NHH

INSTITUTT FOR FORETAKSØKONOMI

DEPARTMENT OF FINANCE AND MANAGEMENT SCIENCE

FOR 5 2010

ISSN: 1500-4066

JUNE 2010

Discussion paper

Grading standards, student ability and errors in college admission

BY

JARLE MØEN AND **MARTIN TJELTA**

Norges Handelshøyskole

NORWEGIAN SCHOOL OF ECONOMICS AND BUSINESS ADMINISTRATION

Grading standards, student ability and errors in college admission*

by Jarle Møen‡ and Martin Tjelta‡‡

ABSTRACT

Grades are important for admission of students in most higher education programmes. Analysing admission and student performance data at a major Norwegian business school, we find that the grading practice of teachers at regional colleges sending students to the school is affected by the average performance of the students being graded. Teachers at colleges recruiting good students from upper secondary school tend to be strict in their grading practice, while teachers at colleges recruiting less good students tend to follow a lenient practice. This has implications for the interpretation of grades and hence for optimal admission procedures. We develop a methodology to assess the consequences of differential grading standards. Approximately ten percent of the students in our data are admitted at the expense of more competent students. We demonstrate costs for the school admitting wrong students and in particular for the rejected students.

KEYWORDS: Grading practices, Differential grading standards, Admission policy

JEL CODES: A20, I21

This is a summary of an article forthcoming in the Scandinavian Journal of Educational Research, Vol. 54(3), 2010, 221–237. The publisher, Taylor & Francis, does not allow systematic distribution of electronic preprints or reprints. Please contact the corresponding author for a complete version. The Scandinavian Journal of Educational Research is available online at http://www.informaworld.com/smpp/title~db=all~content=t713445993.

* This study was originally commissioned by the Norwegian School of Economics and Business Administration as part of the 2003 education reform where the "Siviløkonom" course, was changed from an integrated four-year study programme to a five-year bachelor + master programme. We are grateful to Lars Mathiesen for support and Vivienne Knowles for excellent language assistance. The authors remain solely responsible for the conclusions and all opinions expressed.

‡ Corresponding author. Professor, Department of Finance and Management Science, Norwegian School of Economics and Business Administration, Helleveien 30, N-5045 Bergen, Norway. Ph.: + 47 55 95 96 12 Fax: + 47 55 95 96 50. Email: jarle.moen@nhh.no. Homepage: www.nhh.no/for/cv/moen.

‡‡ Adviser, Finance and Management Department, University of Stavanger, N-4036 Stavanger, Norway. Ph.: + 47 51 83 30 62 Fax: + 47 51 83 10 50. Email: martin.tjelta@uis.no.

Introduction

Grades are important for comparing students, and in particular when ranking students for admission to higher education. We investigate whether the grading practice of teachers is affected by the average quality of the students being graded. Such an effect will have implications for the interpretation of grades awarded by different colleges and for admission to advanced study programmes.

In Norway, as in many other countries, applications to undergraduate studies are decided by a centralised application processing centre.[1] Admission is based on the students' grade point average (GPA) adjusted for the number and composition of credits, work experience and age.[2] No discretionary judgement of the students' academic potential is done. This system is administratively highly efficient, and fair in the sense that it is strictly governed by rules that apply in the same manner for all students regardless of their background and connections. In order for this system to allocate applicants correctly to study programmes, however, it is a prerequisite that all institutions practice the same grading standard. It has been asserted that this is not the case, and our study aims to illuminate this issue.[3]

If the current admission system is fair and efficient, students with equal GPA or admission points should be equally competent to succeed in the field of study they compete for admission to, regardless of which institution they apply from. Within the field of economics and administration, there are a large number of undergraduate programmes with a similar curriculum. The "Siviløkonom" course at the Norwegian School of Economics and Business Administration (NHH) is a very popular place for students to continue from all of these programmes. This makes NHH a suitable "laboratory" to investigate the existence of differing grading practices and their causes and consequences. Our research strategy is to analyse how student performance at NHH varies with previous grades and the characteristics of the colleges where the grades were obtained.

---

[1] The Norwegian Universities and Colleges Admission Service, see http://www.samordnaopptak.no/info/english/. Although centralised application processing centres are found in many countries, there is large heterogeneity in the extent of institutional autonomy concerning admission procedures. In Norway, only a few private colleges govern their own admission at the bachelor level.

