

Nordin, Martin; Heckley, Gawain; Gerdtham, Ulf-G.

Working Paper

The Impact of Grade Inflation on Higher Education Enrolment and Earnings

Working Paper, No. 2019:1

Provided in Cooperation with:

Department of Economics, School of Economics and Management, Lund University

Suggested Citation: Nordin, Martin; Heckley, Gawain; Gerdtham, Ulf-G. (2019) : The Impact of Grade Inflation on Higher Education Enrolment and Earnings, Working Paper, No. 2019:1, Lund University, School of Economics and Management, Department of Economics, Lund

This Version is available at:

<https://hdl.handle.net/10419/260271>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Working Paper 2019:1

Department of Economics
School of Economics and Management

The Impact of Grade Inflation on Higher Education Enrolment and Earnings

Martin Nordin
Gawain Heckley
Ulf-G. Gerdtham

January 2019



LUND
UNIVERSITY

The impact of grade inflation on higher education enrolment and earnings

Martin Nordin^a, Gawain Heckley^b, and Ulf Gerdtham^c

2019-01-10

^aDepartment of Economics and Agrifood Economics Centre, Lund University, Sweden

^bHealth Economics Unit, Lund University, Sweden

^cDepartment of Economics and Health Economics Unit, Lund University, Sweden

Correspondence to:

Martin Nordin

Department of Economics and AgriFood Economics Centre, Lund University

Box 730 – Scheelevägen 15 D

220 07 Lund

Sweden

Email: martin.nordin@agrifood.lu.se

Phone: +46 (0) 46 222 07 90

Abstract

Although grade inflation is unfair and may imply inefficient allocation of human resources, current knowledge of grade inflation effects on individual outcomes is scarce. One explanation is probably the challenge of measuring and estimating causal grade inflation effects. This study examines the consequences of grade inflation at the upper secondary education level on enrolment in higher education and earnings for Sweden. Rigorous diagnostic testing supports our empirical approach. Grade inflation at the school level affects earnings mainly through choice of university and the chosen field of education, rather than through enrolment per se, because attending universities of higher quality and pursuing high-paying fields of education have a substantial impact on earnings. On the other hand, high-skilled students attending upper secondary schools without grade inflation and, unexpectedly, low-skilled women attending “lenient” schools are harmed by this. This causes extensive unfairness and, plausibly, detrimental welfare effects.

Keywords: grade inflation, upper-secondary education, higher education, earnings

JEL-codes: I2, I21, J24

1 Introduction

Grade inflation in the education system exists in many countries. For example, Hurwitz and Lee (2018) show for the US that the proportion of high school students with top grades rose from 39% to 47% in 1998-2016, during a period when SAT scores declined. In Sweden, substantial grade inflation appeared at both the compulsory schooling and upper secondary level when the grading system changed from a relative to a goal- and criterion-referenced grading system in 1997 (Cliffordson, 2004; Wikström and Wikström, 2005; Gustafsson and Yang Hansen, 2009; Björklund et al., 2010). At the same time, school competition in Sweden increased due to a reform giving students the right to choose a school outside their local catchment area. At the upper secondary school level, grades increased by almost 10 percentage points between 1997 and 2003, without a corresponding increase in standardized tests such as PISA.

This study investigates the impact of upper secondary grade inflation on enrolment in higher education and earnings for the period of massive grade inflation in Sweden. An inflated grade captures a specific grade variation that changes the ranking of students without changing the skill distribution per se. We argue that a potential positive impact of inflated grades is accompanied by a negative impact for those harmed by grade inflation.

Grade inflation is a policy concern because it can cause mismatches in the labor market and can be unfair. If grade inflation reduces the informative signal of grades, the distribution of skills in the population may be allocated inefficiently. Additionally, grade inflation leads to a change in the ranking of individuals, affecting the probability of being accepted for high-paying fields of education or higher-quality universities. However, a positive effect can also be expected: grades may affect perceived self-efficacy, thereby

having a real effect on educational attainment. Feedback has been shown to have a positive impact on subsequent test scores (Bandiera et al., 2015).¹ Nevertheless, grade inflation generates inequities between and within cohorts and, since school grades are a strong predictor of university achievements in Sweden (Björklund et al., 2010), university selection based on incorrect grading is likely to impose welfare costs. However, to our knowledge, the welfare costs of grade inflation have not been assessed empirically. We do not calculate the welfare costs in this study, but suggestive evidence shows that they may be substantial.

Grade inflation is a consequence of misaligned incentives. Grades are inflated for the benefit of students, teachers, and schools, but at the expense of the common good, since: i) schools may receive increased funding and attract better students at the expense of other schools, ii) teachers may get better evaluations at the expense of other teachers, and iii) students receive better grades at the expense of students attending schools without or with little grade inflation. Because grades are bounded from above, grade inflation also means more concentrated grades at the upper tail. It has been shown that inflated grades help mediocre students at the expense of good students (Chan et al., 2007).

To measure grade inflation, we compare students' grades with results from the Swedish Scholastic Assessment Test (SweSAT), which is taken and administered outside schools. Importantly, we argue that the SweSAT does not suffer from the same misaligned incentives as school grades and therefore represents a measure of skills not impacted by grade inflation. Around 55 percent of each student cohort takes SweSAT. By aggregating the difference between upper secondary school grades and the SweSAT results at school level, we use the change in schools' grade inflation over time to identify a grade inflation

¹Studies by Wuepper and Lybbert (forthcoming), Bandiera et al. (2015), Diamond and Persson (2017) and Hvidman and Sievertsen (2017) all show different motivational effects of past achievements.

effect. A caveat is that school grades capture a broader set of skills than SweSAT, which mainly captures cognitive skills. Thus the difference between school grades and SweSAT at the *individual* level is not a measure of grading leniency. At the school level, however, systematic divergence between grades and SweSAT results is a strong indication of grade inflation. Because systematic sorting of pupils into upper secondary schools based on unobserved non-cognitive skills (i.e., not related to parental/migration background and compulsory grades, which we control for) is unlikely, the main biasing factor is changes in schools' production of non-cognitive abilities over time. However, our results show no evidence of systematic changes in schools' production of skills. Our approach is a development of that described by Wikström and Wikström (2005), who investigate the relationship between school competition and grade inflation.²

We show that grade inflation mainly impacts investment in higher education. To some extent inflated grades enhance university achievements, suggesting a self-efficacy effect, but the main impact is on the choice of university and field of education. Then, indirectly, a better university and a higher paying field of education implies higher earnings. The signaling value of school grades therefore appears to be limited to university selection, as the direct impact on earnings is small.

The major contribution of our study is that we consider the consequences of exogenous re-ranking of individuals' grades. Such re-ranking may have significant welfare costs because of its detrimental impact on higher education as a "sorting mechanism". We show that highly skilled students at schools without grade inflation are harmed by this re-ranking and that low-skilled women attending "leniently grading"

²They also use the score difference between SweSAT and grades to determine grade inflation, but they lack panel data so their results are based on between-school differences, and not within-school differences.

schools are also affected negatively. The latter finding may be due to a negative self-efficacy effect (Bandura, 1997), as a consequence of the relative downgrading of these women's school grades.

2 Earlier research

In Sweden, the main explanation given for the existence of grade inflation is school competition (Wikström and Wikström, 2005; Vlachos, 2010). However, it could also be caused by students' course choices (Sabot and Wakeman-Linn, 1991) or by teachers wanting good evaluations (Redding, 1998; Johnson, 2003). Grade inflation might also be related to gender or ethnicity. Studies comparing blind and non-blind grading report mixed evidence on gender discrimination. Boys seem to be discriminated against in Israel (Lavy, 2008), but not in Sweden (Hinnerich et al., 2011). In Sweden, Diamond and Persson (2017) report that grading leniency is not related to discrimination on grounds of gender or immigration.

A number of studies have investigated the incidence of grade inflation (see e.g., Wikström and Wikstrom, 2005; Bar et al., 2009; Bauer and Grave, 2011; Butcher et al., 2014; Rey and Looney, 2016). However, only a few studies have investigated the impact of grading leniency on subsequent school achievements. For example, Manacorda (2012) shows that grade retention affects dropout and lower educational attainment up to four to five years after grade failure. However, to our knowledge only Diamond and Persson (2017) and Bagues et al. (2008) estimate the impact of grade inflation on earnings.

