I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14200

# Grades and Employer Learning

Anne Toft Hansen
Ulrik Hvidman
Hans Henrik Sievertsen

MARCH 2021

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14200

# Grades and Employer Learning

**Anne Toft Hansen**
*VIVE*

**Ulrik Hvidman**
*Aarhus University*

**Hans Henrik Sievertsen**
*VIVE, University of Bristol and IZA*

MARCH 2021



ISSN: 2365-9793

**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

# ABSTRACT

## Grades and Employer Learning*

This study examines the labor-market returns of skill signals. We identify the labor-market effect of grade point averages (GPA) by leveraging a nationwide change in the scaling of grades in Danish universities. 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 and the earnings adjustment occurs both within and across firms. Although employers initially screen candidates based on skill signals, our findings suggest that they rapidly learn about worker productivity.

**JEL Classification:** I20, J20, I26

**Keywords:** job-market signaling, employer learning, higher education

**Corresponding author:**
Hans Henrik Sievertsen
Department of Economics
University of Bristol
Priory Road Complex
Priory Road
Office 2C0
Bristol BS8 1TU
United Kingdom

E-mail: h.h.sievertsen@bristol.ac.uk

* We thank Simon Burgess, Thomas S. Dee, Mette Ejrnæs, Fabian Lange, Christine Valente, and Miriam Wüst, as well as the 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, and the seminar participants in Aarhus, Bath, Bristol, Copenhagen, Southampton, and Trondheim, for their valuable comments. 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.

# 1 Introduction

How workers are allocated across jobs has implications for inequality and efficiency at the aggregate level (Roy, 1951; Sattinger, 1975) and employer–employee mismatches are costly at the micro level (Fredriksson, Hensvik and Skans, 2018). A major challenge in the matching process is that labor-market productivity is imperfectly observed by employers. 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 applicant's 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 in terms of the applicant's grade point average (GPA).[2]

In this paper, we examine the signaling value of a GPA at labor-market entry and how employers learn about workers' productivity. To identify the labor-market effect of GPAs, we exploit a grading reform in Denmark that mimics an ideal experiment by creating variation in university graduates' GPAs that is unrelated to outcome-relevant characteristics. 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 post-reform GPA as two individuals with identical pre-reform GPAs could end up with more than a standard deviation difference in their post-reform GPAs. We use this reform-induced variation in grades to identify the effects of GPAs on labor-market outcomes. These features of the reform enable us to produce credible estimates of the effects of grade signals that are unrelated to achievement in a naturally occurring setting.

We use a novel data set containing GPAs for all students at the two largest universities in Denmark, Aarhus University and University of Copenhagen, corresponding to around half of the total population of university students in Denmark.[3] Thus, we examine varia-

[1] Between 2000 and 2016, the proportion of 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 students aged 25 to 34 completed tertiary education in the UK.

[2] Signals of educational achievement in terms of a GPA 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).

[3] About one out of 10 in the population aged 30–35 graduated from either University of Copenhagen or Aarhus University and more than one out of five when looking at the share of this population with a yearly income of more than 60,000 euro (see Appendix Table A.1)

tion in salient grades among a large and diverse group of university students. Moreover, we exploit the detailed Danish administrative data to examine labor-market outcomes over time, which allows us to study how employers learn about productivity.

Our findings show that a reform-induced increase in GPA causes higher earnings in the first two years after graduation. However, this effect diminishes over time, 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 three non-treated cohorts. These supplementary analyses provide strong evidence that the relationship between reform-induced variation in GPAs and labor-market outcomes is not driven by the recoding benefiting specific grade combinations that are rewarded on the labor market.

The effect at labor-market entry is strongest among graduates with parents without a university degree, graduates with a shorter work history, and graduates from Aarhus University (which has a smaller student labor market compared to University of Copenhagen). These findings suggest that the signaling value of the GPA is particularly important for graduates with limited alternative signals in terms of informal links (proxied by parental education) and work history. We also find that the effect is strongest for graduates from majors that are typically more likely to be employed in the private sector and in jobs with greater earnings dispersion. This set of findings indicates that the signaling effect is largest in less regulated labor markets and in labor markets with high variation in wages.

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. 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 adjustment is fastest among workers who change firms, the slower earnings growth is detectable for all workers, including those who stay with the same firm. These findings suggest that the earnings adjustments occur both within and across firms.

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 suggests 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 disappears within three years, and the adjustment occurs both within firms and across firms.

This study contributes to the literature on labor-market sorting and employer learning. While most existing research on signaling and sorting has focused on educational attainment, a few studies have examined the signaling value of educational credentials among students with similar lengths of study. The first group of studies applies experimental curriculum vitae (CV) designs and demonstrates that variation in signals at the intensive margin (i.e., how well you completed the degree in terms of your 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). The second group of studies documents returns to receiving an honors degree classification (, 2015; Khoo and Ost, 2018). The third group of studies exploits discontinuity around the passing cutoff for specific degrees (Clark and Martorell, 2014; Jepsen, Mueser and Troske, 2016). Our study contributes to these studies by identifying the effect of a GPA on actual labor-market outcomes over time for the population of students across educational degrees in university.[4] Our findings also reinforce the existing evidence that employers learn rather quickly about actual productivity (Lange, 2007; Aryal, Bhuller and Lange, 2019).

The remainder of this paper is organized as follows. Section 2 provides the institutional details of the grading reform. Section 3 describes the data., Section 4 presents the research design and empirical strategy, and Section 5 presents the results. Section 6 concludes this paper.

# 2 Background

## 2.1 The 2007 Danish grading reform

In 2006, the Danish government decided to introduce a new grading scale. In August 2007, the 7-point grading scale replaced the old "13-scale" in all educational programs in Denmark.[5] How the grading reform affected students depended on where they were in the educational system on August 1, 2007. If the individual had completed a degree, their entire GPA would be transformed to a new GPA according to a mapping scheme provided by the Ministry of Education. However, for students enrolled in an educational

[4] An alternative signal of productivity is university prestige. Bordón and Braga (2020) exploit university admission test score cutoffs in a regression discontinuity approach and find a substantial initial wage premium of graduating from a prestigious university, which decreases over time.

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

program on August 1, all given grades were recoded to the new scale based on the scheme shown in Table 1, which gave rise to the variation in GPAs that we exploit in this study.

**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<br>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<br>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<br>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 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 the new scale has fewer grades (seven compared to 10), three pairs of grades in the old scheme were collapsed into 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 recoding. Second, the distances between the old and the new grades vary. Although most students were downgraded in absolute terms due to the differences in the two scales, two students with identical pre-reform GPAs could have very different post-reform GPAs because grades were recoded differently.