[2] The GPA adjusted for credits achieved in upper secondary school is sometimes referred to as "school points",

[3] Pål Veiden opened this debate with a feature article in the Norwegian magazine Morgenbladet on July 9th 2004. He warned of "local grade cultures" and predicted a development where "college and university grades would become better the further from Oslo one finds oneself and the poorer the performance of the college or university". Similar concerns have been raised e.g. in Sweden by Collin and Smith (2007) and in Italy by Bagues, Labini and Zinovyeva (2007) and De Paola (2009).

The NHH Siviløkonom course analysed was a four year program divided into 3 study blocks.[4] The first and second blocks lasted 3 semesters each, while the third block lasted 2 semesters. From 1994 until 2002, students from regional colleges competed for direct admission to the second block of the Siviløkonom course. The qualifying college programmes lasted two years and were subject to a common framework with a high degree of overlap with mandatory subjects in the first block of the Siviløkonom course. The colleges further operated with a common grading scale (1-6) and admission to NHH was based on a point system that provided for equal treatment of all colleges.

If similar performances are rewarded similarly across all colleges, a given sum of admission points should represent the same potential for success at NHH. In a regression model used to predict student achievement based on admission points, the college an applicant comes from should reveal no information concerning expected success. Finding that for a given sum of admission points, students from certain colleges perform systematically better or poorer than others would indicate that different colleges grade equally good students differently. This is investigated in the first part of our analysis. We find that college background reveals information concerning academic success at NHH, even if we control for admission points. More specifically, we find that the grading practice of teachers is affected by the average quality of the students being graded. Teachers at regional colleges recruiting good students from upper secondary school tend to be strict in their grading practice, while teachers at colleges recruiting less good students tend to follow a lenient practice. We also find some support for the view that study programmes with many students are stricter than small programmes.

In the second part of the analysis, we quantify the consequences of different colleges using the grading scale differently. We find that approximately ten percent of students admitted directly to the second block of studies at NHH were admitted at the expense of more competent students from strict colleges. The difference between the most and the least strict

[4] The Siviløkonom course described here was phased out in 2006 and replaced by a five year bachelor + master program in compliance with the EU Bologna process. Students who complete a bachelor degree in business administration at a regional college may apply directly to the master program at NHH, and admission to the master program is based on their bachelor grade point average. Hence, the problem we analyze is as relevant after the reform as before. This grade comparison problem also applies outside the field of business administration. There are e.g. master programs in engineering where bachelor students from numerous colleges compete for admission, and there is large variation in entrance requirements at these colleges in terms of admission points from upper secondary school.

colleges is so great that it can take students several semesters to acquire corresponding additional admission points, and those students, which our model classifies as "error admissions", have experienced a higher failure rate and general weaker performance than a control group of marginally qualified, "correct admission" students.

What could cause heterogeneity in grading standards?

There is a substantial literature on grading standards summarized e.g. in Johnson (2003) and Hu (2005). One large strand of the literature deals with grade inflation. This literature focuses on how incentives to award good grades lead to more lenient grading standards in higher education – particularly in the US. Commonly investigated incentives relates to student evaluations, competition for students and allocation of public money. An intertwined strand of the literature deals with grade divergence, i.e. how grading standards evolve differently across different institutions, departments and fields. Grade divergence is part of a larger literature on differential grading standards. One mechanism for differential grading standards not related to incentives was investigated by Goldman, Schmidt, Hewitt and Fischer (1974). Building on Helson's (1947; 1948) theory of adaption level, they found evidence suggesting that teachers anchor their judgement of a specific student's performance by comparing that student's performance to the performance of the other students in the class. This causes the stringency of grading standards to be positively related to average student ability. They point out that this has implications for college admission.

In Norway, teachers at all levels are supposed to grade students according to a "national norm". The role of a national assessment norm is to ensure compatibility of grading standards across institutions, both at the upper secondary school level and in higher education. However, the national norm is not well defined, and it is challenging for teachers to have a strong opinion as to "where the bar is" for a given grade in a given subject at a given level. In this situation it seems plausible that the answers in a batch of examination papers – to some extent – will be assessed comparatively, rather than relative to the national norm. Hægeland, Raaum and Salvanes (2005) find that such an effect manifests itself in Norwegian lower secondary school where they compare final assessment and grades on national exams for students completing 10th grade in 2002 and 2003. Such comparative assessment creates problems if the quality of the students varies systematically from college to college. If an average student attends a college with low admission requirements and therefore has many

weak fellow students, it is easier for this student to distinguish him or herself positively than for a similar student who is the last to be admitted at a college with high admission requirements. Should this line of thought prove correct, the different use of the grading scale may be related to how easy or difficult it is to be admitted to different colleges.

