March 2022

Variation in high school grading: Evidence from Alberta, Canada*

David R. Johnson

Department of Economics

Wilfrid Laurier University

Waterloo, ON, Canada

Abstract:

Low grading standards at a school, over grading, are in place when students with a given level of skills receive higher average school-awarded grades than students with the same skills would receive at a different school with higher grading standards. If over grading occurs in all courses at a school, that school is defined in this paper to have an over grading culture. Estimating relationships between school-awarded grades and average provincial examination grades in up to 10 courses per high school per year over 14 years in Alberta, Canada allows precise measurement of school grading cultures. Variation in high school grading cultures is large enough to affect post-secondary opportunities because average high school grades are the determining factor for entry into more desirable post-secondary opportunities. Schools with low grading standards, the over grading schools, are more likely to be private, rural, and operated primarily for students who return to school after dropping out. Over grading schools have smaller course cohorts and have a smaller percentage of lone parent households and a larger percentage of well-educated parents.

*Very useful comments were received at the Association for Education Finance and Policy meeting in 2020 as well as from colleagues at Wilfrid Laurier University, the University of Waterloo and the National Bureau of Economic Research Education group.

1. Introduction

High school grading cultures, as defined in this paper, exist when students with the same skill level are awarded systematically higher or lower grades across all courses if graded at different schools. In language more often used in the literature, an over-grading school culture has low grading standards not just in one course but in all courses at the school, and an under-grading school culture has high grading standards not just in one course but in all courses at a school. This paper measures high school grading cultures across Alberta high schools and finds consequential variation in grading standards common to all courses between schools. Schools have grading cultures.

High school grades matter enormously in Alberta. High school grades are, with rare exceptions, the most important determinant of undergraduate admission for non-international students. Appendix A presents illustrative variation in average admission grades for the three universities in Alberta continuously open over this period as well as examples of admission averages and cut-off grades for 2022 at these universities. Students with low average high school grades have no chance of admission to the two most desirable universities. High and even very high average grades are required for admission to the more desirable programs in nursing, engineering, or business within any university in Alberta. The grades studied in this paper are for the highest stakes. Students are very aware of the grades required for admission to the various universities and programs. The observed variation in average grades across high school grading cultures is well within the range that moves a student from admission to non-admission in different universities and different programs. The degree program from which a student graduates plays a significant role in their future income prospects.[1] The grading culture of your high school could change your life prospects.

The grading culture of your high school likely matters beyond immediate post-secondary admission prospects. A student might receive their high school diploma at an over grading (low standards) school and not graduate at an under-grading (high standards) school. Students and parents could choose a school based on its grading culture rather than the school's influence on the actual accumulation of student skills. This could lead to a lower level of actual skills for the student and a change in your life outcome. If a student at an over-grading school had a target numerical grade, they could reduce effort in high school and accumulate fewer actual skills facing costs later in life.[2]

This paper estimates relationships between average school-awarded grades and average provincial examination grades with a set of 40,934 school-course-year observations over 10 final-year high school courses at up to 533 schools over 14 academic years from 2005-06 to 2018-19. There are many observations of courses in different subjects at the same school. For each school-course-year observation the average grade on an anonymously graded examination set by the province in that course and the average grade awarded in that course by the teachers at the school to the same students are known. The grades are used to estimate a direct measure of over and under grading, grading standards, at each school that applies to all courses in the school

[1] Finnie et.al (2019) document the large variation, by undergraduate program, in the earnings of graduates from the University of Ottawa over the years studied in this paper. There is no reason to think earnings patterns by program would be substantially different in Alberta then in the locations where University of Ottawa graduates live.

[2] The consequences of low and high standards are discussed further in Gershenson (2018).

over the 14 years.[3] By construction, the average school neither over grades nor under grades. Over grading or under grading is not the behaviour of a specific teacher but of the group of teachers teaching a group of courses at the school over the whole period.

