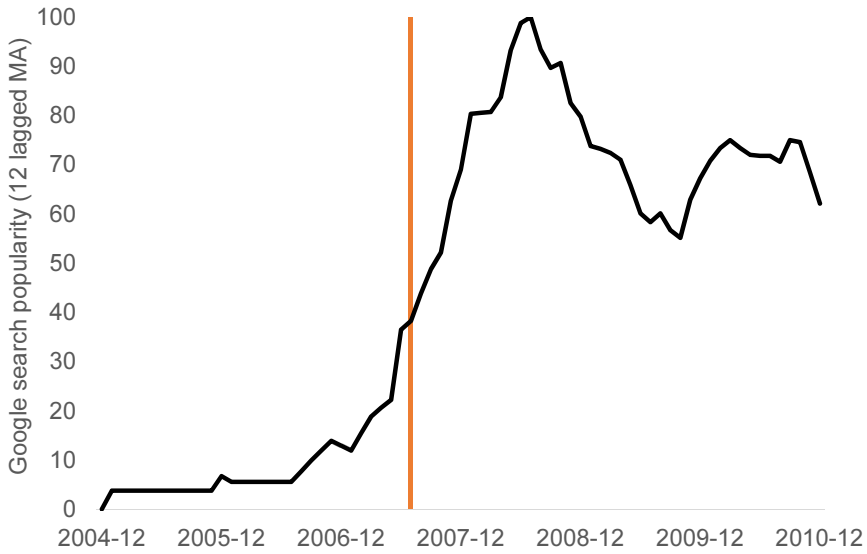
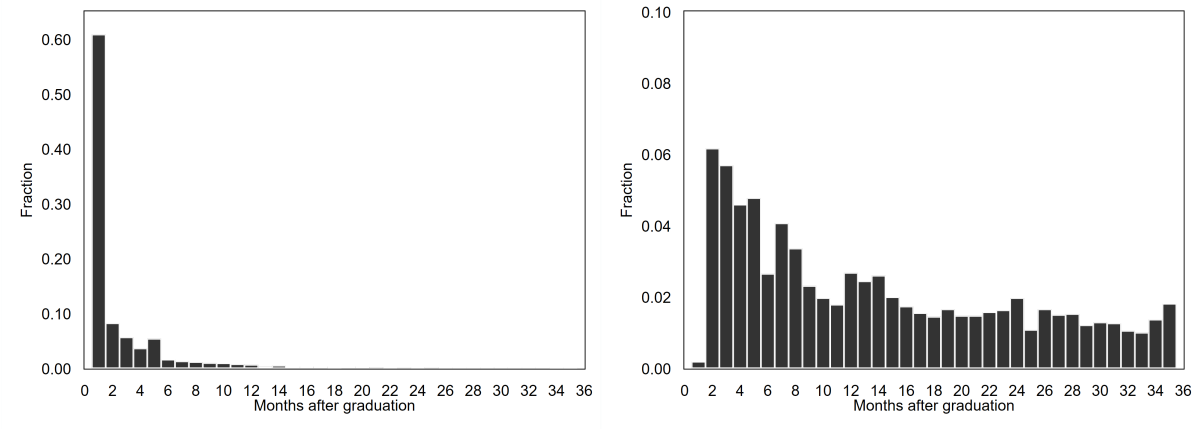


Figure A.5: Google search trend for "den nye karakterskala" (English: "the new grading scale").



(a) Started 1st employment spell

(b) Started 2nd employment spell

Figure A.6: Employment spell timing for treated students.

B 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 correlated or 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.
  - 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.