Consider, for example, a student with grades 8 and 10 on the 13-scale and another

student with two 9s on the 13-scale. They both have a GPA of exactly 9.0 before the recoding. 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 both 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

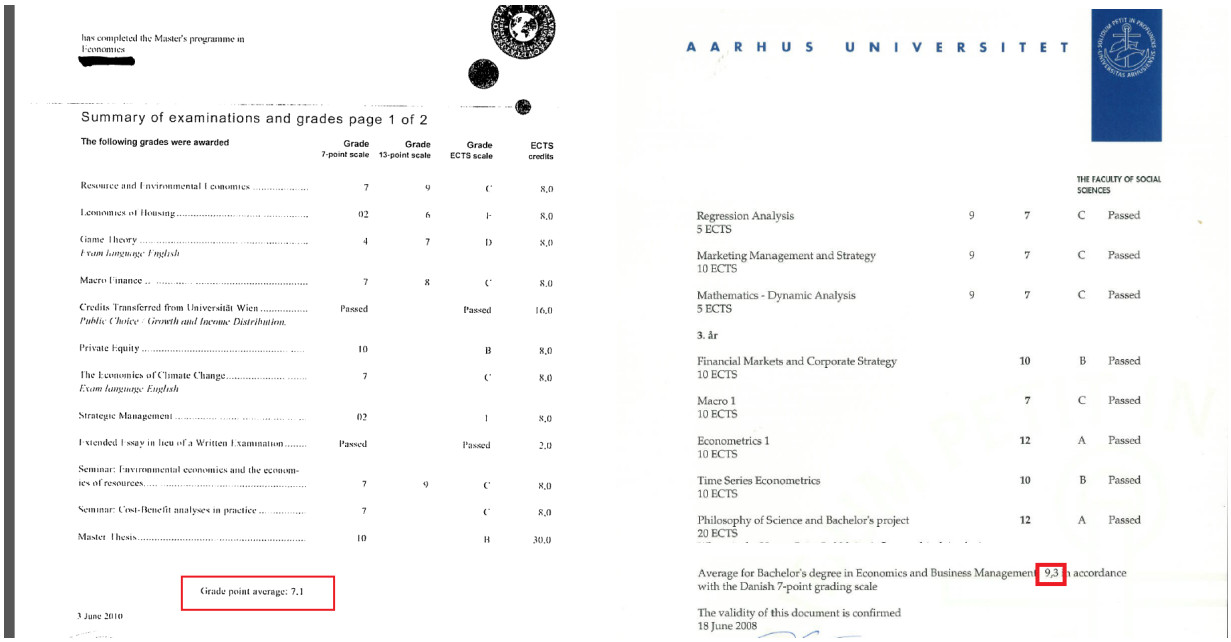
After completing upper secondary education, students can apply for university programs in Denmark. All programs are free, and all students over the age of 18 receive a monthly stipend to pay for their living costs. Enrollment in university programs depends almost exclusively on high school GPA.

Denmark has adopted a three-year bachelor's and a two-year master's structure for most of its university programs.[6] As we focus on the importance of GPAs for labor-market outcomes, we only consider graduates from a 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. Figure 1 shows two examples of diplomas for treated individuals. The diplomas illustrate what the employers observe when making hiring and wage decisions. The diploma shows the student's GPA based on the recoded grades (marked with red).

As Figure 1 shows, the diplomas also 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. 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

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



has completed the Master's programme in Economics

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 | | I | 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

**(a)** University of Copenhagen

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

**(b)** Aarhus University

**Figure 1:** Examples of diplomas for treated individuals

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 University of Copenhagen. Students will be at different stages of the programs at this point, as illustrated in Figure 2. As the treatment (i.e., the reform-induced variation in GPA) is caused by the recoding of their given grades on August 1, 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 2). Specifically, we restrict the sample to the students who had at most 40 ECTS points remaining in 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 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 pro-

grams, we cannot strictly impose such criteria. 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. Finally, to be able to follow the students in the first years on the labor market, we include students who graduated before 2011. This sample restriction only requires students to have completed coursework that was scheduled to take considerably less than a year within three years. As Appendix Figure A.2 shows, 23 percent of the students have graduated by the end of 2007 and 94 percent by the end of 2009. No further sample restrictions are applied, and the final sample consists of 3,813 students. We show that our conclusions are not sensitive to changing the sample selection criteria.

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

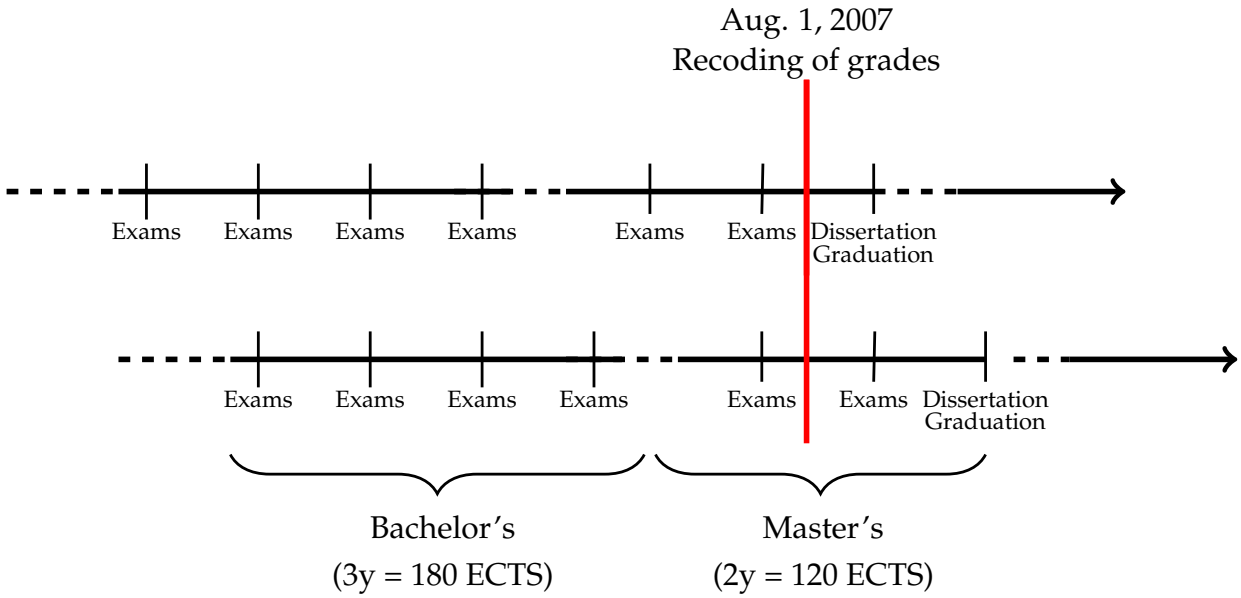



**Figure 2:** University students' exposure to the implementation of the new grading scheme

## 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 data from Statistics Denmark, 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, the 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 on this variable.

Our main labor-market outcome is log total gross earnings in the first five calendar years after graduation. In the model with our main labor-market outcomes, 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 takes the value of 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. In our sample, 65 percent of the students are female and 26 percent of the sample are children of parents with a university degree. Forty-one percent of the sample are students from University of Copenhagen. On average, students were fairly close to graduation (23 ECTS remaining), and the grades given before the recoding accounted for about 70 percent of the overall GPA.

Ninety percent of our sample have positive earnings in the year after graduation. A university graduate in our sample earns on average 43 thousand euros (2015 level, gross) in the first calendar year after graduation (corresponding to around USD 48,000), with 66 percent of them working in the public sector.

## 4 Identification and Estimation

### 4.1 Empirical strategy

In an ideal experiment, we would randomly assign the GPA 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 ideal experiment.

Figure 3a shows the GPA at the time of the grading reform. The horizontal axis shows the GPA before the recoding, and the vertical axis shows the GPA after the recoding. Consider for example the students who had a GPA of 8.0 before the reform. Among those