High schools have strong "grading cultures". The gaps between average grades across all courses awarded by teachers at different high schools to students of the same ability (the same average examination grades in the courses are the measure of ability) are large enough to be of both practical and statistical significance. One standard deviation around a zero mean of the various measures of over and under grading is about three percentage points on grades scaled between zero and one hundred percentage points, the grading scale used for post-secondary admission (see Appendix A). This means it is common that the average school-awarded grade in all courses at one school is roughly three percentage points higher (or lower) in a group of courses at another school when students at both schools have the same average examination grades in the same courses.[4] The school awarded grades are 50% of the weight in the calculation

[3] Appendix B presents averages of school assigned and examination grades in the years studied. There is minor variation and no obvious trend in average examination grades. There is a small increase in average school assigned grades concentrated at the end of the years of study. Similar patterns in school grading cultures are found when academic years are studied individually or in different groupings of academic years.

[4] Grade variation is frequently scaled by the standard deviation of individual student grades when all grades are standardized. In this study average grades across schools are not standardized with this method. Average grades are reported for widely varying numbers of students in courses at each school. Alberta Education reports the standard deviation of individual grades in most courses (excluding 2012-13 and 2013-14 for unknown reasons). In the non-excluded years, the range of standard deviation of individual final grades across all students in a course is between 8.5 and 16.1 percentage points in the different courses and years. The average of the standard deviations is 12 percentage points. If school-awarded grades in one school are 3 percentage points higher than at another school and school-awarded grades are 50% of the weight of final grades, then final grades at the higher grading school are 1.5 percentage points higher, about 0.125 of a standard deviation of individual grades, than at the lower grading school, a change in grades that applies to all diploma courses taken at the school. This is a second way to scale the school over and under grading measures.

of the student's final grades in the courses in the first ten years. Thus, a three-percentage point difference in the average grade awarded in all courses at the school is a one and one-half percentage point difference in the average of final grades, a change easily large enough to move you from one university to another across the admission averages presented in Appendix A or in or out of your desired program as discussed in the text below the table in Appendix A.[5]

The level of over grading and under grading is clearly associated with observable school-level characteristics. The largest quantitative effects show that private schools and non-conventional schools over grade by large amounts, two to three percentage points. Non-conventional high schools enrol mostly students that previously dropped out of high school and are returning to complete a high school diploma.[6] High schools located outside in the two large census metropolitan areas of Calgary and Edmonton, the only large cities in Alberta, over grade by about one percentage point.[7] Three continuous variables describe schools: the percentage of lone parent households associated with the school; the percentage of parents with a university degree associated with the school and the average size of diploma courses taught in the school. There is under grading in schools with larger average course sizes. There is under grading in schools with a larger percentage of lone parent households. There is over grading in schools with a larger percentage of parents with completed university degrees. Since over grading schools

[5] When the weight on school-awarded grades rises to 70% in the last 4 years of the study, the variation in school-awarded grades increases in importance.

[6] The term non-conventional is invented for this paper, not a term used by Alberta Education. Some non-conventional schools are entirely online; others are a mix of online and in-person. They are mostly operated by public-school boards. Conventional high schools have similar proportions of Grade 10, 11 and 12 students. In Alberta, high school starts in Grade 10. Non-conventional high schools are defined as having more than 50% of their students in Grade 12. That percentage is usually much higher, as high as 100% in some cases.

[7] In 2016 the Calgary CMA (Census Metropolitan Area) had a population of 1.4 M; the Edmonton CMA had a population of 1.2M and the next largest urban centre, Lethbridge, had a population of only 117,000.

with lower standards tend to be private, have a smaller proportion of lone parents and a larger proportion of parents with a completed university degree, students growing up with advantages appear to receive further advantages by being more likely to attend a school with an over-grading culture. There is systematic over grading in rural schools. Whether over grading in rural schools is an equity issue depends on whether you perceive growing up in a rural area as a social and economic advantage or disadvantage.