If examination papers in a batch are assessed relative to each other, this may also cause students at colleges with weak professors to be assessed too leniently. Poor course quality will not be fully reflected in poor grades. Good professors normally prefer central institutions with a large faculty, and it is therefore possible that differing grading practices will be indirectly related to the size of the study programmes. A mechanism that points in the same direction is that small classes promote strong bonds between students and teachers. This may create a psychological barrier against strict grading.

Data

The data used encompasses admissions information from 1998 to 2002 and examination information from 1998 to 2003, derived from the student database system "FS" at NHH.

The grade data includes all subjects within the Siviløkonom course for the second and third study blocks from spring 1998 to spring 2003. This encompasses more than 38 000 examination papers handed in by 2313 individuals. Hence we have an unbalanced panel with 16.5 grade observations per individual students. The data base includes 529 "college students" admitted to the second study block of the course with point of departure in a two-year college programme, and 1748 "four-year students", i.e. students admitted directly from upper secondary school to the first study block of the course. See table 1 for further information. The most central part of our analysis is based on the sample of "college students".

[TABLE 1 ABOUT HERE]

Admission data includes both admitted and rejected applicants for the second block of study to the Siviløkonom course for the period 1998-2002. We have complete information on individual admission points, college background and year of application. Admission points represent the students' averages grades from college adjusted for any additional points. We cannot distinguish between "grade points" and "additional points". However, it takes time to

earn additional points, and we know the age of the applicants. This is used to proxy for additional points. Furthermore, we have information on the number of student places for the economic and administrative programmes of the different colleges, as well as the admission point requirements. Descriptive statistics on the different regional colleges included in this study are provided in table 2. We have chosen not to disclose the identity of the colleges used in the analysis, hence they are labelled college A-W. To assure that publicly available information given in table 2 cannot be used to identify whether a particular college follows a strict or lenient grading practice, we have used another set of labels in tables with regression results. Colleges are in these tables labelled "college 1-23" and the order is different from the order in table 2.

[TABLE 2 ABOUT HERE]

The relationship between college background and academic success at NHH

The relationship between college background and academic success at NHH is investigated with the aid of regression analysis.[5] We estimate four different models, reported in table 3. The dependent variable is examination grade at NHH. The grading scale is from 0 to 9, with 9 as the highest grade, 5 as the laudability limit and 2 as the lowest pass mark.[6] Since the grade database is a panel with repeated observations of each student, we allow an unobserved individual-specific effect in the regressions. This individual specific effect will capture each student's ability or potential to do well at NHH beyond what can be inferred through college background and admission points. This is modelled as a normally distributed random effect. The estimation method is Generalised Least Squares (GLS)[7]. Using the Lagrangian Multiplier

[5] See Betts (1997) and Betts and Grogger (2003) for similar analyses of American junior and senior high schools. They find significant differences in grading practice between different schools, and demonstrate that strict grade-setting increases student learning. Figlio and Lucas (2004) find corresponding results in an analysis of variations in grading practices between American elementary school teachers. In a recent paper using data from Italian Universities, De Paola (2009) find that "students obtain lower grades in courses attended by students with higher ability". In line with our hypothesis, she suggests that this is because "instructors evaluate their students on a relative basis and adjust their grading standards to the average level of ability encountered in the classroom".

[6] A new grading scale with grades from A-F was implemented in the autumn of 2003. Our data set extends to, and includes, the spring of 2003.

[7] Generalised Least Squares is a flexible generalization of ordinary least squares. If there is an unobserved individual specific effect in the grade data, the grades a student obtains will be correlated even when we condition on the observed student characteristics. A new observation of a student already in the data set,

test for random effects developed by Breusch and Pagan (1980) and Baltagi and Li (1990), we can clearly reject the null hypothesis of no individual specific effect in the error terms. This, of course, does not assure that our normality assumption is correct, but the estimated coefficients are robust to using ordinary least squares (OLS) in steady of GLS. The OLS estimates do not rely on normally distributed error terms.[8]

We start by investigating whether the college students on average perform poorer or better than the four-year students do. From table 3, first column, it may be seen that there is a statistically significant, but very small (-0.13) difference in average grades between college students and four-year students. This small difference indicates that the admission limit for admission of college students to the second block course of study is practically correctly set. One would like the last to be admitted college student to be as able as the last to be admitted four-year student. We do not have sufficient information to check marginal students against each other and must therefore rely on average differences between the two groups.

In column 2, we investigate whether this very small average difference conceals essential differences between students from different colleges. The dummy variable for college students is therefore, replaced with separate dummy variables for each of the 23 colleges.