**Table 2:** Summary statistics

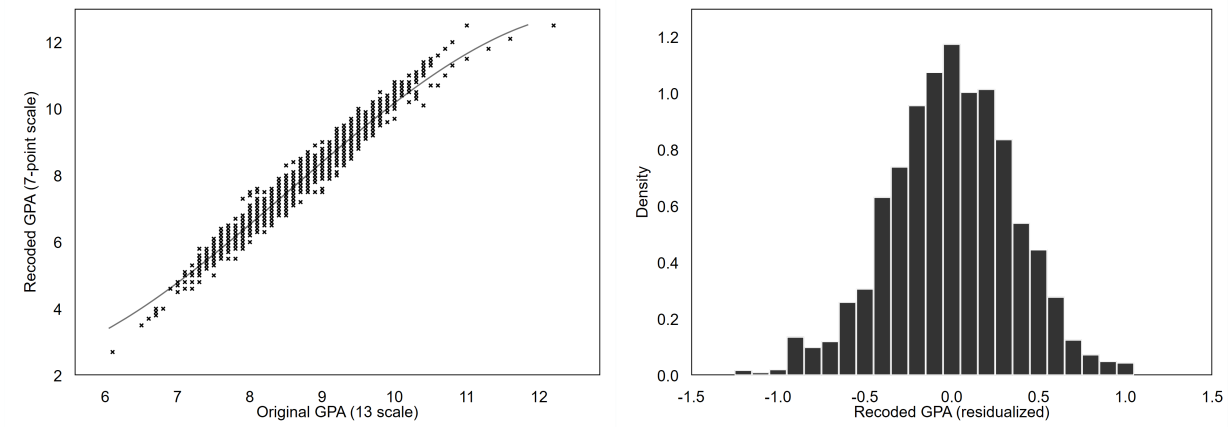
| | Mean | SD | P10 | P50 | P90 |
|---|---|---|---|---|---|
| *A. Background* | | | | | |
| Age at graduation (years) | 30.87 | 5.82 | 27.51 | 29.08 | 32.12 |
| Female | 0.65 | 0.48 | 0.00 | 1.00 | 1.00 |
| Parental disposable income (1000 euros) | 39.41 | 39.78 | 26.56 | 34.30 | 43.05 |
| Parents with 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.

students, some had a GPA of 5.3 and some a GPA of 7.1 after the recoding. This 1.8 GPA difference for students with identical pre-reform GPAs is larger than the standard deviation on the final GPA (recall from Table 2 that the standard deviation on the final GPA is 1.66).

A cubic relationship between pre- and post-recoding GPA fits the variation well, as shown by the solid line in Figure 3a. In our main strategy, the deviation from this fitted line constitutes the treatment variable. Given that a student had a GPA of 8.0 on the old scale, we would predict that the student would have a GPA of 6.0 on the new scale after the recoding. The students with a recoded GPA of 5.3 had a deviation from this guess of -0.7, whereas students with a recoded GPA of 7.1 had a deviation of 1.1. Our goal is to test whether the latter has significantly higher earnings than the former. Figure 3b shows the distribution of the corresponding deviations from the cubic fit. About 50 percent of the students have a deviation of between -0.23 and 0.24 and 98 percent of the sample has

a deviation between -0.90 and 0.88.




**(a)** Pre- and post-recoding GPA **(b)** Residuals

**Figure 3:** 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 three observations are shown.

In practice, we implement our empirical strategy by estimating the following equation using ordinary least squares (OLS):

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

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 reducing or increasing the polynomial degree or using a less parametric approach.[7] Estimating equation (1) corresponds to a two-step procedure, where the first step consists of computing the residuals shown in Figure 3b, and the second step is a regression of earnings on these residuals. $\beta_1$ captures the importance of variation in post-recoding GPA given pre-recoding GPA.

[7] 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 with a linear specification, using a 5 percent cutoff. This is because 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. 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.

Our identification strategy does not rely on any covariates.[8] 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 (i.e., major) studied and the institution. We, therefore, include indicators for the program studied and an indicator for institution (University of Copenhagen or Aarhus University). Second, earnings might also be related to individual characteristics and background. We, therefore, control for age, an indicator for parental origin (non-Western or Western according to the definition by Statistics Denmark), parental income, parental unemployment, parental education, and gender. Parental variables are created as 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 linearly 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 labor-market productivity (e.g., cognitive and non-cognitive skills). To allow for arbitrary correlation within pre-recoding GPA cells, we cluster bootstrap on the pre-recoding GPA level with one decimal point as reported on the diploma.

## 4.2 Identifying assumptions

**Identifying assumption 1.** The causal interpretation of $\beta_1$ 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, it could be the case that the recoding algorithm in itself captures characteristics that are valued on the labor market. Consider the example presented in the previous section. A pre-reform grade combination of two 9s gives a recoded GPA of 7.0, while a pre-reform grade combination of an 8 and a 10 gives a recoded GPA of 8.5. To assess whether this is the case, we conduct placebo tests where we implement the recoding scheme on three non-treated cohorts. In practice, we apply the same sample selection on these cohorts and recode their GPA with exactly the same algorithm, and then run a regression of equation (1) using the same outcome variables. These placebo tests provide evidence on whether the recoding algorithm in itself captures informative about labor-market productivity.

Second, the variation in GPA induced by the reform and the recoding might be correlated with individual characteristics. This could be driven by anticipation effects, if more forward-looking students anticipate the reform and either advance (if they would

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

be punished by the reform) or delay (if they would gain) their studies relative to the reform implementation date.[9] To assess whether individuals who were upgraded are systematically different to those who were downgraded, we show that the reform-induced variation in GPA is unrelated to observable individual characteristics, such as parental background, high school GPA, and undergraduate GPA.

**Identifying assumption 2.** Recall that students completed units and/or their dissertation after the recoding occurred. 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. In related work, we show that high school students reacted to negative variation in their GPA induced by the same reform by increasing their subsequent effort (Hvidman and Sievertsen, forthcoming). If university students responded in the same way, the grades given after the recoding—which enters the overall GPA—could be affected. Importantly, if the students who receive a negative GPA shock due to the grading reform increase their study effort, 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 also course selection after the recoding.

We explicitly test for these post-reform responses in terms of likelihood of graduation, time to graduation, study effort, and choice of elective units. 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 who are very close to graduation at the time of the reform, they have limited time to react to the grading reform.

# 5 Results

## 5.1 Returns to the reform-induced variation in GPA

Table 3 presents the estimated $\beta_1$ coefficients from equation (1), using log earnings in the first five years after graduation as the dependent variables. The coefficient suggests that a one-unit increase in the GPA is related to almost 9 percent higher earnings in the first calendar year after graduation, and almost 8 percent higher earnings in the second

[9] While the reform was announced in advance, we observe relatively little Google search activity before the actual implementation, as shown in Appendix Figure A.3.

calendar year after graduation. The coefficients are both considerably smaller and not statistically different from zero in years three to five after graduation. To get a sense of the magnitude of these coefficients, we should scale them by 0.7 because the recoded GPA on average accounts for 70 percent of the final GPA. A one-unit increase in the final GPA, therefore, causes an increase in earnings of 12 percent in the first year after graduation ($0.087/0.7 = 0.12$). Recall from Table 2 that the standard deviation on the GPA is 1.66. The results therefore imply that a one standard deviation increase in GPA increases earnings by approximately 21 percent in the first year after graduation ($1.66 \times 0.087/0.7 = 0.21$).

**Table 3:** Regression results. Dependent variable: Log earnings year in years one to five after graduation.

| | Year after graduation | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| Recoded GPA | 0.087*** | 0.078** | 0.006 | 0.002 | -0.019 |
| | (0.030) | (0.033) | (0.026) | (0.027) | (0.038) |
| Observations | 3445 | 3465 | 3423 | 3388 | 3366 |
| R-squared | 0.24 | 0.18 | 0.20 | 0.17 | 0.16 |

Notes: Recoded GPA is the overall university GPA measured on the new grading scale from -3 (worst) to 12 (best). All models control for pre-recoding GPA using a third-order polynomial and are estimated with program fixed effects and the full set of covariates, which includes an indicator for institution (Copenhagen or Aarhus), 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 an 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 in parentheses. Asterisks indicate significance at the following levels: * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

In Appendix C, we show that the naive OLS estimate suggests that a one standard deviation increase in GPA is associated with a 6 to 7 percent increase in earnings in the first year after graduation (see Appendix Table C.1). The 95-percent confidence interval from our main specification includes an effect size of 7 percent as a result of a one standard deviation increase in GPA ($1.66 \times (0.087 - 1.96 \times 0.03)/0.7 = 0.07$). If employers have little information beyond what we observe to determine the earnings level at labor-market entry, this could explain why the naive OLS estimate is within the 95-percent interval in the first two years after graduation. Interestingly, the naive coefficient only declines slightly in the first five years after graduation, as shown in Appendix Table C.1, whereas the effect of the reform-induced GPA quickly diminishes.