Studying the grades from Alberta high schools has advantages and disadvantages relative to previous studies of grading standards in the literature. Measurement of grading standards requires an independent measure of students' ability from outside the students' school. In Alberta there are grades awarded on an outside provincial examination, blind graded, in 10 end-of-course (EOC) "diploma courses" taught in Alberta high schools. Diploma courses are the final high school courses in their respective subjects but need not be taken in the final year of high school.[8] There are multiple observations of the different courses at the same school.[9] The measure of grading standards at each school, the numerical measures of over and under grading, apply to the average of the school-awarded grades in all courses. Using multiple courses from a school creates a measure of school grading standards estimated using more than one course, up to 140 course-year observations are available at a school. This improves on Gershenson (2018, 2020) and Tyner and Gershenson (2020) whose end-of-course (EOC) external examination is in one

[8] Ten courses are used in the analysis. The use of the phrase "all courses" here means all courses where there is an EOC examination at the school. There are 12 diploma courses, but very few schools offer the two French courses. Tyner and Larsen (2019) survey the use of EOC examinations in the United States.

[9] The groups of students at a school do not correspond to classes except in a very small high school. School results are reported by academic year and most schools would have classes in a subject in both semesters, some with the same teacher and some with multiple teachers. This is the reason to use the term "school grading culture," the unit of observation is not a class.

course (Algebra I) in North Carolina. Tyler and Gershenson (2020) introduce the term static grade inflation, higher than deserved grades from a school at a point in time compared to another school based on individual student results in one mathematics course with an EOC examination. My Alberta study extends this concept across multiple courses at each school. Betts (1998) and Betts and Grogger (2003) also have only a mathematics grade available for inter-school comparisons. Their external skills measure is the score on a national assessment of more general mathematics skills not directly tied to the curriculum in the different mathematics courses at different schools. Hurwitz and Lee (2017) use composite SAT scores as their independent measures of skills then compare that score to an overall high school GPA. Allensworth and Clark (2020) measure grading standards using the variation in post-secondary success across students from different public high school in Chicago where the students have the same average grade (GPA). This is an indirect measure of school grading culture, a higher college graduation rate at the same GPA is a high grading standard, that is an under-grading school. It is a significant advantage of my paper that the Alberta EOC examinations are so clearly tied to the curriculum of each course. The curriculum is known to teachers, and teachers are expected to prepare students for the curriculum-based examination in the 10 different subjects. The grading data are from 2005-06 to 2018-19, a multi-year sample.[10] The other paper in the literature with the advantage of a one-to-one correspondence between the external examination and the course level material is Figlio and Lucas (2004) who use Florida elementary school system reading and

[10] It is also convenient in estimation and interpretation that the grades on both the Alberta provincial examinations and the course assigned grades use the same 0-100 percentage scale. It is helpful as well that the average examination grade and the average school-awarded grade in each-course-school-year observation are completely separated. In Gershenson (2018, 2020) it is a small complication that the transcript grade includes, with an unknown weight, the EOC state exam grade. The school-awarded grade is not directly observed.

mathematics assessments. As already noted, a second advantage of studying over and under grading in Alberta over Figlio and Lucas (2004) as well as the other papers in the literature is that the courses in Alberta high schools have very clear stakes for students.[11] The EOC examinations in Alberta constitute 50% or 30% of the final grade in courses that are required courses (in different combinations) to graduate high school and many courses are needed for admission to specific and often highly desired programs where positions are limited in number and rationed by high school grades. Students will clearly put effort into these examinations.[12] Finally, it is, in some ways, an advantage that the observations in Alberta are school by course by year. The measurement of grades is at the school level, not at the teacher or classroom level, and some potential selection issues within the school are resolved.[13]

The use of Alberta high school grading outcomes has disadvantages. There are no individual student data. There are no data on longer-term or shorter-term student outcomes of interest, for example, grades in subsequent courses, graduation, post-secondary attendance or

[11] There are no known stakes relating to examination results for teachers or schools. Funding is provincial and allocated per student. Teacher salaries reflect only seniority within their boards. The likelihood of a teacher being removed from their job for a poor student performance on examinations in a public or Catholic school is effectively zero. Teacher's unions in Alberta are very powerful.

[12] The comparison of teacher predicted grades on final examinations in English schools by Murphy and Wyness (2020) has some similarity to the situation in Alberta. In England, the external examination result is the final grade in the course. Their comparisons of predicted grades to final grades also reveals school grading cultures in England in the sense of school-level commonalities across grading variation in predicted grades relative to final grades. Murphy and Wyness (2020) do not stress the existence of school grading cultures.