[TABLE 3 ABOUT HERE]

We see that in part, there are significant differences in achievement between students from different colleges. This may indicate different grading practices in the different colleges, but note from table 2 that several of the colleges are represented by very few students. Moreover, since we have not controlled for the students' admission points, the coincidental composition of the student groups from different colleges may affect the results.

If all colleges use the grading scale alike, a given sum of admission points should represent the same potential for success at NHH. Testing this is a main objective in our analysis. In column 3, therefore, we include the admission points of each student on admission to the

---

therefore, does not contain as much information as a new observation of a student not already in the data set. Assuming random effects, GLS will minimize a *weighted* sum of squared residuals so that this information about the group structure of the data set is used optimally. The weakness of the approach is that the results may be sensitive to the particular assumptions made about the structure of the residuals.

[8] Strictly speaking, the error terms cannot be normally distributed, as our dependent variable is categorical. Grade data, however, are usually well approximated with the normal distribution.

second block course of study at NHH. If all colleges use the grading scale alike, the college an applicant comes from should now not reveal any information concerning expected success.

Since the four-year students are admitted to NHH directly from upper secondary school and do not have comparable admission points to the second block course, these are excluded from the rest of our analysis. We choose as a new reference group the students with background from college no. 1. This is of course an arbitrary choice, but it seems a natural choice as the students from colleges no. 1 are similar to the original NHH reference group, cf. the coefficient on college no. 1 in column 2. Moreover, the group is fairly large so the measured level is not likely to be driven by random variation.

Admission points to the second block course of study at NHH consist of average grade from the two year college programmes and up to 0.4 additional points for work experience, military service and/or other higher education qualifications over and above admission requirements. In column 3, we also control for age. The age variable runs from 0 to 7, where $0 \leq 21$ years, 1=22 years, ... , 6=27 years and $7 \geq 28$ years. The variable measures the number of years the students have had to acquire additional points.

From column 3 we see that when the average grade (admission points) from a previous place of study increases by 1, the anticipated grade at NHH increases by 1.86 ±0.14, all else equal.[9] Note that each grade level on a scale of 1 to 6 corresponds to a 1.67 grade stage on the NHH scale of 0 to 9, and if we only use pass grades, each grade level from 1 to 4, corresponds to two grade stages on a scale from 2-9. Hence, there seems to be a close, one to one, correspondence between grade levels from the previous places of study and the grade levels at NHH. We test the specification for possible non-linearities in admission points and age but find no significant second order coefficient.

As mentioned, the age variable is intended as a proxy for students' additional points on admission. If the additional point awarding activities represent relevant competence building in relation to studies at NHH, the age variable should not have significant impact. The grade average would then be correctly adjusted for greater competence achieved via point awarding activities. The coefficient is however, significantly negative and indicates that the additional

[9] The sign on the coefficient in the table is negative because admission points follow the (old) college grading scale where 1 is the best and 6 the weakest. As a result, good students have a low number of admission points.

points system has been too generous from an academic point of view. All else equal, an increase of 1 year in age at commencement of studies, causes the expected grade at NHH to fall by 0.15 ±0.01. Hansen (2005) makes a similar finding in an analysis of Norwegian medical students.

The most interesting variables in column 3 are the dummy variables identifying the college from which the individual applicant had his or her background. Having controlled for admission points and age, we should see no significant coefficients on these variables if all colleges assess similar performances equally. However, for given admission points and age, there are significant differences in student performance between students with background from different colleges. An F-test clearly rejects a hypothesis that all the college dummies are simultaneously zero. The difference between the best and the poorest student groups is almost a whole grade. Students from college no.18 have on average, a 0.58 poorer grade than students from college no. 1 (the reference group), while students from college no. 13 have, on average, a 0.32 better grade. This suggests that colleges use the grade scale differently.

The regression reported in column 3 has the highest level of explanation in the table, but R^2 is nevertheless, only 19 %. This indicates a wide grade variation around the predicted level, and is to be expected. The model does not account for random variation in student preparation, student performance, difficulty of the exam and random errors in the censoring process. In addition, the students' admission points do not perfectly measure their academic potential, cf. our discussion of the error term on page 6.

We have hypothesised that colleges with a lenient grading practice will be characterised by low admission requirements and few student places. This is tested in column 4 where the dummy variables for which college a student applies from are replaced by the colleges' characteristics. We have included admission requirements for non-open colleges, a dummy for whether a student is from an open course of study, i.e. a programme where all students with a diploma from upper secondary school are accepted, and the number of student places. All characteristics are measured two years prior to admission to NHH such that they agree with the intake of the individual student in our sample.[10]

[10] Gain of additional points and delay in earlier studies create a degree of imprecision in the timing, but this is hardly a problem given that the college characteristics are relatively stable over time.