Diamond and Persson (2017) use an innovative bunching estimator to identify grading leniency in grade nine in Sweden. Based on data from a nationwide standardized test in mathematics, they show that teachers manipulate the mathematics test score of

students who are close to a grade threshold and experience “a bad day” in relation to their performance in nationwide tests in Swedish and English taken a couple of weeks before the mathematics test. A key distinction between our study and that by Diamond and Persson (2017) is that those authors consider two narrow points of the skills spectrum, individuals just below the ‘pass’ and just below the ‘pass with distinction’ thresholds, whereas we consider the whole skill distribution, including students negatively and positively impacted by the re-ranking.

Diamond and Persson (2017) show that manipulation increases compulsory and upper secondary GPA in Sweden, the probability of completing upper secondary education in time (at age 19), the probability of enrolling in higher education, and earnings (at age 23). However, in the Italian university system, where funding is based on student performance, grading leniency is negatively associated with employment and wages (Bagues et al., 2008), plausibly due to selective grading leniency. A related study by Maurin and McNally (2008), who investigate exogenous lowering of examination thresholds in higher education in France³, shows that grading leniency enables pupils to invest in more education than would otherwise have been possible and that it increases earnings substantially.

3 Grades, grade inflation, and SweSAT

3.1 Grade inflation and school competition

With the introduction of the goal- and criterion-referenced grading system in 1994, grades obtained by school leavers in Sweden started to increase. The first cohort with new grades graduated in 1997. Figure 1 shows separately the increase in average grades for men and

³Maurin and McNally (2008) use the student riots in 1968, where normal examination procedures were abandoned.

women studying on an academic or a vocational track⁴. For academic track students the general grade inflation stopped in 2003, but for vocational track students it continued for another year. The increase in average grades between 1997 and 2003 was about 30 percent larger than the (substantial) gender difference in grades. Because a corresponding increase in knowledge has not been found (Cliffordson, 2004; Gustafsson and Yang Hansen, 2009; Björklund et al., 2010) and because international tests such as PISA, TIMSS, and PIRLS actually report a decrease in mathematics, science, and reading comprehension (Skolverket 2012), the consensus is that the increase is due to grade inflation. Grades stopped increasing after 2004, probably due to growing awareness of the phenomenon, which culminated in a review of the goal- and criterion-referenced grading system by the Swedish National Audit Office in 2004.

Figure 1 about here

For our study population (students enrolled on an academic track) we assess the grade inflation for different parts of the grade distribution. Figure 2 we shows the relative change in the share of students with a grade above 12, above 15, and above 18 (where 20 is the top grade), which about 70%, 30% and 5% of students receive in 1997, respectively. As can be seen, the relative increase is much larger for the group of students obtaining a grade above 18 (an increase of almost 250%) than, in particular, for the group obtaining a grade above 12 (an increase of only around 20%).⁵ Moreover, whereas average grades remained constant after 2003, the share of students with top grades (over 18) kept increasing until 2006.

Figure 2 about here

⁴Others reporting an increase in average grades are e.g., Gustafsson and Yang Hansen (2009) and Vlachos (2010).

⁵Figure A1 shows that the grade inflation is due to a sharp increase in the highest grade (pass with special distinction). Fewer courses are graded with just a pass, and for women the share of courses with the second highest grade (pass with distinction) also decreased.

The introduction of the new grading system made grade inflation possible. In the former relative grade system, where grades were intended to follow a normal distribution on national level, substantial grade inflation could not exist. Swedish grade setting is based on classroom assessment and the teacher alone sets the grade. Standardized tests are given in Swedish, English, and mathematics, but the tests are not high-stakes examinations and function only as an aid for scale calibration. Thus, with incentives favoring grading leniency, grade inflation could definitely have flourished in the Swedish schooling system.

The incentive (competition between schools) was also amplified during the same period (Wikström and Wikström, 2005). From the academic year 1992-1993 onwards, all Swedish students gained the right to choose a school outside their local catchment area, with student funding moving with the student. This school reform increased school competition and the number of private schools in Sweden increased dramatically. In upper secondary education, the number of private schools increased by almost 200 percent between 2001 and 2009, and the share of students choosing a private school increased from 2 to about 20 percent in 2009 (Skolverket, 2012). However, according to Vlachos (2010), school competition is a general phenomenon and private schools are only responsible for a minor part of the grade inflation.

3.2 Grades and SweSAT

There are two ranking systems to decide who is admitted to higher education in Sweden. Individuals are separately ranked according to their average upper secondary grade and their score in SweSAT. Almost 30 percent of those accepted into higher education are accepted on the basis of SweSAT.

The upper secondary grading system contains four grade levels: no-pass, pass, pass with distinction, and pass with special distinction. The final grade is calculated as followed:

$$\sum_{m=1}^M \frac{\text{Course credits}_m \times \text{Grade points}_m}{\text{Course credits for full program}}$$

where M stands for the number of courses and the grade points allocated are: 10 points for pass grades, 15 points for pass with distinction grades, and 20 points for pass with special distinction grades. From 1997 to 2002, a full upper secondary program consisted of 2150 course credits for academic tracks and 2370 course credits for vocational tracks. In 2003 the program structure was revised, so a full program now consists of 2500 course credits for both types of track.⁶ A pass grade is not a requirement for getting a final grade; to achieve a final grade, completing three years of upper secondary education is necessary. However, to be eligible for tertiary education, students need pass grades in 90 percent of their courses.

SweSAT functiones as an instrument of selection into higher education. The test is optional and administered twice a year by the Swedish Council for Higher Education (i.e., it is administered outside schools). SweSAT is similar to the American SAT (Wikström, 2005) and captures cognitive abilities (Carlstedt and Gustafsson, 2005). In 1996-2011, the test consisted of five subtests: Swedish reading comprehension (READ), English reading comprehension (ERC), vocabulary (WORD), data sufficiency (DS), and interpretation of diagrams, tables and maps (DTM). The difficulty level of SweSAT varies between years and the test score is therefore normalized to a scale between 0.0 and 2.0 (Stage and Ögren, 2009).⁷

⁶Before 2003, a course could be of any size and schools offered a wide variety of courses. From 2003, courses gave either: 50, 100, 150, 200, or 250 course credits (with some exceptions of even larger courses) and the programs contained a larger share of compulsory courses.

⁷In the econometric specification, national time dummies capture variations in difficulty level.

SweSAT can be taken several times, including before finishing upper secondary education, and as a selection instrument to higher education only the best result counts. Repeated test taking has a positive effect on the test result (Henrikson and Törnkvist, 2002) and therefore both the number of times the test is taken (learning effect) and the timing (age effect) have to be considered when computing a ‘cleaner’ measure. We return to this issue in section 6, where we explain how we construct schools’ mean SweSAT.

4 A model of skill signaling error on labor market outcomes

In this section, we set up a model of how grade inflation in the Swedish education system can be isolated and linked to labor market outcomes. We note that the school grade of student i , taught at school j , is a combination of cognitive skills, C , non-cognitive skills, NC (e.g., motivation, self-control, conscientiousness, and organizational skills), school’s educational investment, E , and grade inflation, π . Our outcome variables, Y , are thus given by the following function:

$$Y_{ij} = S_{ij}^{Grade}(C_{ij}, NC_{ij}, E_j, \pi_j) \quad (1)$$

To provide some structure, we assume that school grades are a linear combination of cognitive and non-cognitive skills, plus an error term:

$$S_{ijt}^{Grade} = C_{it}(C_{it-3}, E_j, \pi_{jt}) + NC_{it}(NC_{it-3}, E_j, \pi_{jt}) + \varepsilon_{ijt} \quad (2)$$

The educational investment E (three years of upper secondary education) explains the progression of skills, C and NC , between time $t-3$ (when beginning upper secondary education) and time t (when graduating from upper secondary education). The school’s assessment of both C and NC at time t is also determined by grade inflation π . We do not observe each school’s level of grade inflation. However, using different scholastic measures we do observe each individual’s cognitive skills, C_i , and non-cognitive skills,

NC_i . We also observe some proxy measures of each school's educational investment function, E_j . In principle, we are therefore able to back out π under the structure outlined in model (2).

To do this, we use SweSAT, a standardized test not subject to the incentives driving grade inflation in school grades. To isolate school level grade inflation, we assume that the SweSAT measures the school level skills in the same way as upper secondary school grades do, yielding:

$$\overline{S_{jt}^{Grade}} - \overline{S_{jt}^{SweSAT}} = C_{it}(\pi_{jt}) + NC_{it}(\pi_{jt}) \quad (3)$$

The difference between average upper secondary school grades and average SweSAT scores can be considered an upper bound of grade inflation. However, it may not be valid to assume that SweSAT represents the full set of skills captured by upper secondary school grades. To deal with this problem, we use a control strategy where we also condition on compulsory school grades (capturing C_{it-3} and NC_{it-3}). This assumes that any divergence between school grades and SweSAT is constant across cognitive and non-cognitive skills as distributed at the beginning of upper secondary school, at $t-3$ ⁸. In the results section, we test our method of isolating grade inflation. By the inclusion of information on school cohort's migration share and mean parental background, we show that the variation in school grade inflation is not linked to student background characteristics.