[13] Under the very specific model in the appendix, all selection issues are resolved when the skills measured by the examination and the school-awarded grade completely overlap in each course. That is a very strong assumption but, to some degree, this strong assumption is implicitly made in the rest of the literature whenever teacher-awarded grades and outside grades are compared. The appendix makes this assumption explicit in a linear model relating grades to skills. The over and under grading measures are perfectly valid if the teacher grade and the examination grade each have a linear relationship to the same group of skills with positive but possibly different weights on each skill component.

completion, or earnings. Much of the existing literature continues past the creation of measures of over and under grading, the measurement of grading standards, to measure the relationship between grading standards and future outcomes. In panels of individual student data with annual assessments, both Figlio and Lucas (2004) and Gershenson (2020) show that higher standards increase learning growth rates in a value-added model. Betts (1998) and Betts and Grogger (2003) use individual data to show that students exposed to high grading standards earlier in high school increase skills as measured on the national mathematics assessment and make larger gains in mathematics skills during high school. They find no effect on graduation, college attendance or earnings. With individual data, these papers can disaggregate results by type of student, a student's place in the initial ability distribution or a student's race. This analysis cannot be done in Alberta where all grades are aggregated to the school-level unit of observation.

The study of the Alberta high school grading standards contributes to describing and understanding the interaction between observed measures of grading standards and observable school characteristics. Most of the literature measures the change in the gap between the high school grades and the single external measure of skills over time while sorting high schools by Social and Economic Status (SES) characteristics. Hurwitz and Lee (2017) compare the change in SAT scores to the change in average grades in a group of courses across a large group of American schools. Their key result is finding a large increase in Grade Point Average (GPA) when there is no change in the SAT scores (dynamic grade inflation in their language). A major concern is that the largest amount of GPA growth in found in schools that are richer and whiter as well as private (including religious private). Betts and Grogger (2003) are also concerned with the impact of differential grading standards in mathematics at different points in the skill distribution and by race. Gershenson (2018) focuses on changes in differential grading standards when schools are sorted by income and race. Disadvantaged students experienced less grade inflation in his relatively short sample. Grade inflation further disadvantaged already disadvantaged students. There is a very large literature how (static) grade inflation in cross section does or does not affect disadvantaged students more. The gaps between the school-assigned grade and the blind-graded external measure of skills (usually a course related examination of some type) from a single point in time are studied. Results vary as to whether students with lower social and economic status are over graded, under graded or neither. Studies finding over grading of lower SES students (lower standards) include Botelho, Madiera and Rangel (2015), Gibbons and Chevalier (2008), Himmler and Schwager (2012) and Rauschenberg (2014). Studies finding under grading of lower SES students (higher standards) include Marcenaro-Gutierrez and Vignoles (2015) and Rangvid (2015). Lavy (2018) finds no relation between SES measures of disadvantage and grading standards. Burgess and Greaves (2013), using very detailed English data, find relations between teacher assigned grades relative to examination grades and ethnic characteristics where some ethnic groups receive lower grades and other ethnic groups receive higher grades from teachers' conditional on the student's grade on the blind-graded external examination.

The paper proceeds as follows. The model of school assigned grades conditional on examination grades is estimated and results are presented. A conclusion follows. There are several appendices.

2. Method

The relationship estimated between the average school awarded grade in course j at school k in year t, denoted $\overline{\mathrm{S}}_{jk(t)}$, and the average examination grade in course j at school k in year, denoted $\overline{\mathrm{E}}_{jk(t)}$, is

$$\overline{\mathrm{S}}_{jk(t)} = \mathrm{D}_{j(t)} + P_k + \gamma_j \bar{E}_{jk(t)} + \varepsilon_{jk(t)} \qquad (1)$$