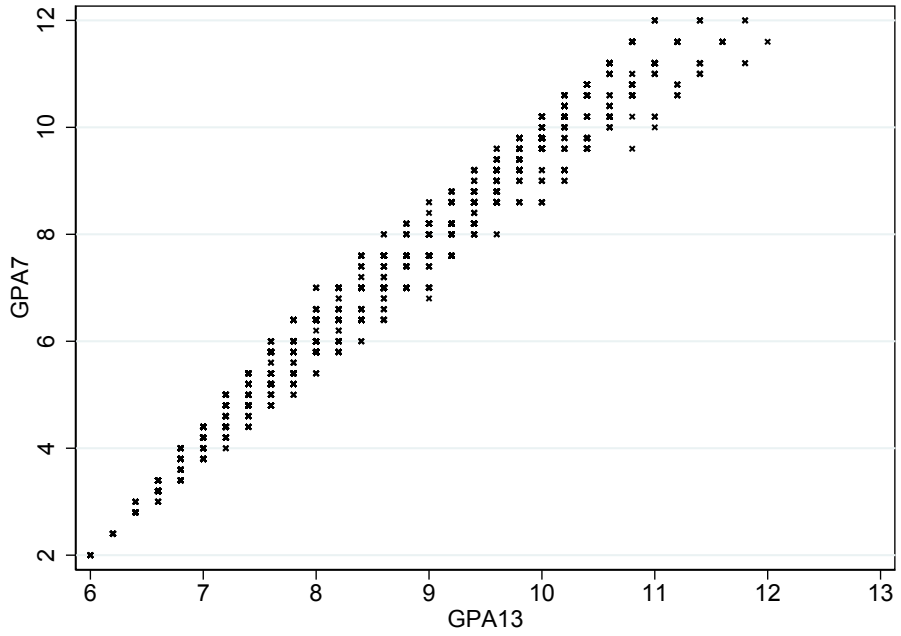


Figure B.1: A simulated example of the relationship between GPA7 and GPA13

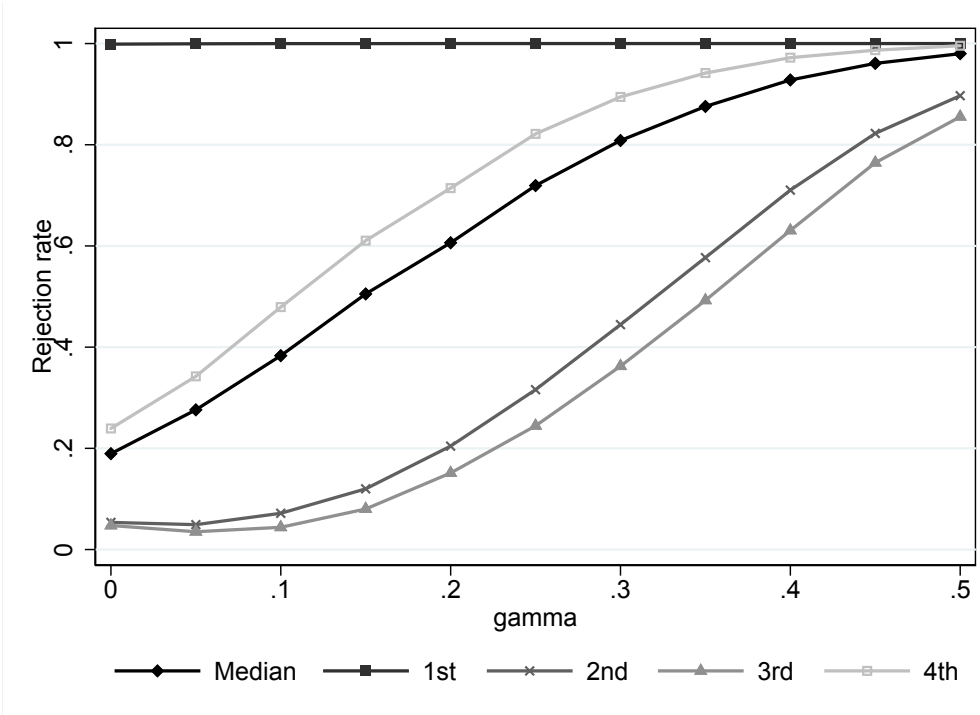
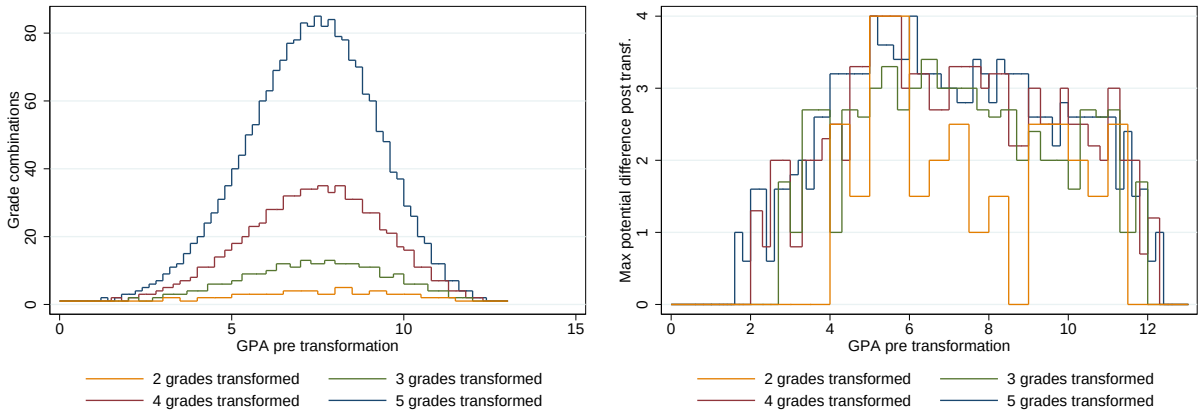


Figure B.2: Rejection rates.