Next, we have to consider that school production, E , may vary between schools and over time. The available evidence is that school resources decreased, or stayed the same, during the period (the share of teachers with a teaching qualification and teachers' (real)

⁸On the other hand, conditioning on compulsory school grades may underestimate grade inflation. Thus, there is potentially a selection problem in that students may choose schools on their perception of π_{jt} which correlate with background characteristics that explain our outcome variable. We could have conditioned on compulsory grades already in (3), but for this reason we prefer to show the impact of controlling for compulsory grades, separately.

wages have decreased and the teacher/pupil ratio has remained constant). Moreover, there has been no major change in the variation in school resources between schools (Holmlund et al., 2014). Thus, while a change in school quality is a potential problem, the change is small and should have a minor impact on grades and our measure of grade inflation. In the main model specification, we include school fixed effects to control for variation in school quality across schools and therefore only consider within-school variation. This leaves potential changes in school quality and resources over time. To address this, we also control for a set of resource-related school characteristics and school-specific linear time trends, and in a sensitivity test we control for the individual progression of grades between junior and senior school years.

In summary, we combine two approaches to isolating grade inflation. The first is a differencing approach using a standardized test as a benchmark for actual skills, the difference between this and school grades being a measure of grade inflation. This is combined with a control strategy that, according to our model of grade inflation, should also be able to isolate grade inflation. We therefore combine two grade inflation isolation strategies in order to provide convincing evidence of the impact of grade inflation on enrolment in higher education and earnings.

5 Data

We use a population sample of pupils receiving a final grade from upper secondary education in the years 1997 to 2004⁹. By using SweSAT, which is taken by only about 15% of students on a vocational track, to construct our measure of grade inflation, we

⁹We have data for younger cohorts too but, because the main outcome variable in this study is income, a reasonable amount of time must pass between graduating from upper secondary education and when income is measured, i.e., the individual has to have time to finish higher education and enter the labour market.

restrict the sample to students on an academic track¹⁰. We also restrict the sample to individuals who have attended Swedish compulsory education. The studied cohorts are, principally,¹¹ born between 1978 and 1985, and the sample consists of 280,364 individuals.

With matched education data and labor market data obtained from Statistics Sweden (SCB), we follow these cohorts up until 2013. The register of final grades from compulsory school, the register of grades from upper secondary school, and the register of higher education grades describes the educational attainment of these individuals, while the Longitudinal Integration Database for Health Insurance and Labor Market Studies (LISA) describes their labor market outcome. The Multigenerational Register links individuals to their parents, while the population and housing censuses provide data on parental education and income during childhood.

6 The empirical approach

6.1 School-level divergence between grades and SweSAT

As a first step, we calculate each individual's mean SweSAT, because the mean is likely to have better precision than a single test score¹². Thus, we use the mean residual SweSAT when considering both the timing and the sequence of test taking. First, we estimate the test score S_{ist}^{SAT} :

¹⁰For each cohort, around 23 percent are not in the register of grades from upper secondary school (but they are in the LISA database) because they fail to achieve a final grade from upper secondary education. The majority are drop-outs from upper secondary education and only a minor share leave school after compulsory education.

¹¹Since some students (9.2 percent) finish upper secondary education at an older age than the typical graduation age of 19 and a small share (1.2 percent) finish at a younger age, there are some older and younger individuals in the data. However, removing individuals not graduating at age 19 has no impact on the results in this study.

¹²An alternative would be to use the first test score, but for some individuals this is taken before finishing upper secondary education.

$$S_{ist}^{SAT} = \alpha + \gamma_{is} + \delta_{it} + S_{ist}^{SAT_res} \quad (4)$$

on two sets of dummy variables, γ_{is} and δ_{it} , and save the residual test score, $S_{ist}^{SweSAT_res}$. γ_{is} indicates whether the test is being taken for the first ($s=1$), the second ($s=2$), or the s :th time. The test is given two times a year (in spring and fall) and δ_{it} indicates the time before or after finishing upper secondary education. However, since we want the test score to be measured reasonably close to the graduation date, we use only tests taken in the spring before graduating upper secondary education in June, or tests taken up until three years after graduating upper secondary education. Second, we calculate the mean residual test score $\overline{S_{it}^{SweSAT_res}}$ for each individual and standardize this measure. We thus have a single test score measure for each individual, cleansed from the learning and the age effect. The next step is to compute the school-level divergence between grades and SweSAT. To do this, we regress the upper secondary grade on $\overline{S_{it}^{SweSAT_res}}$:

$$S_{it}^{Grade} = \overline{S_{it}^{SweSAT_res}} + D_{it} \quad (5)$$

and save the residual, D_{it} . The standardized version of this measure is our individual divergence between the grade and SweSAT. By averaging up this measure to school cohort level, we obtain, $\overline{D_{jt}}$, a measure that captures the difference between a school cohort's cognitive ability level (measured by SweSAT) and the cohort's average upper secondary grade. Because grade inflation may differ between men and women (Hinnerich et al., 2011), the measure is averaged up separately for men and women.

Figure 3 shows the yearly mean standardized grade and the national mean grade divergence, $\overline{D_t}$, for the population taking the test. These measures follow each other closely. However, in comparison with the overall change in grades (shown in Figure 1), we observe increasing school grades until 2006 in the test-taking population. This is

probably due to selection in test taking (where test takers have higher grades on average), in combination with continued grade inflation at the top of the grade distribution. Selection in test taking is a potential concern, but in section 7 we analyze the problem and find that it has no impact on the results in this study.

Figure 3 about here

The somewhat higher slope observed for \overline{D}_t indicates that grade inflation was higher than the observed increase in grades in the study period. This is in line with the PISA, TIMSS, and PIRLS test results, which show that while average grades have increased, knowledge has in fact decreased.¹³

6.2 The econometric model

With the following model, we use \overline{D}_{jt} to determine the probability of individual student i investing in higher education and earnings, Y_{ijt} :

$$Y_{ijt} = \alpha_j + \delta_t + \overline{D}_{jt} + \sigma \overline{P}_{jt} + \gamma \overline{M}_{jt} + \theta \overline{CG}_{jt} + \overline{GB}_{jt} + \pi_k TK_{jtk} + \rho X_i + \varepsilon_i \quad (1)$$

In this model, α_j and δ_t are school and finishing year fixed effects, respectively. The finishing year dummies capture the national cohort variation in educational attainment and earnings.¹⁴ The school fixed effect means that we estimate a within-school grade inflation effect.

¹³With a yearly normalized test score, a yearly variation in knowledge cannot be picked up by the yearly tests alone. However, since we compute the individual grade inflation from several tests, the knowledge variation is picked up. For example, if a person takes the test in two consecutive years, an equally good achievement will be better rewarded in the second year if the mean knowledge level has decreased in the test-taking population. Changes in the general knowledge level thus affect the mean individual test score, $\overline{S}_{it}^{SweSAT-res}$.

¹⁴We also tried including birth cohort dummies and late graduation year, but these had no impact on the model or the grade inflation effect.

With mean compulsory grade, \overline{CG}_{jt} , parental income and education, \overline{P}_{jt} , migration share, \overline{M}_{jt} , and mean, we control for changes in the inflow. Even if we use a gender-specific grade inflation measure, we also control for the gender balance, \overline{GB}_{jt} .

Moreover, by including the share of students choosing j different tracks, TK_{jtk} ,¹⁵ we capture differences in the difficulty level of passing courses on certain tracks. If some tracks are easier, a larger inflow into these tracks may imply increased grades at the school. To capture the school production, we include the teacher/pupil ratio, share of teachers with a teaching qualification, and share of permanently employed teachers (measured at municipal level by the Swedish National Agency for Education). School-specific linear time trends are included as an additional sensitivity test.

To get a reliable measure of the school-level grade inflation, we use a sample of schools that contain at least two classes per cohort, i.e., these schools should contain at least 320 pupils finishing upper secondary education during the period (with eight cohorts, the average number of students per cohort is 40, clearly over the typical class size in Sweden of around 30 students). The mean is 188 pupils per cohort, i.e., around six classes per school.