Course fixed effects, $D_{j(t)}$, vary by year and course. Variation in $D_{j(t)}$ across courses j within a year is interpreted as mostly uninteresting variation in scaling of both school and examination assessments across different courses as skills are transformed into grades differently in different courses (the detailed model behind equation (1) is presented in Appendix C). Variation in $D_{j(t)}$ within the same course "j" across time is more controversial. The provincial EOC examinations are scaled to be equal in difficulty from year to year so that variation in the average provincial grade is supposed to measure changes in the average skills of students in the province from year to year. But it is certainly possible, even likely, that the examinations are not scaled perfectly. This means if students did not vary in quality as a group from year to year, then the same examination grade would correspond to different skill levels by subject between years and $D_{j(t)}$ would vary over time. School awarded grades could also vary in their average from year to year at the same student skill level.[14] These issues are further explored in Appendix C but are not

[14] This paper does not focus on the changes, if any, in grading standards over time. See Appendix B for some very aggregate evidence that average examination grades remained similar over time. There is a small increase in school-awarded grades over time. These changes appear to be a smaller part of the story in Alberta in diploma courses, possibly because the presence of the external examination acts to prevent or at least reduce grade inflation in diploma courses. There is some evidence that all school-awarded grades jumped in the four years after 2015-16. That period coincides with the re-weighting of school-awarded grades from 50% to 70% of the final grade. If a change occurred at all schools with the reweighting, then the $D_{j(t)}$ terms would have a shared component across all courses within each year or across the last four years. That analysis

important in this paper beyond requiring time-varying course fixed effects in the estimation of (1). Allowing the slope coefficients γ_j in (1) to vary by course is conceptually important (see Appendix C). This coefficient allows the relationship between changes in the average school awarded grade and the average examination grade to differ by course.

The parameters of interest in this study are the school fixed effects, P_k. Note there is no time parameter on P_k. The school fixed effects are in place across the years for the period of estimation. If at school k, the value of P_k is three, then average school-awarded grades in all courses offered at the school are three percentage points higher than the average school-awarded grades predicted for that school controlling for course fixed effects that vary by year and controlling for that school's actual average grades on the examinations in each course in each year. A value of $P_k = 3$ says school k is an over grading school. This school has low grading standards across all courses. The intuition: when the school-awarded grade and the examination linearly reward a common set of skills (Appendix C), the examination grade is a sufficient control variable that allows accurate estimates of the over and under-grading parameters, the values of $\widehat{P}_k$. The average examination grade functions as a control variable even in the presence of variation in home inputs to learning and selection of students into schools when the two grades reward a set of overlapping common skills even with different weights. When a value of $\widehat{P}_k$ (the estimate of P_k) is significantly positive at school k, this is an over grading (low standards) school. When $\widehat{P}_k$ is significantly negative, this is an under-grading (high standards) school. When $\widehat{P}_k$ is not significantly different from zero, this is a school that neither under grades nor

is left for future work and in this paper is modelled as part of changes in the course-year fixed effects $D_{j(t)}$ in the last four years as part of variation in course-year fixed effects across all years.

over grades. The school fixed effect measures over and under grading, the school's "grading culture" parameter.

Equation (1) is estimated using the techniques pioneered by Abowd, Kramarz and Margolis (1999) – henceforth AKM. The AKM model estimates a set of course fixed effects, the values of $\widehat{\mathrm{D}}_{j(t)}$, that vary by both course and year. The AKM model also estimates the slope coefficients relating the change in the school awarded grade to the change in the examination awarded grade, the values of $\hat{\gamma}_j$, and a set of school fixed effects, $\widehat{\mathrm{P}}_k$. AKM estimates robust standard errors of the school fixed effects and the slope coefficients.[15] In AKM the school fixed effects are normalized to have an average value of zero across schools. If $\widehat{\mathrm{P}}_k$ = 3 then the average grade awarded by teachers in all courses at school k is estimated to be three percentage points higher than the predicted school-awarded grades on those courses. The prediction is made using the average examination grades observed at that school in that course and the course-year fixed effects for those courses.

To discover if school-level characteristics are systematically associated with larger and smaller values of $\widehat{\mathrm{P}}_k$, the estimated values of $\widehat{\mathrm{P}}_k$ are regressed on school characteristics using[16]

$$\widehat{\mathrm{P}}_k = a_0 + a_1 I_k + a_2 SES_k + a_3 C_k + \varepsilon_k \qquad (2)$$

I_k is a series of indicator variables active when school k is a specific type of school: an unconventional school, a private school or an urban school. SES_k represents two continuous SES

[15] The routine in Stata is felsdvregdm.