## 5.2 Validity of the research design

**A. Placebo tests: Does the recoding algorithm predict labour 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. We implement the same sample selection criteria and the same recoding algorithm for cohorts in July 2002, July 2003, and July 2004 as for the actual treated cohorts in 2007 and conduct exactly the same analysis as shown in Table 3. In other words, we consider all students who had at most 40 ECTS remaining 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 then we estimate equation (1). The results are shown in Table 4. We find no detectable correlation between the recoded GPA and earnings in any of the years for any of the three cohorts. This 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:** Regression results on placebo cohorts. Dependent variable: Log earnings 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.

**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 to test whether the reform-induced variation in GPA is related to any observable characteristics. Column (1) of Table

**Table 5:** Regression results on covariate balance. Dependent variables in column header.

| | Female | HS GPA | BSc/BA GPA | — Parents' — | | | Predict. earn. |
|---|---|---|---|---|---|---|---|
| | | | | Income | Unempl. | Uni. degr. | |
| | (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 in parentheses. Asterisks indicate significance at the following levels: * $p< 0.1$, ** $p < 0.05$, and *** $p< 0.01$.

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 of the students.

**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. 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 elective units as a consequence of the recoding. Fourth, we test whether the recoding affected time to graduation.

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. Columns

(1) and (2) show that the reform-induced variation in GPA do 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[10] and performance in subsequent assessments.

**Table 6:** 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 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. All models are estimated with the full set of covariates (see notes for Table 3). Standard errors clustered at the pre-recoding GPA level in parentheses. Asterisks indicate significance at the following levels: $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, and $^{***}$ $p< 0.01$.

**D. Are the results sensitive to model specification?** We now assess whether our findings are sensitive to the empirical specification. Figure 4 shows point estimates and confidence intervals for the coefficient on $\beta_1$ 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.

[10]To assess whether the students select different types of elective units, we create a measure of how demanding the subsequent units are. We do this by running a regression of exam marks 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 their high school GPA, a student receives a higher grade. In other words, the fixed effects capture unit difficulty (or marking 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.

All 72 specifications shown in Figure 4 lead to positive and significant point 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.

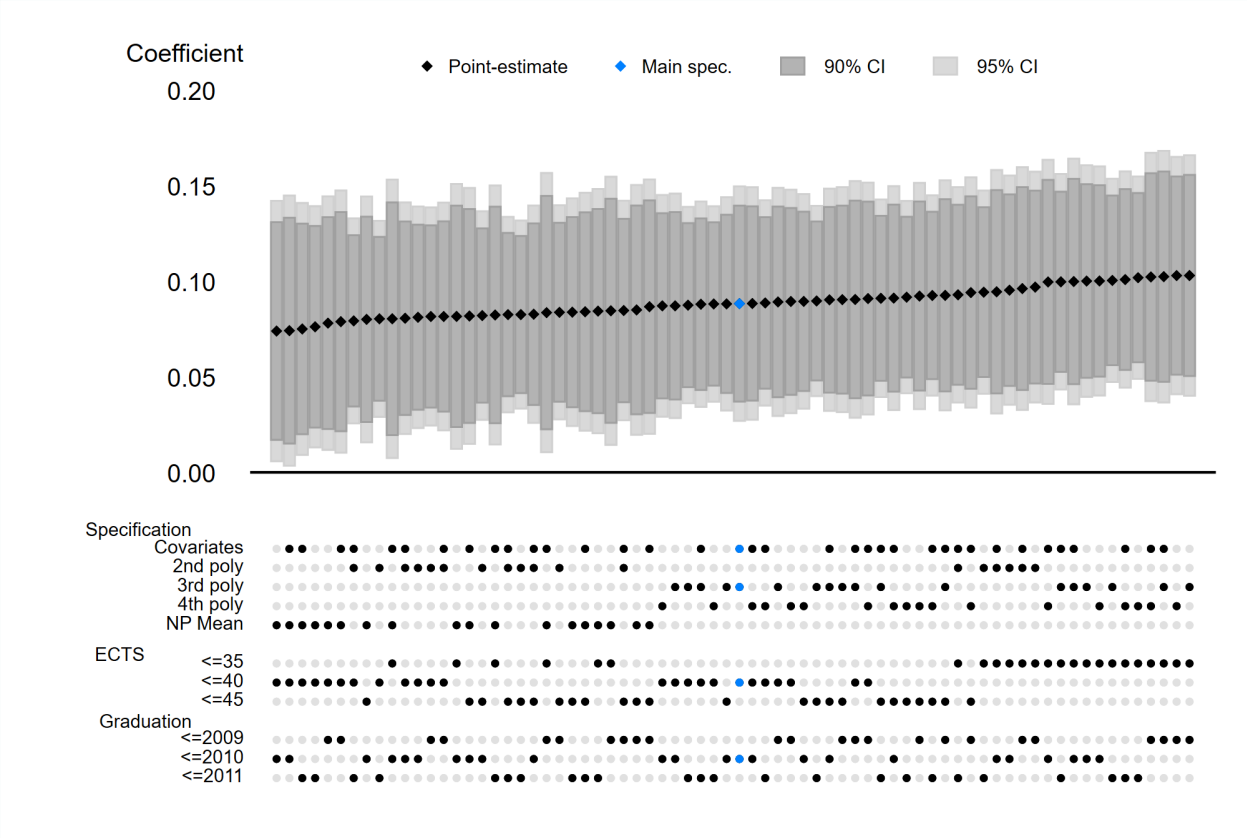



**Figure 4:** Specification curve.

Notes: The chart shows the 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.

**E. Is the relationship symmetric?** So far, we have been assuming that there is 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 5 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.