7 Results

7.1 Investigating the grade inflation measure

To begin with, we investigate \overline{D}_{jt} . Table 2 shows the impact of different sets of control variables on \overline{D}_{jt} . In column (1), for males, and (5), for females, we find that 47% and 58% of the within variation in \overline{D}_{jt} is explained by a national time trend. This demonstrates a

¹⁵The share finishing a certain track in relation to the total number of students attending the school, i.e., including pupils on both vocational and academic tracks. This means that changes in selection into academic and vocational track are also controlled for.

common increase in grades. In columns (2) and (6), where we include the school-level variation in gender balance, compulsory grades, and track, the within R^2 increases. It is mainly compulsory grades that matter: inflow, measured as the change in mean compulsory grades, has a clear impact on $\overline{D_{jt}}$. A larger (smaller) share on science (social science) tracks also has a negative (positive) impact on D_{jt} (but the impact of track on the within R^2 is small). This is expected, as science is known to be more difficult than social science and it has an impact on grading. Columns (3) and (7) show that changes in the inflow, measured at the school level with parental background and migration share, has no impact on $\overline{D_{jt}}$. Finally, in columns (4) and (8) we find that our school characteristics have no effect on $\overline{D_{jt}}$. Thus, the variation in $\overline{D_{jt}}$ is mainly related to differences in the inflow of skills (measured by mean compulsory grades) and partly to how students allocate themselves between different tracks. Importantly, the variation in $\overline{D_{jt}}$ is not related to socioeconomic factors (parental income and education) and migration status, indicating that, given control for compulsory school grades, there is no evidence of students self-selecting to schools with higher grade inflation. Also, the Swedish compensatory system for allocating resources (more resources go to schools in segregated areas) therefore does not explain the grade inflation. This is supported by Holmlund et al. (2014) that resources are not allocated in a more compensatory way over time. We conclude that the remaining variation consists of random noise and within-school variation in grade inflation. The sensitivity results in section 8 support this conclusion.

Table 2 about here

7.2 Estimating the grade inflation effect on higher education and earnings

Table 3 shows the effect of $\overline{D_{jt}}$ on enrolment in higher education and earnings. The effect corresponds to an increase in grade inflation of one standard deviation, which is not

uncommon. For around 16% of schools, the increase in \overline{D}_{jt} between 1997 and 2003 was above one standard deviation.

On comparing the columns, we find (again) that mean compulsory school grade has an impact on the results: on including compulsory school grades (in columns (2) and (8)) the effect of \overline{D}_{jt} decreases significantly, particularly for enrolment. For men, column (2) shows that a standard deviation increase in \overline{D}_{jt} increases the probability of enrolling in higher education by around 2.9%. For women, column (6) shows no impact of \overline{D}_{jt} on enrolment. For earnings, the effect of \overline{D}_{jt} is 175.5 and 110.7 for men and women, respectively. These estimates correspond to impacts of 5.1% and 4.5% (measured at mean earnings for each gender). Family background, gender, and migration share (added in columns (3) and (9)), the school quality indicator (added in columns (4) and (10)), and school specific linear time trends (added in columns (5) and (11)) have no impact on the effect of \overline{D}_{jt} .

Table 3 about here

It is interesting to compare the \overline{D}_{jt} effect with the effect of upper secondary grades (standardized and aggregated at school level) on enrolment and earnings. Conceptually, the difference between estimates is the bias in the causal grade effect on enrolment and earnings. For enrolment, we find in columns (6) and (12) that the grade effect is much larger than the effect of \overline{D}_{jt} , but for earnings the estimates are more comparable. In fact, for men the \overline{D}_{jt} effect and the grade effect on earnings are similar. This indicates that cognitive ability has a large impact on enrolment (and therefore there is a large difference in estimates on removing the cognitive ability variation with SweSat), but that cognitive ability matters less for earnings (hence the similar estimates).

Figures 4 and 5 show the effect of \overline{D}_{jt} on enrolment and earnings longitudinally, i.e., each year after graduation, respectively. For enrolment in higher education, we find that the effect is large one to three years after graduation (Figure 4). The effect then decreases with time and already four years after graduation D_{jt} has no effect for women's enrolment. Seven years after graduation, the effect becomes insignificant for men and negative for women. For women, this explains the insignificant effect seen in Table 3. Thus, in the long run, grade inflation has no impact on higher education for women, but grade inflation makes women enroll earlier. For men, the \overline{D}_{jt} effect becomes negative eventually, but the overall effect on enrolment is still positive. For earnings (Figure 5), the effect of \overline{D}_{jt} shows up six to seven years after upper secondary education (i.e., around the time when students start entering the labor market after having finished higher education) and after 14-15 years the effect is almost SEK 40,000 for men and SEK 15,000 for women. Thus, for men the long-term effect is more than twice as large as the overall effect in Table 3, but for women the effect does not increase with time to the same extent.

Figures 4 and 5 about here

7.3 Investigating the mechanism behind the grade inflation effect

In Table 4, we further investigate the effect of \overline{D}_{jt} . First, in columns (1)-(2) for men and (3)-(4) for women, we estimate the effect on years of schooling and the probability of obtaining a degree¹⁶ for the sample who enroll in higher education. These outcomes reflect overall university achievements. If \overline{D}_{jt} captures grade inflation and not scholastic skills, the \overline{D}_{jt} effect on performance is possibly a self-efficacy effect. For years of

¹⁶A degree here means completing at least 15 years of schooling, which is equivalent to reaching the number of tertiary course credits required for a Bachelor's degree. However, without fulfilling the program-specific requirements (including a Bachelor's project or dissertation) a Bachelor's degree is not necessarily achieved.

schooling, we see a small effect (one standard deviation of grade inflation increases schooling by around 0.1 year), but for the probability of getting a degree the impact is larger (2-3%).

Next, we examine the effect of \overline{D}_{jt} on earnings for the group who enrol in higher education (columns (5)-(6)) and the group who do not enroll (columns (7)-(8)). We find that the grade inflation effect is mainly driven by the group who enroll in higher education (and therefore the particularly large impact on earnings in this sub-sample).

Table 4 about here

An inflated grade increases the probability of being accepted to a school of higher quality or to a high-paying field of education. As a next step, we therefore condition the sample on enrolment and estimate the effect of \overline{D}_{jt} on: i) the probability of going to an old (more reputable) university¹⁷, ii) students' mean upper secondary grades at these universities (indicating higher quality), and iii) students' mean upper secondary grades in different fields of education. As Table 5 shows, \overline{D}_{jt} has a clear impact on the choice of university and field.

Table 5 about here

In Table 6, we investigate the effect of \overline{D}_{jt} on the probability of investing in different fields of education. We find that an inflated grade increases the probability of studying in high-paying fields such as economics and law, engineering (for men), and medicine, which is shown by Kirkeboen et al. (2016) to have a substantial impact on earnings. Grade inflation also decreases the probability of studying in tracks with low admission scores such as teaching (for women) and computer science. For some reason, grade inflation also decreases the probability of women studying science.

¹⁷The earnings premium has been found to be higher for old universities (Lindahl and Regnér, 2005).

Table 6 about here

We have shown that grade inflation is much higher at the top of the grade distribution (see Figure 2). The impact of grade inflation is therefore, plausibly, larger for students with higher grades. We analyze this hypothesis by estimating the effect of \overline{D}_{jt} for students in different parts of the scholastic distribution. Because upper secondary school grades are endogenous, compulsory school grades are used to classify the students into an upper- and a lower-skill group. For each cohort and each gender, the students are divided into those above and below the national mean compulsory school grade (Table 7). This shows that the \overline{D}_{jt} effect is larger for the group with compulsory school grades above the mean. For men with grades below the mean, the effect on earnings is positive, but for women we actually find a negative impact on enrolment.

Table 7 about here

It is uncertain why there is a negative effect on enrolment for women, but it could be caused by a negative self-efficacy effect (Bandura, 1977). There is extensive research showing negative self-efficacy effects (see e.g., Bandura and Locke, 2003; Wuepper and Lybbert, 2017), where a low belief in one's capability affects one's goals. In this context, attending a school where, primarily, high-skilled students are rewarded with inflated grades means a relative downgrading of one's own achievements (even if the inflation has a positive impact compared with *other* schools). This in turn may be demoralizing and affect beliefs in one's efficacy and thus future goals, here affecting the investment in higher education. A negative effect for women is consistent with findings on the PISA test of lower levels of self-efficacy in mathematics and science among girls (OECD, 2015). This may also explain the negative grade inflation effect on studying science for women.