We see that, all else equal, the expected grade at NHH increases by 0.29 ± 0.09 when the admission requirement to the college the student attended increases by 1 grade. Admission requirement is measured by school points which are the average grade from upper secondary school. The number of student places is also influential. The expected grade at NHH increases by 0.11 ± 0.03 per hundred students of the programme from which they come, all else equal. One possible interpretation is that small schools with close contact between teachers and students reward students' performances more generously than large schools with greater distance between teachers and students.

It is perhaps counter-intuitive that the dummy variable for open studies is positive. Apparently, the expected grade at NHH is 1.22 *higher* for students with background from open studies. However, for these students, the school points variable with coefficient 0.29 is missing and set to zero. Hence, these two variables need to be evaluated together. The interpretation of the two coefficients is that open studies impose a student quality corresponding to an admission limit of 1.220/0.0293 = 41.64 school points. This is not extremely low, so open studies do not necessarily imply that the students are weak. Both the numerator and denominator come with a rather wide confidence interval, however, so the point estimate is rather uncertain.

Alternative interpretations

Our results are consistent with different colleges using the grade scale differently. It is, however, important to consider whether our results are consistent with other explanations as well. Various mechanisms have been suggested to us, all built on the belief that some colleges manage to lift students to a higher level of competence than other colleges on the grounds of good teaching quality. There is certainly reason to believe that there is variation between study programs with respect to teaching quality, but this in itself cannot account for our results. The relevant question is whether some colleges give their students some sort of quality that is not reflected in the students' grades, even if all schools use the same grading standard. This quality element must help the students do well at NHH. It can be good working habits, motivation, extracurricular knowledge or the development of critical thought. Such qualities seems closely related to the distinction between a deep-level and a surface-level approach to learning, first made by Marton and Säljö (1976a). A deep-level approach to

learning will lead to long term recall and make students more able to apply their knowledge outside the exam situation.

In theory, another possible mechanism is that some schools systematically fail to bring out the potential in their students so that when they start at NHH they tend to do better than expected, given their previous grades. From an econometric point of view, both mechanisms imply that the individual specific component in the error term is correlated with the college dummies and the school characteristics. Whatever the mechanism, table 3, column 4, shows that students doing better than expected tend to come from colleges with high admission limits and many student places.

Given the data available, we cannot rule out that the alternative mechanisms described above are relevant. We do not find it plausible, however, that differences in teaching quality are driving our results. One reason is that we see no *a priori* theoretical underpinning for why students doing better than expected should come from colleges with high admission limits and many student places. Another reason is that we find evidence consistent with different grading standards also within NHH. In Møen and Tjelta (2005, section 9) we present an analysis of 90 000 examination results from 1995 to 2003 for students in the Siviløkonom course. We group the examination codes for six electives, five fields of specialisation as well as "various subsidiary courses" and the "Siviløkonom thesis". In a corresponding manner to the analysis of differences in examination results between students from different colleges, we use average grades from mandatory courses in the second study block as an indicator of the student's ability. Thereafter we analyse how well students perform in the elective topics and in the specialised fields of the final study year. Our analysis reveals that subject fields that recruit strong students tend to award lower grades and the opposite. There is no reason to think that teaching quality vary systematically within NHH in a manner that can explain this pattern.

Consequences of differential grading standards

Table 3, columns 3 and 4 show that there is a systematic difference between students admitted with the same sum of points, but with background from different colleges. Our interpretation is that this principally reflects different grading standards. One may then argue that students from colleges practicing particularly strict grading should have extra compensatory points and students from colleges practicing particularly lenient grading should have corresponding

penalty points. It is hardly possible to construct a perfect admission system, and some degree of unfairness should therefore be accepted without spending resources on developing new routines. Before suggesting a policy change, therefore, it is important to quantify the extent of the problem. One way to measure the extent of the problem is to assess the number of students who would be affected by introducing an "ideal" admission system. Since we have admittance information for both admitted and rejected applicants, such an ideal admission can be simulated based on our regression results.

The rest of this chapter, and the associated tables 4, 5 and 6, are not included in this summary version of the paper. See information on the front page. Our main findings are as follows:

The simulation results indicate that a scant 10 % of the students are "error admissions". For approximately half of the colleges, the grades are so divergent that students would need to spend one to two semesters to earn the corresponding additional points.
The proposed error admission students have lower GPA, higher failure rate, more examination attempts and less progress than the students' in the control group.