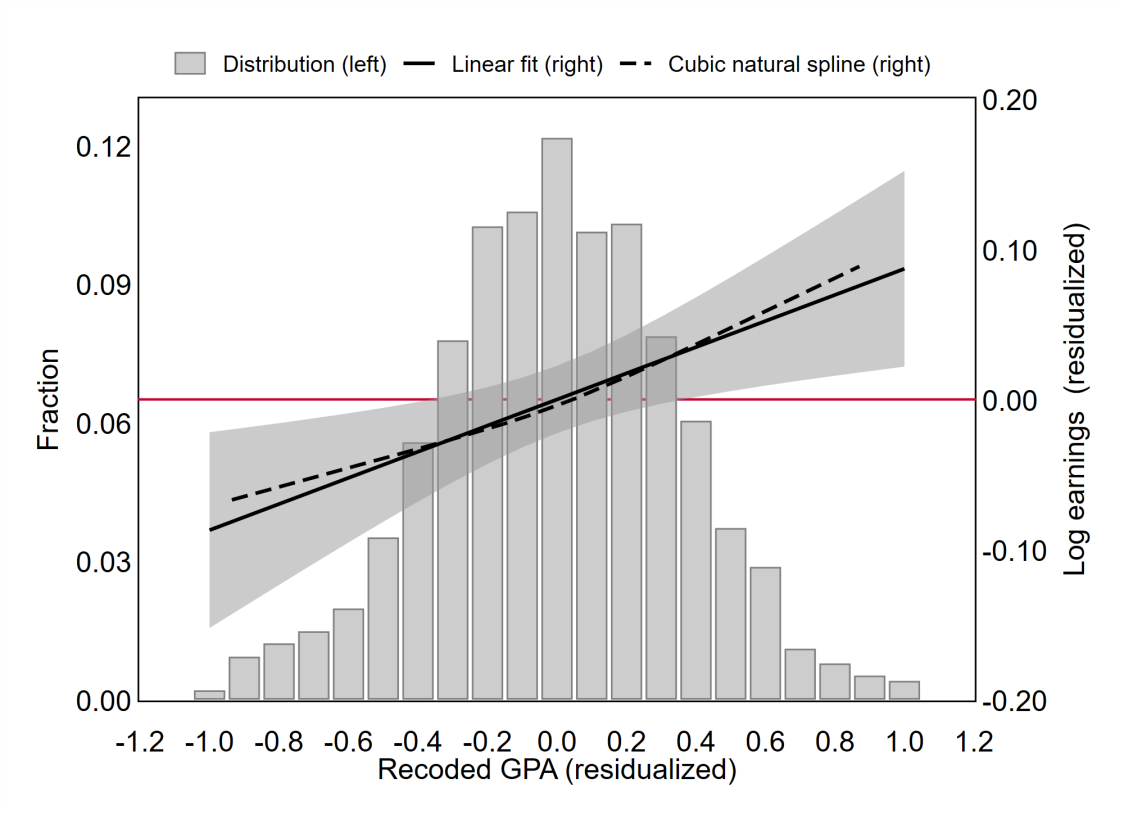



**Figure 5:** The relationship between log earnings in the calendar year after graduation and reform-induced variation in GPA.

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.

## 5.3 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 presents the results of both mechanisms that could drive the effect 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), but instead of using log earnings as the dependent variable, we estimate the model with different labor-market outcomes as the dependent variable. The first row of results shows that there is no effect on the extensive margin of earnings.

The second and third rows of Table 7 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 of Table 7 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, rows 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 unrelated 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.

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

One could expect that the reform-induced GPA plays a bigger role in less regulated labour 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. Again, in line with the hypothesis on the importance of the labor-market structure, we find larger effects for majors with higher

**Table 7:** Regression results. Other labour 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 in parentheses. Asterisks indicate significance at the following levels: * $p < 0.1$, ** $p < 0.05$, and *** $p< 0.01$.

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. One might expect that graduates with better networks on the relevant job market—as proxied by parental education—are 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. It is not uncommon for students in Denmark to have part-time jobs that are relevant to their studies and later careers. In line with this hypothesis, 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 from Aarhus University. As the student labor market is considerably larger in Copenhagen, especially within the social sciences and business, we treat the latter variable as a proxy for less labor-market experience.

**Table 8:** 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 $\beta_1$ 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 also find that the employers rapidly learn about productivity and that the earnings adjustments occur both within and across firms. We find no evidence of an effect on the likelihood of employment just after graduation or on job changes within the first five years on the job market. Furthermore, we find 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 and bunching, for example, could make 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, 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 finding that grades are used as a signaling device suggests that such external factors could have implications for labor-market outcomes.[11] Third, Bar, Kadiyali and Zussman (2009) show that students select into electives 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. Fourth, and finally, while the setting that we study is unique, grading reforms are 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.

More generally, our findings provide evidence of the importance of skill signals in the

[11] 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)). Our study provides evidence of one source of "noise" in university graduates' signals of educational achievement.

labor market and key insights in terms of 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


**Aryal, Gaurab, Manudeep Bhuller, and Fabian Lange.** 2019. "Signaling and Employer Learning with Instruments." National Bureau of Economic Research Working Paper 25885.

**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.

**Christensen, Jens Jakob, and N Westergard-Nielsen.** 1999. "Wages and human capital: the Danish evidence." *Returns to Human Capital in Europe*, 29–42.

**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.

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

**Hvidman, Ulrik, and Hans Henrik Sievertsen.** forthcoming. "High-Stakes Grades and Student Behavior." *Journal of Human Resources*.

**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.

**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.

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

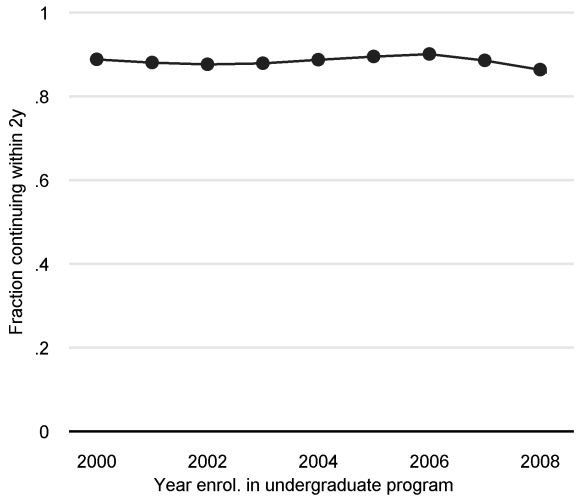
**Sattinger, Michael.** 1975. "Comparative advantage and the distributions of earnings and abilities." *Econometrica*, 43(3): 455.

**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.

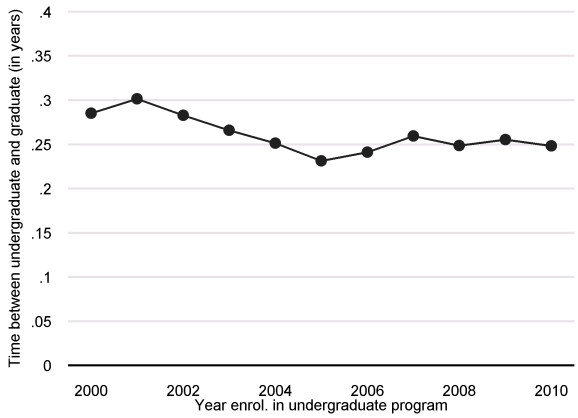
**Spence, Michael.** 1973. "Job market signaling." *The Quarterly Journal of Economics*, 355–374.

**The earnings returns to graduating with honors — Evidence from law graduates.** 2015. "The earnings returns to graduating with honors — Evidence from law graduates." *Labour Economics*, 34: 39 – 50.

# A Appendix




**(a)** Share continuing within 2y




**(b)** Time btw degrees

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

| | 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.1:** Educational background for the population aged 30-35, 2008-2012