8 Sensitivity tests

8.1 Selection in test taking

On estimating the relationship between schools' share of students taking SweSAT and \overline{D}_{jt} (with the same main model as before), we find a negative relationship (data not shown). This is an expected finding, since with more grade inflation the incentive to take the test is smaller. If there is selection in test taking, this is a potential problem. For example, if the probability of taking SweSAT is correlated with scholastic ability, high-ability students attending low grade-inflation schools may be overrepresented as test takers. This implies $\text{Cov}(\pi_{jt}, C_{ijt}) < 0$ in the population taking SweSAT, implying a negative relationship between S_{jt}^{Grade} and S_{jt}^{SAT} . That is, for a school without grade inflation, SweSAT value increases because the students taking the test become more selective over time (high achievers see their grade become less competitive). The school's average SweSAT value is therefore endogenous and the grade inflation effect may be biased.

Figure 6 shows that the relationship between test taking and upper secondary grade is hump-shaped. Up to a grade of around 16 (about 25% of the students on academic tracks have a higher grade), the probability of taking the test increases with the grade, while thereafter the probability decreases.

To test whether selection in test taking is a problem, we add to our main model a variable measuring variation in test taking at schools. We also interact the test taking variable with average compulsory grades, \overline{CG}_{jt} (linearly and quadratically¹⁸). Table 8

¹⁸To consider the hump-shaped relationship between the probability of taking the test and upper secondary grade, we include the interaction quadratically. Because upper secondary grade is used to construct the grade inflation measure, we prefer to interact the share of test-taking students with compulsory grades.

shows the results of this test. We find that the share of students taking the test is related to the share enrolling in higher education, indicating a willingness to invest in higher education, but that inclusion of the share taking the test (and the interactions) has no impact on the effect of \overline{D}_{jt} . We conclude that selection in test taking does not affect our results.

Table 8 about here

8.2 Testing whether school production biases the results

To capture a school production effect, in another sensitivity test we add a measure of individuals' progression of grades. For a particularly productive school, grades are, reasonably, lower in junior than in senior years of upper secondary education. We are able to separate out courses taken the first year (about 80% of first-year courses) and, by measuring the difference in grades between the first-year courses and the second- and third-year courses, we try to capture the progression in grades, i.e., school production. We find that adding grade progression has only a minor impact on our \overline{D}_{jt} effect (Table 9). However, grade progression has a positive impact on enrolment for both genders and earnings for men (but, strangely, a negative effect on earnings for women).

Table 9 about here

Finally, we include the inflation in core subjects as a sensitivity test. In mathematics, Swedish, and English, there are national tests. These tests probably decrease the grade inflation in these core subjects (although the results are used only for calibration). However, if our measure of grade inflation is mainly caused by variations in the student body or school production, such variation is likely to be captured by the core subjects too.

Thus, by including separate grade inflation measures for each of the core subjects (same construction as for the overall grade inflation measures), we assess whether our results are caused by true grade inflation or not. A decrease in the $\overline{D_{jt}}$ effect indicates that our measure of grade inflation is biased. However, the results show that adding $\overline{D_{jt}}$ in core subjects increases the overall grade inflation effect, rather than decreasing it (Table 10). This result may be related to the reported decrease in mathematics and reading ability (Skolverket, 2012). Moreover, including $\overline{D_{jt}}$ in core subjects separately (without including the overall grade inflation measure) into the model shows that inflation in these core subjects has no impact on enrolment or earnings (results not shown). While this exercise is a weak test of exogeneity, it shows at least that our results are not affected by changes in school production of core subject skills.

Table 10 about here

9 Conclusions

Although grade inflation is common in most countries and at most educational levels, current knowledge of grade inflation effects on individual outcomes is scarce. The main explanation for the lack of research on this topic is probably the challenge of measuring grade inflation and estimating causal grade inflation effects. This study examines the association between grade inflation at upper secondary education level and higher education and earnings. The constructed grade inflation measure captures grade inflation at the school level and does not appear to include a component biasing our estimates. Confounding explanations for grade inflation are changes in the inflow of students and changes in school production. Our findings demonstrate that the grade inflation is not caused by school production, while controlling for past school achievements shows that other background factors also have no impact on the results. Moreover, we show that core

subjects (mathematics, English, and Swedish), where there are national tests for grade calibration, are not causing the grade inflation effect. Instead, the effect is caused by grade increases for subjects without national tests, where grading is, plausibly, more arbitrary. This finding makes us confident that we assess true grade inflation.

Grade inflation involves re-ranking of individuals and this aspect seems to have a substantial impact on the payoff of higher education, but only affects men's probability of investing in higher education. Instead, the main mechanism explaining the large impact of inflated grades on earnings is the impact on choice of university and field of education. In other words, even if the change in grades at enrolment is relatively small, since grades also affect the choice of education we observe a large impact on earnings. An unexpected negative effect of grade inflation on enrolment is shown for low-skilled women attending "leniently grading" schools. Because grade inflation is larger at the top of the grade distribution, low-skilled individuals mainly observe relative downgrading of their own achievements. We argue that this may cause a negative self-efficacy effect for women.

Because the re-ranking of individuals involves a benefit for those who gain higher grades at the expense of others with similar or even higher skills, it may have a negative impact on the productive potential of society. If nothing else, it is unfair.

References

- Bandiera, O., Larcinese, V. and Rasul, I. (2015). "Blissful ignorance? A natural experiment on the effect of feedback on students' performance", *Labour Economics* 34: 13–25.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

- Bandura, A. and Locke, E. (2003). “Negative self-efficacy and goal effects revisited”, *Journal of Applied Psychology* 88 (1): 87–99.
- Bar, T., Kadiyali, V. and Zussman, A. (2009). “Grade information and grade inflation: The Cornell Experiment”, *Journal of Economic Perspectives* 23(3), 93–108.
- Bauer, T. and Grave, B. (2011). “Performance-related funding of universities: Does more competition lead to grade inflation?”, IZA DP No. 6073.
- Bagues, M., Labini, M. and Zinovyeva, N. (2008). “Differential grading standards and university funding: Evidence from Italy”, *CESifo Economic Studies* 54(2): 149–176.
- Björklund, A., Fredriksson, P., Gustafsson, J-E. and Öckert B. (2010). “Den svenska utbildningspolitikens arbetsmarknadseffekter: vad säger forskningen?”, IFAU Rapport 2010:13.
- Borghans, L., Meijers, H. and ter Weel, B. (2008). “The role of noncognitive skills in explaining cognitive test scores” *Economic Inquiry* 46 (1): 2–12.
- Borghans, L., Golsteyn, B., Heckman, J. and Humphries, J. (2016). “What grades and achievement tests measure”, *Proceedings of the National Academy of Science (PNAS)*, 13354–13359.
- Butcher, K., McEwan, P. and Weerapana, A. (2014). “The effects of an antigrade-inflation policy at Wellesley College”, *Journal of Economic Perspectives* 28(3), 189–204.
- Carlstedt, B. and Gustafsson J. (2005). “Construct validation of the Swedish Scholastic Aptitude Test by means of the Swedish Enlistment Battery”. *Scandinavian Journal of Psychology* 46(1): 31–42
- Chan, W., Hao, L. and Suen, W. (2007). “A signaling theory of grade inflation”, *International Economic Review* 48(3): 1065–1090.

- Cliffordson, C. (2004). "Betygsinflation i de målrelaterade Gymnasiebetygen", *Pedagogisk Forskning i Sverige* 9: 1–14.
- Diamond, R. and Persson, P. (2017). "The long-term consequences of teacher discretion in grading of high-stakes tests", NBER Working Paper No. 22207.
- Duckworth, A.L. and Seligman M. (2005). "Self-discipline outdoes IQ in predicting academic performance of adolescents", *Psychological Science* 16 (12): 939–944.
- Gustafsson, J-E. and Yang Hansen, K. (2009). "Resultatförändringar i svensk grundskola". In: *Vad påverkar resultaten i svensk grundskola? Kunskapsöversikt om betydelser av olika faktorer*, Skolverkets rapport, 40–85.
- Henriksson, W. and Törnkvist, B. (2002). "The effects of repeated test taking in relation to the test taker and the rules for selection to higher education in Sweden", Report No. 41 (University of Umeå, Department of Educational Measurement).
- Hinnerich, B., Höglén, E. and Johannesson, M. (2011). "Are boys discriminated in Swedish high schools?". *Economics of Education Review* 30(4), 682–690.
- Holmlund, H., Häggblom, J., Lindahl, E., Martinson, S., Sjögren, A., Vikman U. and Öckert, B. (2014). "Decentralisering, skolval och fristående skolor: resultat och likvärdighet i svensk skola" IFAU Report No 2014:25.
- Hurwit, M. and Lee, J. (2018). "Grade inflation and the role of standardized testing". In Buckley, J., Letukas, L. and B. Wildavsky (ed.): *Measuring Success: Testing, Grades and the Future of College Admissions*. Baltimore: Johns Hopkins University Press.
- Hvidman, U. and Sievertsen, H. (2017). "Grades and Student Behavior," *Mimeo*.
- Johnson, V. (2003). "Grade Inflation: A Crisis in College Education". Springer.
- Kirkeboen, L., Leuven, E. and Mogstad, M. (2016). "Field of study, earnings, and self-selection". *Quarterly Journal of Economics* 131(3): 1057–1111.