Concluding remarks

In our main analysis, we have examined the academic performance of 529 individuals from 23 regional colleges admitted midway to the Siviløkonom course at a major Norwegian business school, NHH. Information on which college a student has as background helps us predict the student's academic performance at NHH even when we control for the student's admission points. Most likely this reflects that similar performances are rewarded differently at different colleges. The difference between the most and the least strict colleges is so large that it can take students several semesters to gain corresponding additional points. Furthermore, the analysis shows that as many as ten percent of the students admitted to the second study block at NHH from 1998 to 2002, may have been admitted at the expense of more competent students from stricter colleges. The students that our model classifies as "error admissions" experience a higher failure rate and a generally weaker performance than a control group of marginal "correct admission" applicants.

In this study we look at a sample of the best undergraduates only; those applying for admission to a leading Norwegian business school. Strictly speaking, we can therefore only

draw conclusions on how the college characteristics influence grading at the top end of the grade distributions. In Møen and Tjelta (2005, section 9), however, we analyse grade setting in the different subject areas at NHH and look at performance across the complete grade distribution. There we find corresponding tendencies. [11] The colleges that graded the applicants in our main sample teach a broad set of subjects, mostly in a classroom setting. We therefore believe that the grading problem we point to is representative for most grading situations.

Misallocation of talent can have important consequences. We know, for example that different types of education give very different private returns measured in salary. From the point of view of society, it is also important to allocate human capital in the most effective way possible.

How can the situation be improved? The differences in college grading standards appear to be systematically related to the quality of the student body and the size of the programme. Undergraduate programmes that recruit weak students award better grades for given performances. The same applies to some programmes with few student places. One solution is to move away from a centralised admission system and rather allow all institutions to assess applicants according to their own judgements. This is practiced for example, in Britain and the USA. This will probably bring about a widespread use of international, standardised tests and inflict significant costs on both the applicants and the institutions. A more obvious remedy is to implement measures that help censors form a realistic picture of what represents the national norm for different grades at each level and subject field.[12] One instrument is the use of external censors. However, all grades used in this analysis were awarded in collaboration with external censors. Hence external censorship is not sufficient. What may be the problem is that individual teachers and colleges choose their own external censors. In this way, a pattern of cooperation may develop whereby colleges with lenient grading standards

[11] In an analysis that is not reported, we also find that the failure limit is influenced by the quality of the students who take an examination or choose a particular subject field. Unfairness regarding who fails and who passes a course of study can therefore present an equally serious side of the problem as unfairness regarding admission to further studies.

[12] The Norwegian Association of Higher Education Colleges (UHR) is working on this at a general level. In order to harmonise the understanding of what the different grades involve, general "requirement specifications" have been drawn up. Within economic-administrative studies, the national council for economic and administrative fields (NRØA) has developed a subject-specific grading description. UHR's national faculty meetings have in addition appointed a reference panel that is to keep records of grade statistics and monitor its subject fields. However, none of these measures represents direct or continuous feedback to the individual censor and the grade setting of those concerned.

choose censors from other colleges with equally lenient grading standards and correspondingly so in strict colleges. A better remedy, which at the same time will produce interesting data for the authorities, is exams that are co-ordinated across schools nationwide. This is feasible in subject fields where many colleges offer similar programmes. One can hardly expect that all censors will have the same interpretation of the grading scale without some mechanism that provides systematic feedback to censors as to "where the bar is". In a companion analysis, we find that even within a large institution like NHH, it is difficult to achieve consistent grading standards across different study areas. Presently there is a lacking awareness of this problem.[13]

As a final remark, we would like to return to the alternative interpretation of our results discussed after table 3. We cannot entirely rule out that differences in teaching quality affect our findings as well as different grading standards. One particular possibility is that some study programs promote deep-level learning while others promote surface-level learning. Marton and Säljö (1976b) suggest that whether students choose one learning strategy or the other depends on how the assessment systems are designed. Co-ordination of exams across schools may be a tool to promote exam questions that reward deep learning. This way, improved average exam quality and learning may be a side benefit associated with the solution we propose.

[13] This is evident from the comments made by university and college rectors and administrators in the debate referred to in footnote 3.

References

Bagues, Manuel, Mauro Sylos Labini and Natalia Zinovyeva (2008). Differential Grading Standards and University Funding: Evidence from Italy. *CESifo Economic Studies*, **54**(2), 149-76.

Betts, Julian R. (1997). Do Grading Standards Affect the Incentive to Learn? Discussion Paper 22-97, Department of Economics, University of California, San Diego.