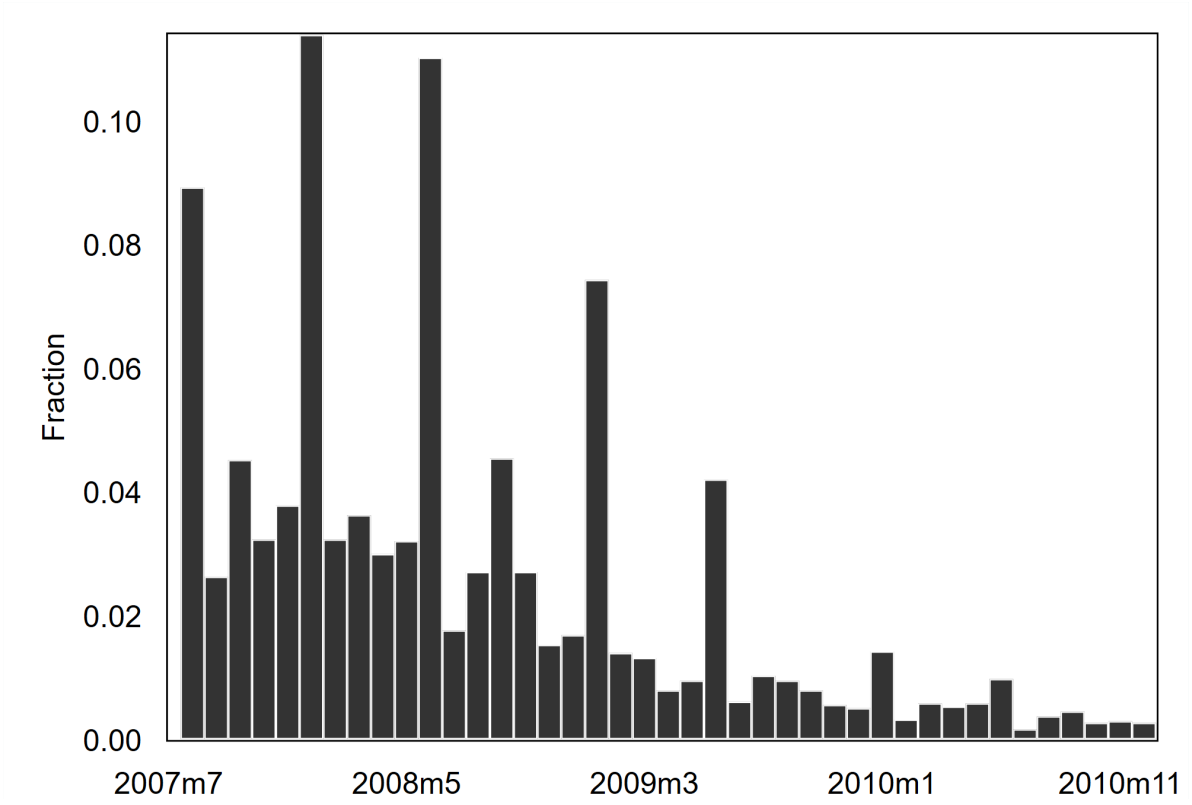



**Figure A.2:** Time of graduation for treated students.

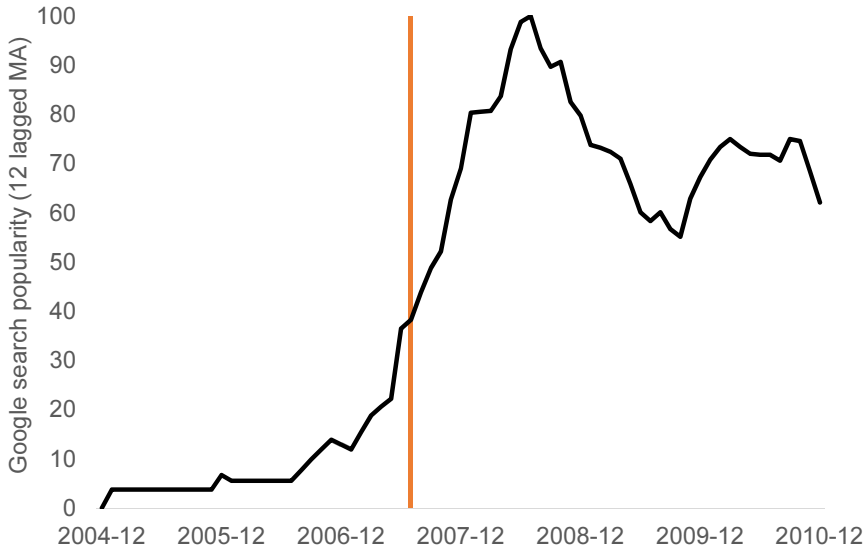
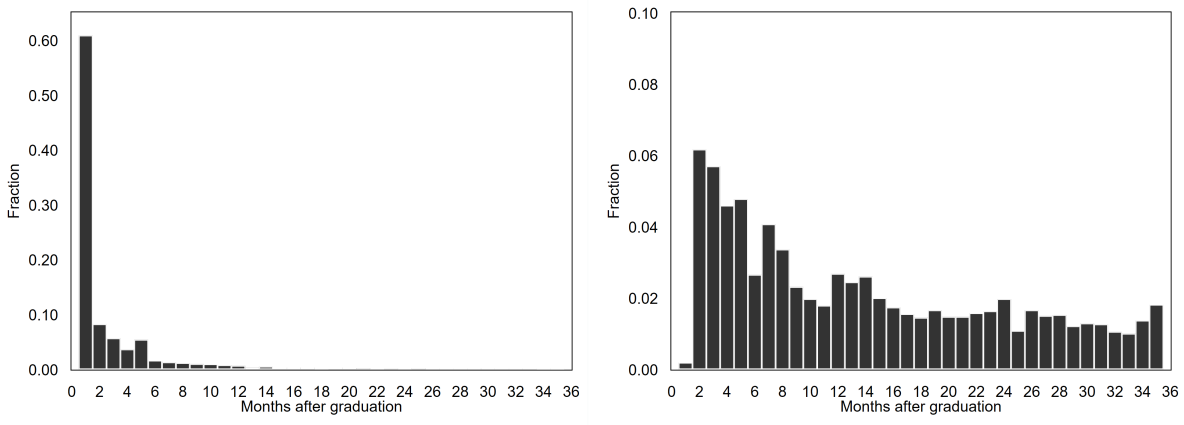



**Figure A.3:** Google search trend for "den nye karakterskala" (English: "the new grading scale").




**(a)** Started 1st employment spell

**(b)** Started 2nd employment spell

**Figure A.4:** 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 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 $\gamma$ 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.