- Lavy, V. (2008). “Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment”, *Journal of Public Economics* 92, 2083–2105.
- Lindahl, L. and Regnér, H. (2005) “College choice and subsequent earnings: Results using Swedish sibling data”, *Scandinavian Journal of Economics* 107: 437–457.
- Manacorda, M. (2012). “The cost of grade retention”, *The Review of Economics and Statistics*, 94(2): 596–606.
- Maurin and McNally (2008). “Vive la Révolution! Long-term educational returns of 1968 to the angry students”, *Journal of Labor Economics* 26:1–33.
- OECD (2015). “The ABC of Gender Equality in Education Aptitude, Behaviour, Confidence”, PISA, OECD Publishing.
- Redding, R. (1998). “Students' evaluations of teaching fuel grade inflation”, *American Psychologist* 53(11), 1227-1228.
- Rey H-J. and Looney, A. (2016). “Measuring inflation in grades: An application of price indexing to undergraduate grades”, *Economics of Education Review* 55: 220–232.
- Sabot, R. and Wakeman-Linn, J. (1991). “Grade inflation and course choice”, *Journal of Economic Perspectives* 5(1): 159–170.
- Skolverket (2012) Betygsinflation – betygen och den faktiska kunskapsutvecklingen. Dnr 2012:387.
- Stage, C. and Ögren, G. (2009). Högskoleprovet våren och hösten 2009 Provdeltagargruppens sammansättning och resultat, BMV Rapport 40, Institutionen för beteendevetenskapliga mätningar, Umeå universitet.
- Tholin, J. (2003). “A Funny Dance”. Doctoral thesis, Borås, The University College of Borås, Department of Education.

- Vlachos, J. (2010). Betygets värde. En analys av hur konkurrens påverkar betygssättningen vid svenska skolor, Konkurrensverket rapport 2010:6.
- Wikström, C. and Wikström, M. (2005). "Grade inflation and school competition. An empirical analysis based on the Swedish upper secondary schools". *Economics of Education Review* 24, 309–322.
- Wikström, C. (2005). "Grade stability in a criterion-referenced grading system: the Swedish example", *Assessment in Education: Principles, Policy & Practice* 12(2): 125–144.
- Wuepper, D. and Lybbert, T. (forthcoming). "Perceived self-efficacy, poverty and economic development," *Annual Review of Resource Economics*.

Appendix

Parental education in 1990 is measured as years of schooling. For 12 percent who have missing education level, a dummy variable is included. Parental income is an average of (positive) incomes in 1985 and 1990. The first-generation immigrant dummy is child's region of origin: Nordic countries, EU28, Non-EU28 countries and Russia, North America and Oceania, Africa, Asia, and South America. Second-generation immigrants are born in Sweden, but their parents are both born abroad. We use the same region of origin as for first-generation immigrants, but we include also a dummy for having a parent with mixed origin.

Tables and Figures

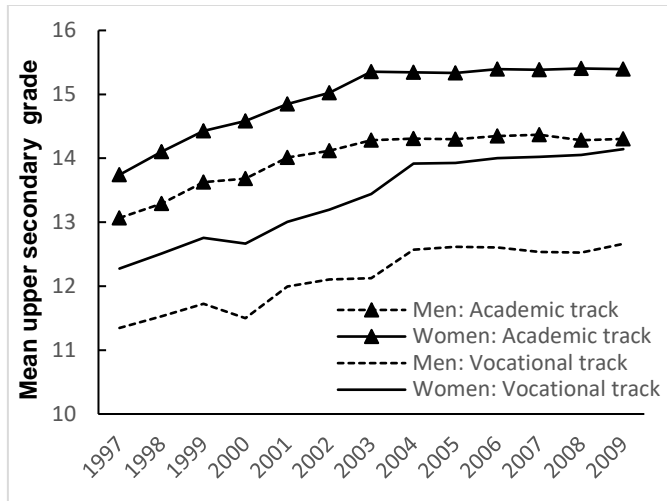


Figure 1. Mean upper secondary grades obtained by men and women in Sweden, 1997-2009.

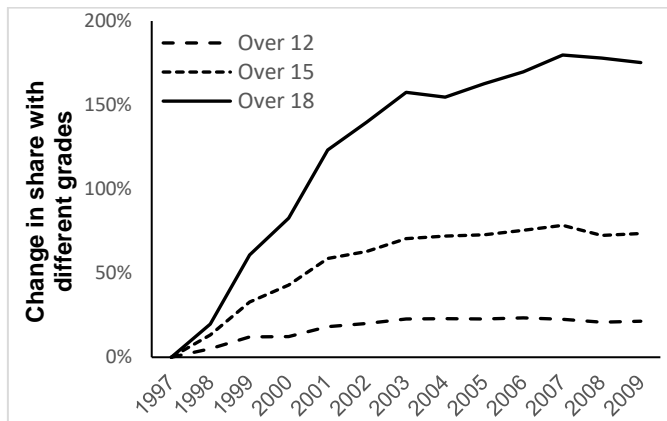


Figure 2. Change in the share of Swedish students with a grade above 12, 15, and 18 (where 20 is the top grade), 1997-2009.

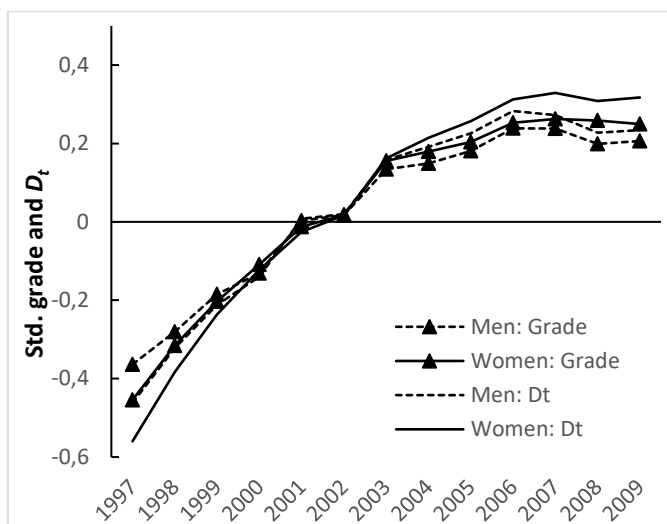
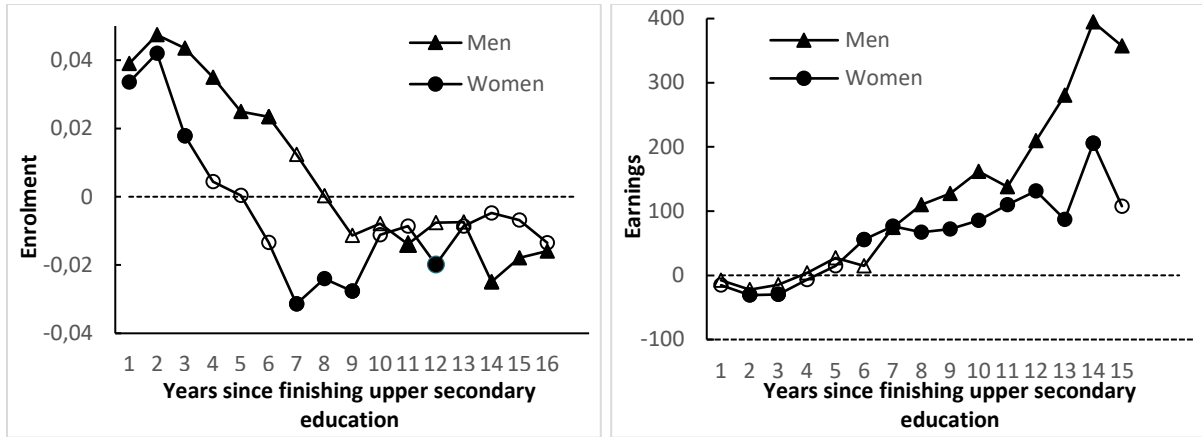


Figure 3. Mean standardized upper secondary grade and mean standardized D_{jt} in the test-taking population of Swedish students, 1997-2009.