Betts, Julian R. and Jeff Grogger (2003). The Impact of Grading Standards on Student Achievement, Educational Attainment, and Entry-Level Earnings, *Economics of Education Review*, **22**(4), 343-52.

Baltagi, Badi and Qi Li (1990). A lagrange multiplier test for the error components model with incomplete panels, *Econometric Reviews*, **9**(1), 103-7.

Breusch, Trevor S. and Adrian R. Pagan (1980). The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics, *Review of Economic Studies*, **47**(1), 239-53.

Collin, Sven-Olof and Elin Smith (2007). The Relationship Between Grade and University Prestige – Reflections From a University College, Kristianstad University College Working Paper 2007:7.

De Paola, Maria (2009): Are easy grading practices induced by low demand? Evidence from Italy, Munich Personal RePEc Archive Paper No 14425.

Figlio, David N. and Maurice Lucas (2004). Do High Grading Standards Affect Student Performance? *Journal of Public Economics*, **88**(9-10), 1815-34.

Goldman, Roy D., Donald E. Schmidt, Barbara Newlin Hewitt and Ronald Fisher (1974). Grading Practices in Different Major Fields, *American Educational Research Journal*, **11**(4), 343-357.

Johnson, Valen E. (2003). *Grade Inflation: A Crisis in College Education*, New York: Springer

Hansen, Marianne Nordli (2005). Privatskole og karakterer på medisinstudiet, *Tidsskrift for Den norske lægeforening*, **125**(16), 2216-18.

Helson, Harry (1947). Adaption-level as a frame of reference for psyco-physical data, *American Journal of Psycology*, **60**(1), 1-29.

Helson, Harry (1948). Adaption-level as a base for a quantitative theory of frames of reference, *Psycological Review*, **55**(6), 297-313.

Hu, Shouping (ed.) (2005). *Beyond Grade Inflation: Grading Problems in Higher Education*, ASHE Higher Education Report, **30(**6), San Francisco: Jossey-Bass.

Hægeland, Torbjørn, Oddbjørn Raaum and Kjell G. Salvanes (2005). Pupil achievement, school resources and family background, IZA Discussion Papers 1459.

Marton, Ference and Roger Säljö (1976a). On qualitative differences in learning I. Outcome and process, *British Journal of Educational Psychology,* **46**, 4-11.

Marton, Ference and Roger Säljö (1976b). On qualitative differences in learning II. Outcome as a function of the learner's conception of the task, *British Journal of Educational Psychology,* **46**, 115-27.

Moen, Espen R. and Lone Semmingsen (1996). Utdanning og livsløpsinntekt, SNF-rapport 96/96, Oslo.

Møen, Jarle and Martin Tjelta (2005). Bruker ulike høgskoler karakterskalaen ulikt? En analyse av sammenhengen mellom skolebakgrunn og faglig suksess ved NHH, *Økonomisk forum*, No. 6, 33-45.

Table 1: Number of examination observations, spring 1998 – spring 2003

	Number of examinations	Number of students	Number of examinations per student
All students			
Electives	4 621	2 041	2.26
Second block	25 494	2 313	11.02
Third block	8 119	1 426	5.69
Total	**38 234**	**2 313**	**16.53**
Four-year students			
Electives	3 894	1 620	2.40
Second block	19 380	1 784	10.86
Third block	5 761	1 055	5.46
Total	**29 035**	**1 784**	**16.28**
College students			
Electives	727	421	1.73
Second block	6 114	529	11.56
Third block	2 358	371	6.36
Total	**9 199**	**529**	**17.39**

Table 2: Information on colleges with applicants to the second study block at NHH 1998-2002

College	Total admission 1998-2002	Min. yearly admission	Max. yearly admission	Rejected applicants 1999-2002	Min. yearly rejection	Max. yearly rejection	Min. admission limit	Max. admission limit	Min. study places	Max. study places
A	71	10	17	37	3	13	42.4	46.2	260	280
B	54	0	25	53	4	23	–	–	–	–
C	47	8	12	52	9	20	Open	46.9	235	245
D	47	6	11	39	0	16	Open	44.2	50	50
E	42	5	10	26	4	9	47.3	52.3	50	60
F	32	5	8	25	3	10	Open	37.0	105	115
G	31	2	8	35	5	15	Open	41.1	90	110
H	28	3	9	35	6	11	45.0	48.0	125	125
I	28	3	8	17	1	10	Open	45.4	60	75
J	23	2	8	32	6	10	39.6	45.2	60	60
K	23	1	7	26	4	11	Open	41.1	110	115
L	16	2	5	21	4	7	Open	41.1	90	165
M	13	1	4	12	0	6	Open	38.0	55	65
N	13	0	6	35	5	12	41.4	47.1	130	130
O	11	0	5	5	0	3	33.3	42.8	35	35
P	9	0	4	3	0	2	Open	Open	60	90
Q	9	0	6	7	1	3	Open	Open	80	130
R	8	1	2	8	0	5	Open	Open	65	80
S	6	0	3	19	0	13	–	–	–	–
T	6	0	3	0	0	0	–	–	–	–
U	5	0	2	15	1	9	Open	Open	100	100
V	5	0	2	3	0	1	Open	Open	20	55
W	2	0	1	7	0	5	Open	41.6	120	150
All	529	96	117	512	114	146	Open	52.3	20	280