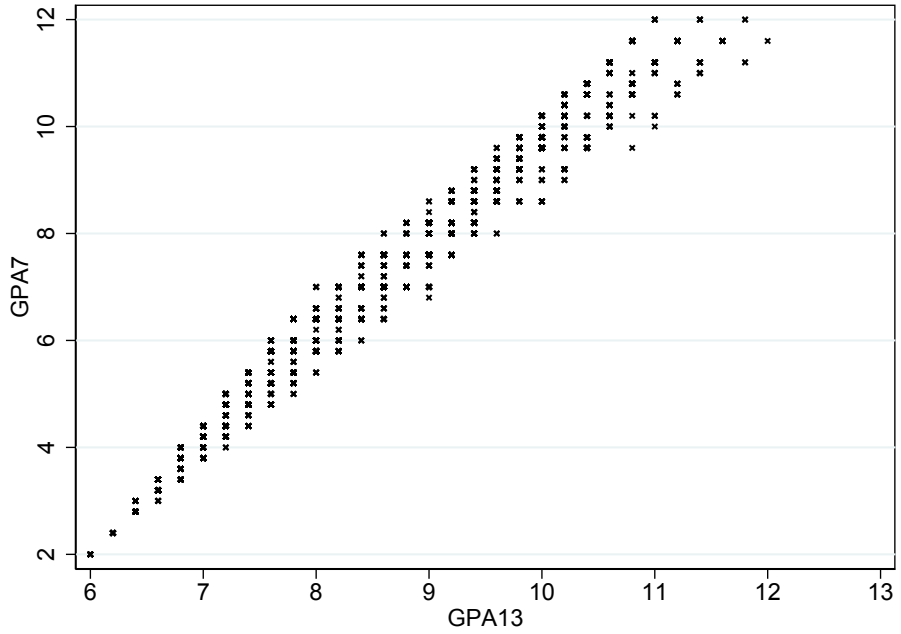



**Figure B.1:** A simulated example.

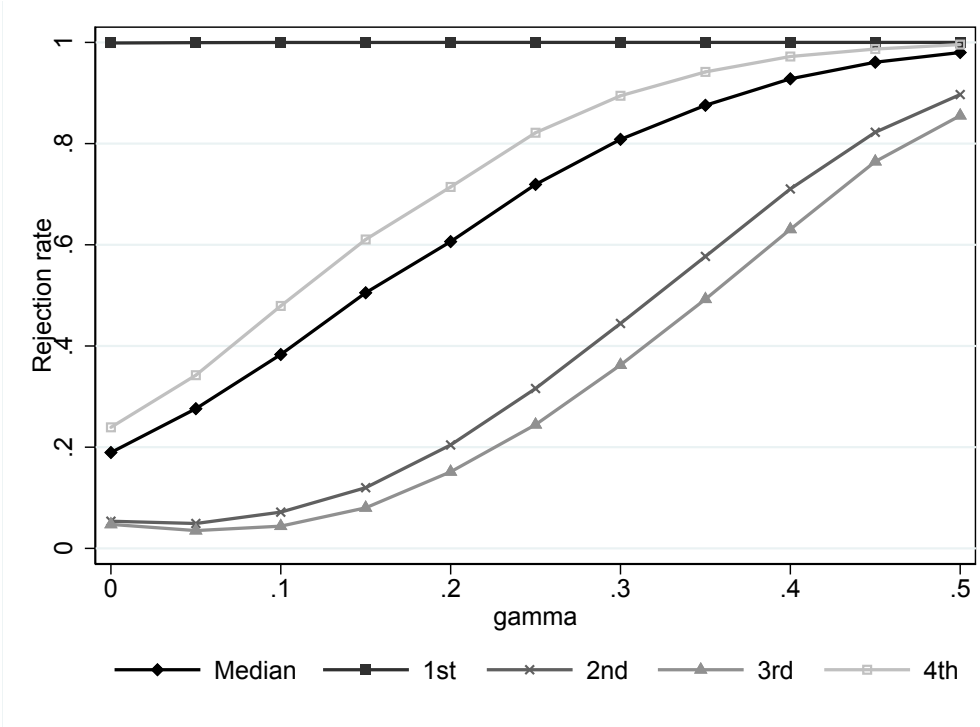



**Figure B.2:** Rejection rates.

# C 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 show that a higher GPA is associated with higher earnings.

Before we turn to the association between GPA and earnings, Figure C.1 shows the overall variation in the GPA. The average GPA in the covered cohorts was 9.27 on a 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.

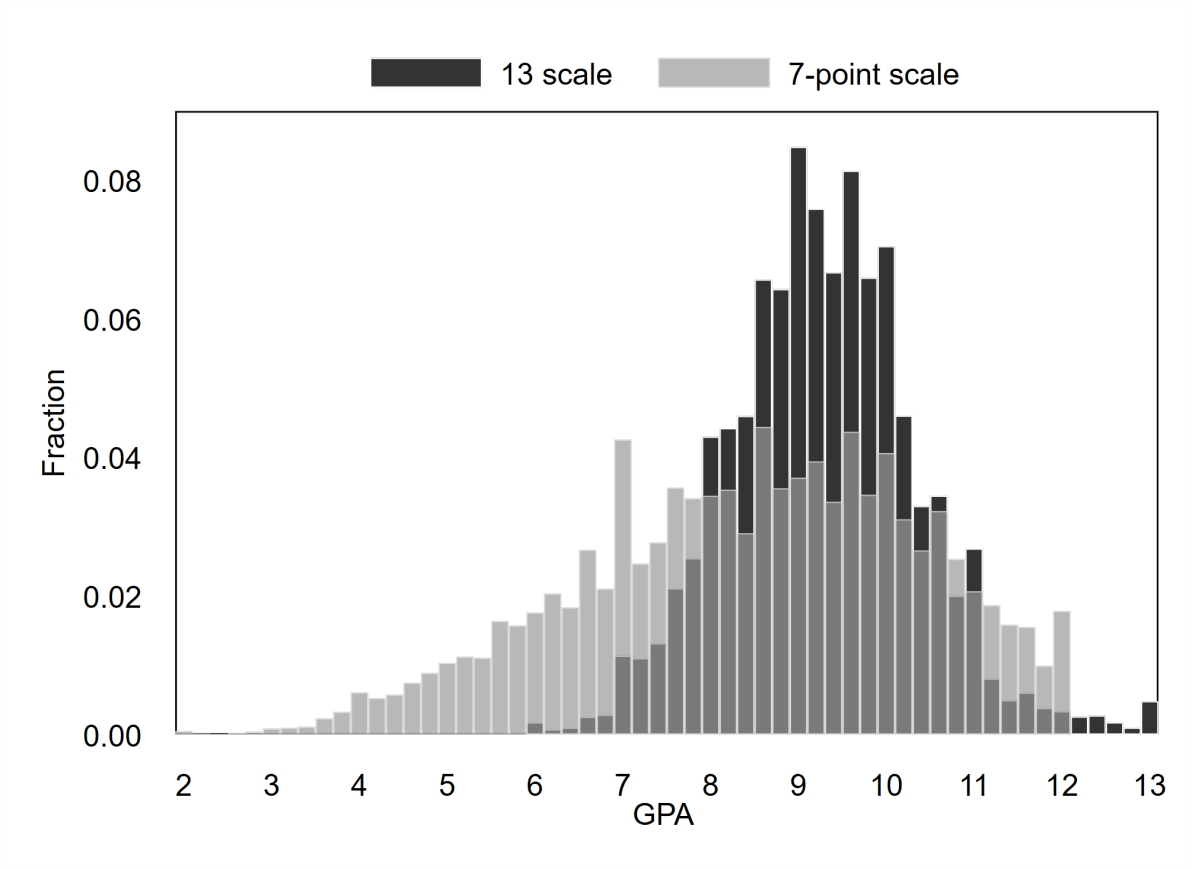



**Figure C.1:** The GPA distribution for university graduates.

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 C.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 7 percent higher earnings. In the fifth calendar year after graduation a one standard deviation higher GPA is associated with four percent higher earnings. By comparison,

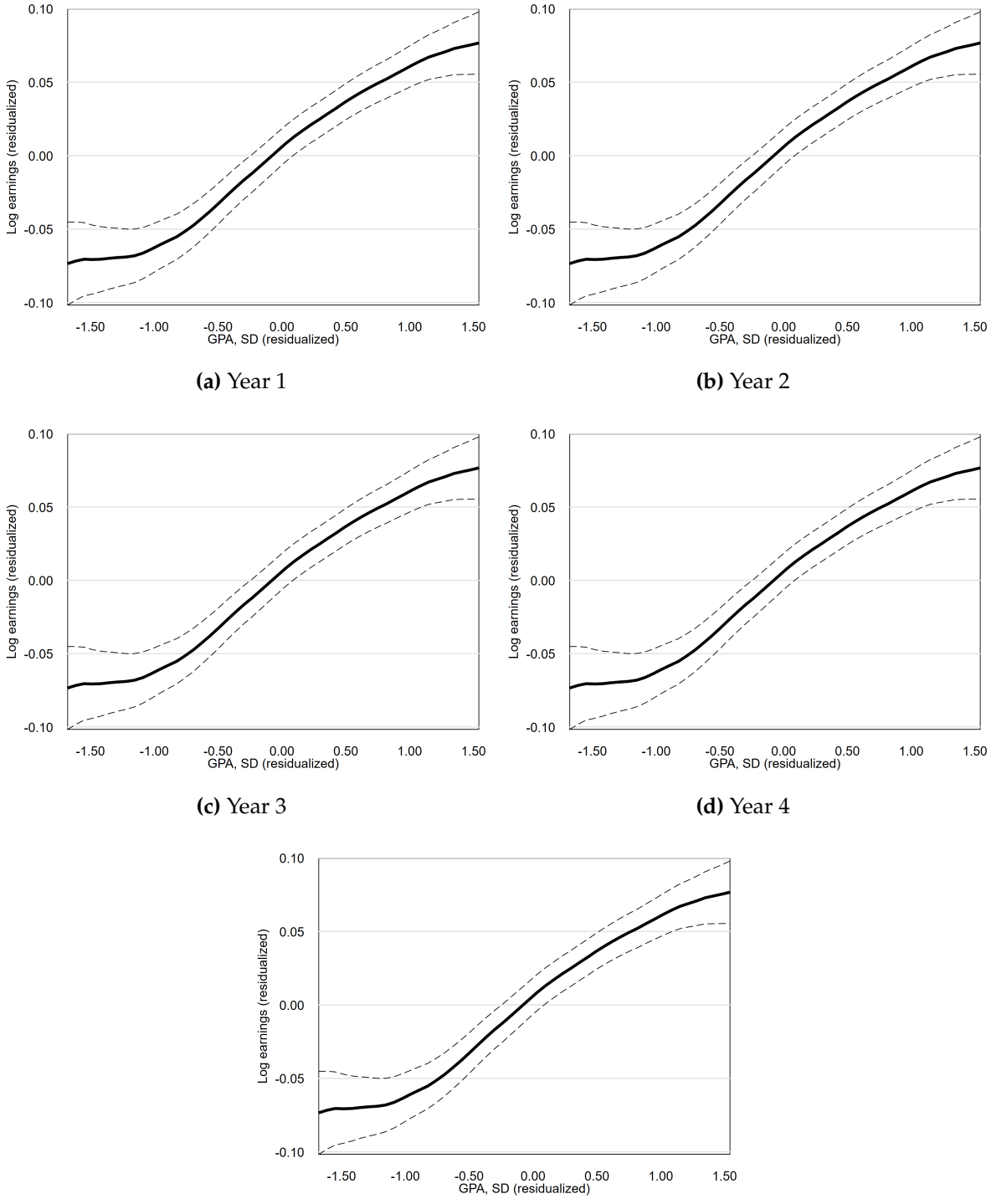
Christensen and Westergard-Nielsen (1999) finds that one additional year of schooling is associated with 4.5 percent higher earnings in Denmark.