[16] To keep the notation less intensive, variation between the true value of P_k and the estimated value of $\widehat{\mathrm{P}}_k$ is incorporated into ε_k in expression (2). Equation (2) is similar to regressions used in the AKM literature for explanations of productivity that vary by workers across characteristics of firms. Dhuey and Smith (2018) estimate a similar regression to explain school fixed effects that vary with the tenure of specific school principals.

variables. One is a measure of the percentage of all households associated with the school that are lone parent households. The other is a measure of the percentage of adults associated with the school with a completed university degree. C_k is a measure of the average number of enrollees in diploma courses in the school over the period studied. The error terms in (2) are treated as heteroscedastic and robust standard errors are estimated for the parameters of interest.[17]

3. Results

Table 1 presents descriptive data on all 533 high schools with diploma courses operating in Alberta between 2005-06 to 2018-19. Many schools (certainly not all) schools operate in all years.[18] Conventional schools are high schools with enrolment in Grades 10, 11 and 12 in roughly equal numbers. High school students in Alberta enter a conventional high school in Grade 10 and leave after obtaining sufficient credits and the right set of courses to graduate, usually, but not always, after three years.[19] Conventional schools educate most students. The

---

[17] It is also possible to substitute equation (2) directly into equation (1) and estimate

$$\bar{S}_{jk(t)} = a_0 + D_{j(t)} + \gamma_j \ \bar{E}_{ijk(t)} + a_1 I_k + a_2\, SES_k + a_3\, C_k + \varepsilon_{jk(t)}$$

The course-specific time-varying constant term, the sum of $a_{0(t)}$ and $D_{j(t)}$, is not of central interest. One advantage of the equation above is that, unlike the AKM estimator, the equation can be estimated with and without weights following the suggestions of Solon and Woodridge (2015). Standard errors can be generated clustering the observations in different ways. One difference between the equation in the footnote and equation (2) in the text is that equation (2) in the text uses one observation per school, implicitly weighting all schools equally in the estimates of a_1, a_2 and a_3. In the equation above, schools vary by the number of examinations in the group of years included and schools with more examinations can receive a larger weight in the estimates of a_1, a_2 and a_3. Estimates of a_1, a_2 and a_3 using this methodology are available on request.
[18] Results from courses with less than 6 students enrolled in a year are not reported for privacy reasons, otherwise results are reported publicly by course. There is considerable population growth in Alberta over this period so new schools are opened every year.
[19] There is a significant drop out rate from Alberta high schools. There are also students who remain in conventional high schools for a fourth year and are counted as Grade 12 students. Such

remaining students are found in non-conventional schools, a concept already introduced as schools where all or a very large proportion of their students are in Grade 12. These schools meet the needs of students who have dropped out and then return to complete high school courses required to graduate. There are many non-conventional schools, but their total enrolment is small. Table 1 also shows a relatively large (for Canada) private school sector in Alberta.[20] In the last column, the number of schools of each type that have results for 10 or more courses are presented. Estimates of the school fixed effects in the AKM framework exist with only one observation from a school but the estimates might be considered more interesting and have more meaningful standard errors when there are multiple courses from a school. Some results are presented later restricting the sample by excluding schools with less than 10 courses over all years to check for robustness on this dimension.

The second last and third last columns of Table 1 present averages of examination and school awarded grades from different types of schools. Averages of school assigned grades are much higher than the average grades awarded by the markers on the anonymously graded provincial examination in every case.[21] Average examination grades are much higher at private schools than at non-private schools. This is not surprising when the social and economic advantages of students attending private schools are presented later in the paper.

---

students remain to obtain additional credits, to increase their average grade or for social or athletic reasons.

[20] In Alberta private schools take fees from parents and receive a provincial grant per student taught. The grant per student is lower than the grant received by a public school. It is unclear whether resources per student are higher or lower at private schools than public schools. The Catholic school system in Alberta is not part of the private school system. It receives full funding and does not charge fees. I exclude the very small and fully publicly funded francophone system from this study.