C Additional Material - Not for Publication



(a) Grade combinations

(b) Maximum difference in post-transformation GPA

Figure D.1: Combinations and maximum difference, given GPA, and number of transformed grades.

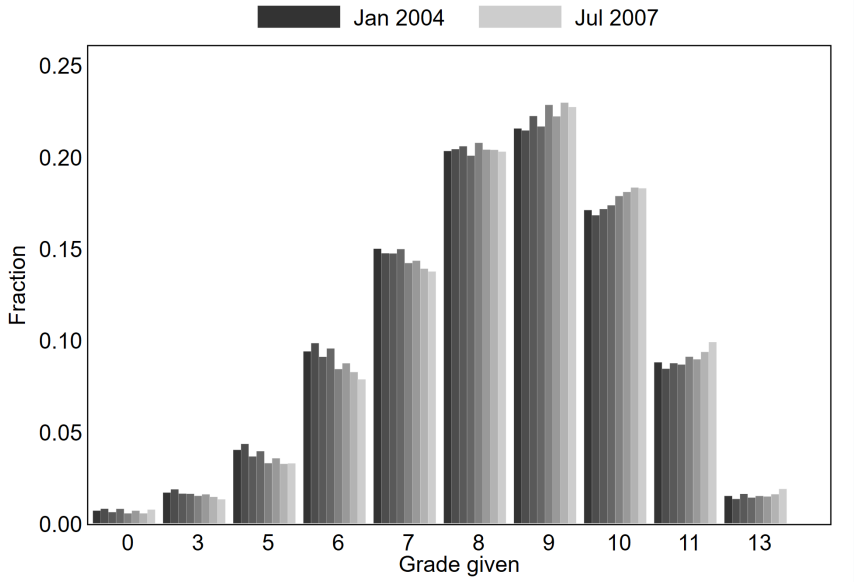


Figure D.2: Relative grade frequency.

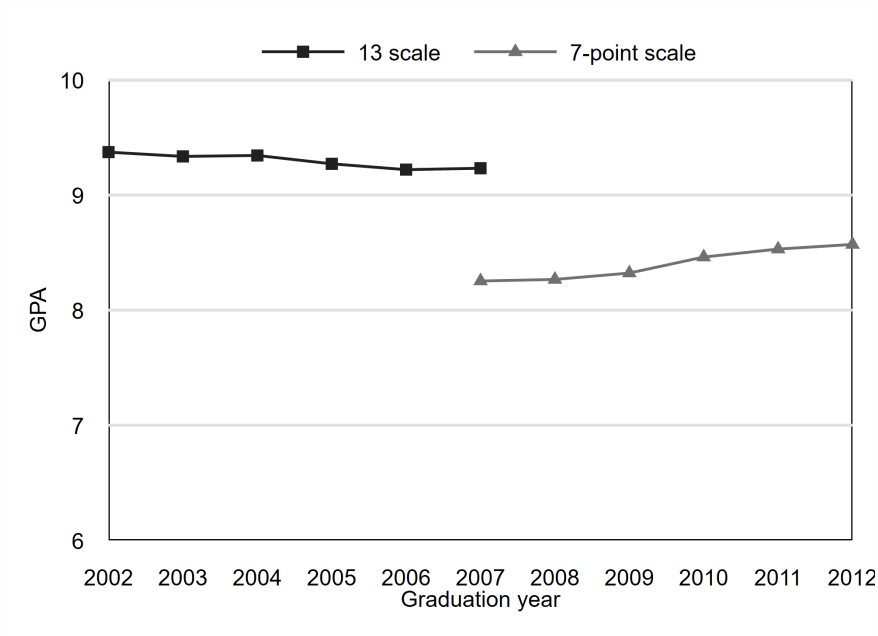


Figure D.3: Average university GPA by year of graduation.

Notes: The sample consists of all students who commenced and finished their postgraduate studies in the years 2000-2012 at the University of Copenhagen or Aarhus University.