Figures 4 and 5. Longitudinal relationship between the grade inflation impact on enrolment and earnings (filled markers show significant effects).

Table 2. Estimating the relationship between grade inflation (measured as \overline{D}_{jt}) and different covariates

	Grade inflation among men				Grade inflation among women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of women		0.0200 (0.0951)	0.0088 (0.0956)	0.0235 (0.0978)		-0.0118 (0.0800)	-0.0128 (0.0800)	0.0022 (0.0832)
Mean compulsory grade		0.514*** (0.0399)	0.523*** (0.0405)	0.520*** (0.0409)		0.405*** (0.0378)	0.407*** (0.0387)	0.407*** (0.0399)
Father's mean income			0.0106 (0.0694)	0.0364 (0.0745)		0.0779 (0.0637)	0.0532 (0.0599)	0.0748 (0.0623)
Mother's mean income			-0.0350 (0.0600)	-0.0222 (0.0608)		-0.0237 (0.0590)	-0.0254 (0.0556)	-0.0331 (0.0570)
Father's mean education			3.45e-05 (0.0219)	-0.0035 (0.0225)		0.0043 (0.0195)	-0.0145 (0.0189)	-0.0158 (0.0189)
Mother's mean education			0.0082 (0.0274)	0.0066 (0.0279)		0.0503* (0.0259)	0.0071 (0.0237)	0.0089 (0.0242)
Migrant share			0.283 (0.223)	0.344 (0.225)		0.0179 (0.214)	0.102 (0.214)	0.190 (0.221)
Humanities track		0.166 (0.225)	0.188 (0.229)	0.180 (0.229)		0.210 (0.170)	0.219 (0.167)	0.186 (0.169)
Science track		-0.246** (0.0968)	-0.236** (0.0953)	-0.260*** (0.0990)		-0.320*** (0.0903)	-0.319*** (0.0908)	-0.372*** (0.0900)
Social science track		0.217** (0.0966)	0.228** (0.0950)	0.231** (0.0979)		-0.0073 (0.0834)	-0.0078 (0.0833)	-0.0080 (0.0868)
Technology track		0.0415 (0.182)	0.0655 (0.181)	-0.0310 (0.175)		-0.0942 (0.175)	-0.0872 (0.175)	-0.133 (0.172)
Share with teaching qualification				0.0013 (0.0016)				0.0012 (0.0015)
Share perm. employed teachers				-8.05e-05 (0.0022)				-0.0013 (0.0021)
Teacher-student ratio				0.0206 (0.0131)				0.0083 (0.0107)
School- and time dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	128,893	128,893	128,893	123,914	151,471	151,471	151,471	145,402
Within R-squared	0.472	0.539	0.541	0.541	0.583	0.625	0.626	0.626

Notes: The dependent variable is \overline{D}_{jt} , school-level divergence between upper secondary grades and SweSAT at the gender level. When controlling for parental education, we also control for the share of missing values in the educational variable. ***Significant at 1% level. **Significant at 5% level. *Significant at 10% level. Clustered (on schools) standard error in brackets.

Table 3. Estimating the effect of grade inflation (measured as \overline{D}_{jt}) on enrolment in higher education and earnings.

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<u>Enrolment in higher education</u>								
\overline{D}_{jt}	0.0678*** (0.009)	0.029*** (0.008)	0.028*** (0.008)	0.0307*** (0.00772)	0.0300*** (0.00833)		0.024*** (0.008)	-0.010 (0.007)
School-level grade variation						0.083*** (0.007)		
Mean compulsory grade		0.166*** (0.012)	0.159*** (0.015)	0.158*** (0.0126)	0.152*** (0.0146)	0.054*** (0.015)		0.143*** (0.011)
Gender and migrant share	no	no	yes	yes	yes	yes	no	no
Parental background	no	no	yes	yes	yes	yes	no	no
Tracks	no	yes	yes	yes	yes	yes	no	yes
School quality indicators	no	no	no	yes	no	no	no	no
School linear time trends	no	no	no	no	yes	no	no	no
R-squared	0.005	0.007	0.007	0.007	0.010	0.008	0.002	0.004
<u>Earnings</u>								
\overline{D}_{jt}	236.0*** (33.00)	175.5*** (33.96)	186.3*** (33.84)	185.4*** (34.72)	166.7*** (35.74)		182.8*** (28.17)	110.7*** (27.66)
School-level grade variation						183.4*** (29.99)		
Mean compulsory grade		265.3*** (53.76)	211.5*** (54.47)	219.5*** (55.04)	290.6*** (60.87)	50.40 (68.50)		319.3*** (33.68)
Gender and migrant share	no	no	yes	yes	yes	yes	no	no
Parental background	no	no	yes	yes	yes	yes	no	no
Tracks	no	yes	yes	yes	yes	yes	no	yes
School quality indicators	no	no	no	yes	no	no	no	no
School linear time trends	no	no	no	no	yes	no	no	no
R-squared	0.025	0.025	0.025	0.025	0.028	0.025	0.009	0.008
Observations		128,893		123,914		128,893		151,471

Notes: The dependent variable in the upper and lower panel is enrolment in higher education and earnings. \overline{D}_{jt} is the school-level divergence between upper secondary grades and SweSAT at the gender level. All models include school and time fixed effects. When controlling for parental education, we also control for the share of missing values in the educational attainment variable. Error in brackets. ***Significant at 1% level. **Significant at 5% level. *Significant at 10% level.

Table 4. Estimating the effect of grade inflation (measured as \overline{D}_{jt}) on different subsamples of students

	Sample enrolling in higher education						Sample without higher education	
	Years of schooling		Prob. Exam		Earnings		Earnings	
	Men	Women	Men	Women	Men	Women	Men	Women
\overline{D}_{jt}	0.139*** (0.0241)	0.0932*** (0.0232)	0.0303*** (0.0082)	0.0188** (0.0076)	202.6*** (40.62)	153.7*** (30.38)	90.21* (53.60)	-64.67 (52.79)
Observations	97,444	124,665	97,444	124,665	97,444	124,665	31,449	26,806
R-squared	0.012	0.010	0.008	0.005	0.027	0.008	0.014	0.011

Notes: The dependent variables are years of schooling, the probability of reaching an exam in higher education and earnings. \overline{D}_{jt} is the school-level divergence between upper secondary grades and SweSAT at the gender level. All models include school dummies and dummies for year when finishing upper secondary education. Controls include family background, the gender and migration share, track, and compulsory grades (all measured at the school level). Clustered (on school) standard error in brackets. ***Significant at 1% level. **Significant at 5% level. *Significant at 10% level.

Table 5. Estimating the effect of grade inflation (measured as \overline{D}_{jt}) on the probability of attending an old university and mean grades at the university and in the chosen field of education

	Old university		High-quality university		High-grade field	
	Men	Women	Men	Women	Men	Women
\overline{D}_{jt}	0.0206** (0.0085)	0.0370*** (0.0085)	0.0824*** (0.0208)	0.104*** (0.0210)	0.0601*** (0.0212)	0.0762*** (0.0180)
Observations	97,444	124,665	97,444	124,665	97,444	124,665
R-squared	0.002	0.001	0.003	0.002	0.001	0.001

Notes: The dependent variables are the incidence of studying at an old university, mean grade at the university, and the mean grade in the field of education. \overline{D}_{jt} is the school-level divergence between upper secondary grades and SweSAT at the gender level. The sample is conditional on enrolment in higher education. All models include school dummies and dummies for year when finishing upper secondary education. Controls include family background, the gender and migration share, track, and compulsory grades (all measured at the school level). Clustered (on school) standard error in brackets. ***Significant at 1% level. **Significant at 5% level. *Significant at 10% level.