The admission limit is points for admission from upper secondary school to the two-year economic and administrative study programme at the respective colleges from 1996 to 2000 (two years prior to admission to the second study block at NHH). Student places are the number of student places available for admission to the two-year economic and administrative study programme at the respective colleges from 1996-2000 (two years prior to admission to the second study block at NHH). “Open” means either that the course is defined as open, or that all qualified applicants are admitted. Information on admission limits and number of student places is derived from different annual publications of the “Application hand book” published by the centralised application processing centre (Samordna opptak). Admission information from college B and T is not available. “College S” is not a single college, but includes students with grounds for admission based on study at several schools. Rejected applicants are only included in the selection if admission points can be calculated. Information on rejected applicants’ college background is not available for 1998. These observations are therefore excluded. Since part of the information given in this table is publicly available, the colleges are in later tables labeled “college 1-23” (with a different sequence) to assure anonymity.

Table 3: The effect of students' admission requirements on examination grades achieved at NHH

	(1)		(2)		(3)		(4)	
Constant	4.597***	(0.028)	4.597***	(0.028)	8.555***	(0.363)	6.973***	(0.474)
College student‡	-0.128***	(0.033)						
College no. 1‡			-0.118	(0.119)				
College no. 2‡			0.209**	(0.096)	0.195*	(0.109)		
College no. 3‡			-0,456***	(0.118)	-0.324***	(0.120)		
College no. 4‡			0.029	(0.124)	0.054	(0.124)		
College no. 5‡			-0.322**	(0.144)	-0.154	(0.134)		
College no. 6‡			-0.033	(0.153)	0.109	(0.139)		
College no. 7‡			-0.107	(0.152)	-0.145	(0.140)		
College no. 8‡			-0.059	(0.144)	-0.080	(0.134)		
College no. 9‡			-0.005	(0.167)	0.252*	(0.147)		
College no. 10‡			-0.482***	(0.165)	-0.480***	(0.146)		
College no. 11‡			-0.345	(0.224)	-0.249	(0.184)		
College no. 12‡			-0.264	(0.200)	-0.365**	(0.168)		
College no. 13‡			0.103	(0.220)	0.315*	(0.179)		
College no. 14‡			-0.212	(0.247)	-0.039	(0.200)		
College no. 15‡			0.34	(0.269)	0.246	(0.214)		
College no. 16‡			-0.618**	(0.262)	-0.512**	(0.204)		
College no. 17‡			-0.244	(0.282)	0.121	(0.220)		
College no. 18‡			-0.779**	(0.357)	-0.588**	(0.272)		
College no. 19‡			1.176**	(0.569)	0.338	(0.429)		
College no. 20‡			-0.089	(0.351)	-0.087	(0.264)		
College no. 21‡			-0.111	(0.112)	0.139	(0.118)	1.582	(0.380)
College no. 22‡			-0.338	(0.321)	0.089	(0.245)	1.529	(0.439)
College no. 23‡			-0.564*	(0.323)	0.099	(0.250)	1.523	(0.441)
Admission points					-1.857***	(0.144)	-1.812***	(0.142)
Age					-0.150***	(0.014)	-0.142***	(0.014)
Open study‡							1.220***	(0.375)
School points/10							0.293***	(0.085)
Student places/100							0.109***	(0.034)
no. of observations	38 234		38 234		9 199		9 199	
R^2	0.069		0.075		0.187		0.172	

The dependent variable is examination grade. Regressions are estimated by random effects GLS. Standard errors are given in parenthesis. Dummies for 22 different subjects are included in all regressions but not reported. BEA200 (financial accounting and financing) is the excluded category. In the last column, the dummies for colleges 21-23 are kept since we do not have data for admission requirements and number of student places for these colleges.

‡ indicates dummy variables

* means significant at 10% level ** means significant at 5% level *** means significant at 1% level