**Table C.1:** Regression results. Dependent variable: log earnings

| | Year after graduation | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| All | 0.066*** | 0.057*** | 0.057*** | 0.049*** | 0.041*** |
| | (0.014) | (0.010) | (0.009) | (0.010) | (0.011) |
| *N* | 37217 | 37596 | 37390 | 37127 | 36807 |
| Pre 2007 | 0.071*** | 0.048*** | 0.056*** | 0.047*** | 0.037*** |
| | (0.022) | (0.013) | (0.011) | (0.013) | (0.013) |
| *N* | 14337 | 14560 | 14583 | 14493 | 14384 |
| Post 2006 | 0.063*** | 0.062*** | 0.058*** | 0.050*** | 0.044*** |
| | (0.011) | (0.010) | (0.009) | (0.011) | (0.012) |
| *N* | 22880 | 23036 | 22807 | 22634 | 22423 |

Notes: The table shows the coefficients from regressing log earnings in years 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 parentheses. Asterisks indicate significance at the following levels: * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

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




**(a)** Year 1

**(b)** Year 2

**(c)** Year 3

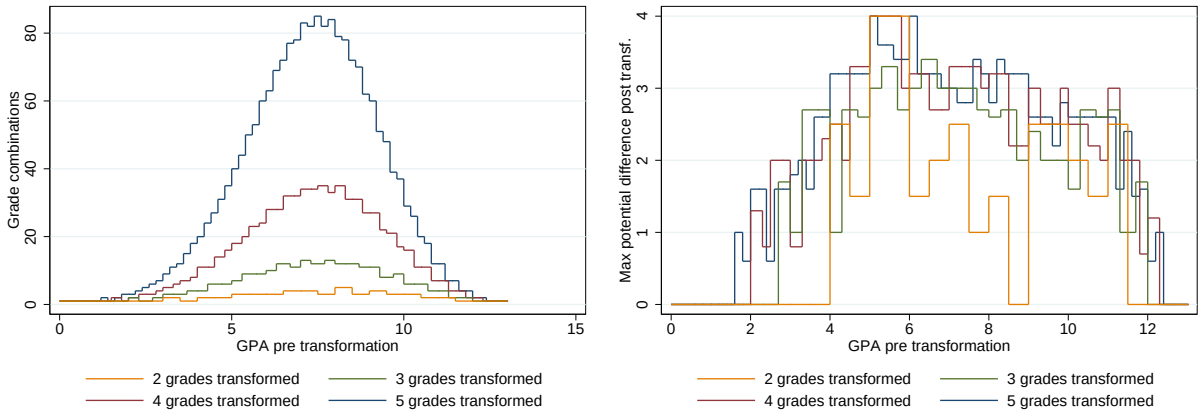
**(d)** Year 4

**(e)** Year 5

**Figure C.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.

# D 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.

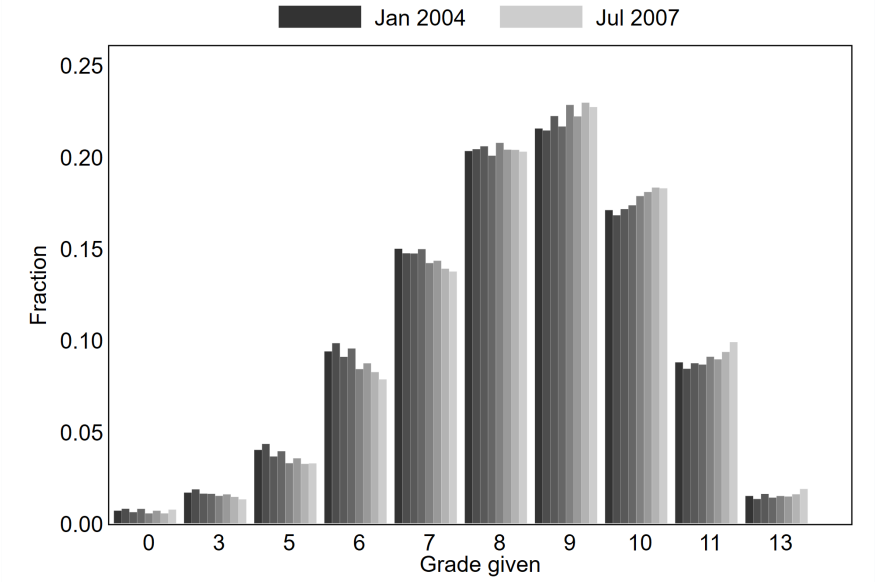



**Figure D.2:** Relative grade frequency.

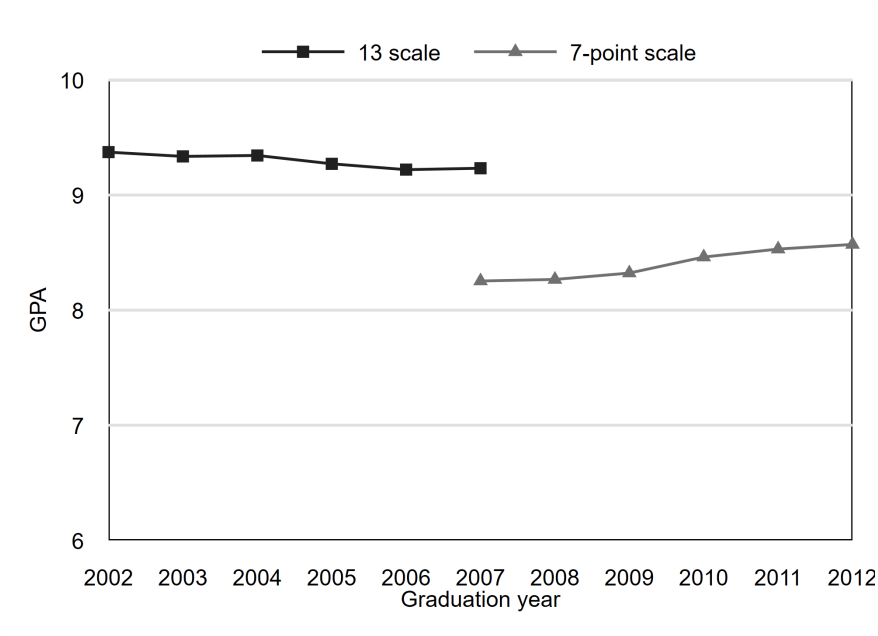



**Figure D.3:** Average university GPA by year of graduation.

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.