Table 6. Estimating the relationship between grade inflation (measured as \overline{D}_{jt}) and chosen field of education

	Teaching	Humanities	Social science	Economics and law	Science	Computer science	Engineering	Agriculture	Health	Medicine	Services
Men											
\overline{D}_{jt}	-0.001 (0.004)	-0.005 (0.003)	0.003 (0.004)	0.017*** (0.006)	-0.001 (0.003)	-0.007** (0.003)	0.018** (0.008)	-0.000 (0.001)	-0.002 (0.003)	0.008*** (0.002)	-0.000 (0.003)
Women											
\overline{D}_{jt}	-0.009* (0.005)	-0.000 (0.004)	-0.008 (0.005)	0.0112 (0.007)	-0.006** (0.003)	-0.005*** (0.002)	0.006 (0.005)	0.001 (0.002)	-0.001 (0.006)	0.005* (0.003)	-0.001 (0.003)

Notes: The dependent variables are the incidence of studying in certain fields of education in higher education. \overline{D}_{jt} is the school-level divergence between upper secondary grades and SweSAT at the gender level. All models include school dummies and dummies for year when finishing upper secondary education. Controls include family background, the gender and migration share, track, and compulsory grades (all measured at the school level). Clustered (on school) standard error in brackets. ***Significant at 1% level. **Significant at 5% level. *Significant at 10% level.

Table 7. Estimating the effect of grade inflation (measured as \overline{D}_{jt}) on enrolment and earnings for individuals below and above mean compulsory school grades

	Men		Women	
	Sample with compulsory grade below mean	Sample with compulsory grade above mean	Sample with compulsory grade below mean	Sample with compulsory grade above mean
<u>Enrolment in higher education</u>				
\overline{D}_{jt}	0.0056 (0.0125)	0.0533*** (0.0085)	-0.0396*** (0.0110)	0.0166** (0.0077)
Observations	63,324	65,569	73,693	77,778
R-squared	0.004	0.003	0.003	0.001
<u>Earnings</u>				
\overline{D}_{jt}	128.2*** (40.02)	247.7*** (52.42)	10.04 (34.01)	208.1*** (41.87)
Observations	63,324	65,569	73,693	77,778
R-squared	0.022	0.024	0.008	0.008

Notes: The dependent variables are enrolment in higher education and earnings (in 2013). \overline{D}_{jt} is the school-level divergence between upper secondary grades and SweSAT at the gender level. All models include school dummies and dummies for year when finishing upper secondary education. Controls include family background, the gender and migration share, track, and compulsory grades (all measured at the school level). Clustered (on school) standard error in brackets. ***Significant at 1% level. **Significant at 5% level. *Significant at 10% level.

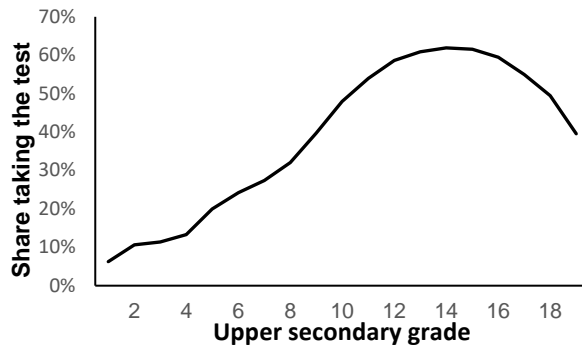


Figure 6. Relationship between the share of students taking the SweSAT test and the average upper secondary grade.

Table 8. Testing whether selection in SweSAT test taking alters the effect of grade inflation (measured as \overline{D}_{jt}) on enrolment and earnings

	Higher education		Earnings	
	Men	Women	Men	Women
\overline{D}_{jt}	0.0278*** (0.0071)	-0.0041 (0.0065)	182.4*** (34.03)	116.6*** (27.68)
Share taking the SweSAT	0.224*** (0.0270)	0.139*** (0.0371)	9.422 (121.0)	175.6 (107.8)
Share taking the SweSAT \times Mean compulsory school grade	-0.0323 (0.0480)	0.103 (0.0690)	182.8 (272.4)	34.58 (200.5)
Share taking the SweSAT \times Mean compulsory school grade ²	-0.0328 (0.0272)	-0.0802*** (0.0277)	136.3 (161.4)	-26.68 (96.35)
Observations	128,893	151,471	128,893	151,471

Notes: The dependent variables are enrolment in higher education and earnings (in 2013). \overline{D}_{jt} is the school-level divergence between upper secondary grades and SweSAT at the gender level. All models include school dummies and dummies for year when finishing upper secondary education. Controls include family background, the gender and migration share, track, and compulsory grades (all measured at the school level). Clustered (on school) standard error in brackets. ***Significant at 1% level. **Significant at 5% level. *Significant at 10% level.

Table 9. Investigating whether the individual progression of grades affects the effect of grade inflation (measured as \overline{D}_{jt})

	Enrolment in higher education		Earnings	
	Men	Women	Men	Women
\overline{D}_{jt}	0.0221*** (0.0076)	-0.0114* (0.0066)	168.0*** (33.82)	114.8*** (27.73)
Progression of grades	0.0153*** (0.0021)	0.0050** (0.0017)	48.02*** (5.608)	-16.11*** (4.224)
Observations	128,507	151,127	128,507	151,127
R-squared	0.010	0.004	0.026	0.009

Notes: The dependent variables are enrolment in higher education and earnings (in 2013). \overline{D}_{jt} is the school-level divergence between upper secondary grades and SweSAT at the gender level. All models include school dummies and dummies for year when finishing upper secondary education. Clustered (on school) standard error in brackets. ***Significant at 1% level. **Significant at 5% level. *Significant at 10% level.

Table 10. Investigating whether the overall effect of grade inflation (measured as \overline{D}_{jt}) can be explained by grade inflation in core subjects

	Enrolment in higher education		Earnings	
	Men	Women	Men	Women
\overline{D}_{jt}	0.0413*** (0.0095)	0.0079 (0.010)	224.6*** (44.78)	130.6*** (41.34)
\overline{D}_{jt} in mathematics	-0.0251*** (0.0079)	-0.012 (0.007)	16.78 (35.31)	-19.20 (27.84)
\overline{D}_{jt} in English	-0.0075 (0.0068)	-0.005 (0.0067)	-16.74 (38.43)	21.15 (26.21)
\overline{D}_{jt} in Swedish	0.0077 (0.0080)	-0.011 (0.007)	-73.13* (42.41)	-24.68 (28.90)
Observations	128,876	151,471	128,876	151,471
R-squared	0.007	0.004	0.025	0.009

Notes: The dependent variables are enrolment in higher education and earnings (in 2013). \overline{D}_{jt} is the school-level divergence between upper secondary grades and SweSAT at the gender level. All models include school dummies and dummies for year when finishing upper secondary education. Clustered (on school) standard error in brackets. ***Significant at 1% level. **Significant at 5% level. *Significant at 10% level.

Appendix

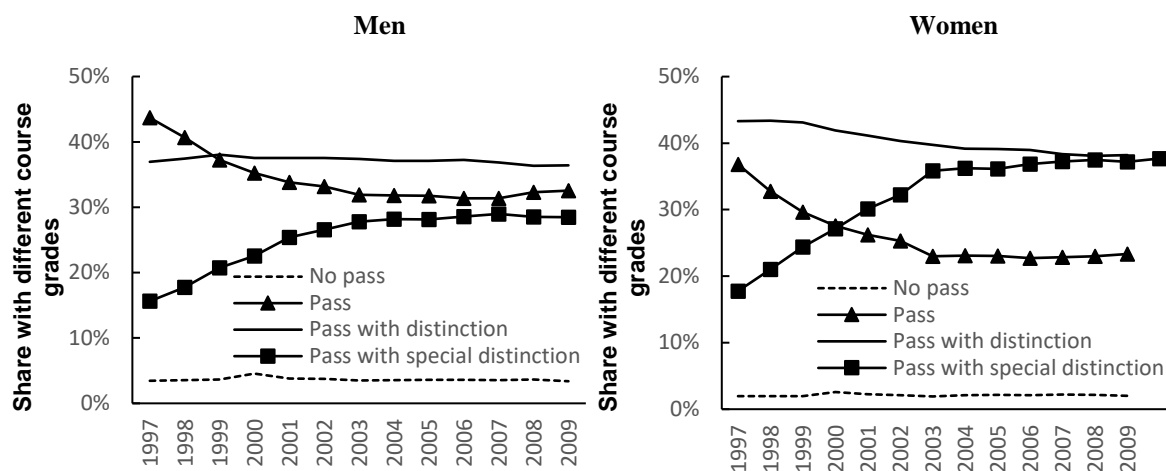


Figure A1. Share of Swedish students obtaining “no pass”, “pass”, “pass with distinction”, and “pass with special distinction” upper secondary grades, 1997-2009.