[21] This is consistent with other literature comparing teacher grades to grades on examinations graded anonymously.

Table 2 presents AKM estimates of equation (1) using the three groupings introduced in Table 1. The estimates of the course-specific slope coefficients γ_j from equation (1) are estimated including all schools.[22] The slope parameters in equation (1) are precisely estimated. There are positive and precisely estimated relationships between an increase in the average grade in a course in a school in a year awarded at the school by the school's teachers and an increase in the average grade awarded to the same students on the provincial examination. It would be very surprising (perhaps even shocking) if such relationships did not exist. The provincial examination is based on the course's provincial curriculum in each subject. Teachers are supposed to teach and assess the provincial curriculum as tested on the examination. The estimated slope coefficients vary by course and vary slightly by subset of schools. The table introduces the concept of "university courses." The six "university" courses are the courses more closely associated with admission to university post-secondary programs.[23] It is conceivable that behaviour around awarding school grades in the "university courses" could be different that in the other courses where the stakes for students might be considered lower. This is not the case.

The estimated school fixed effects, the measures of over grading and under grading measuring the school's grading culture, are the parameters of central interest. Recall that an

[22] The values do not change if the schools with less than 10 courses are excluded. The excluded schools are a very small proportion of the observations.

[23] The requirements to graduate from high school are credits in one of the two English courses, 30-1 or 30-2; one of the two mathematics courses 30-1 or 30-2; one of the two social studies courses 30-1 or 30-2 and one of the four science courses. However, admission to more competitive university programs generally require the 30-1 versions of English, Math and Social Studies courses as well as one or more of the three named sciences rather than the more general Science 30 course. There are 12 diploma courses in Alberta. There is a diploma course in the French Language, that is French for non-francophones, and a diploma course in French for francophones. Both French courses are taught at very few schools and are excluded from the analysis. The term "university" course is not a formal term used by Alberta Education. That term was created for this paper.

estimate of a school's fixed effect of plus (minus) three says that average school assigned grades in any course at that school is three percentage points higher (lower) than at another school with the same average examination grades. The school fixed effects have an average value of zero by construction in the AKM estimator. Table 2 shows that for the set of schools including schools with less than 10 courses, the standard deviation of estimated school fixed effects is about three percentage points. Figure 1 (upper left) presents histograms of the school fixed effects using all schools and all courses and then histograms that separate out conventional schools, private schools and rural schools that are used later. The histogram in the upper left corner shows that for all schools, school fixed effects have values from -10 to +10 percentage points. These are not small values. The count of schools is on the vertical axes of Figure 1. Many schools have values of over and under grading parameters, the school fixed effects, that would matter to student average grades. The third last column of Table 2 reports more than half the estimated school fixed effects have robust t-statistics larger in absolute value than 2 in the AKM framework. The last two columns of Table 2 present the standard deviation of school fixed effects and the proportion of those fixed effects statistically different from zero when the sample is restricted to schools with more than 10 courses. The school fixed effects are less variable in the restricted sample but for the all-schools and conventional schools, a slightly higher percentage are statistically different from zero.

It is easy to imagine a situation where there are no school grading cultures with practical consequence.[24] One might even expect this to be the case. There are up to 140 courses at the same school in the 14-year period. Many different teachers might contribute to the school-

[24] The histogram(s) in Figure 1 would have schools bunched, for example, within one percentage point of zero. School fixed effects could still be statistically significant, just not of much practical consequence.

awarded grades in a school over 14 years. In many schools there could be more than one teacher covering a course. You would expect that even within a school, some teachers have low standards (are over graders) and some teachers have high standards (are under graders) relative to the provincial standard set on the examinations within a course-subject areas or across course subject areas. The fact that many schools have fixed effects that are statistically different from zero spread far from zero over many groupings of courses and schools is strong evidence of the existence of consequential grading cultures at Alberta high schools. Providing a method that documents the size of these school fixed effects and grading cultures is a key contribution of this study.

Before moving to discussion of the estimates of equation (2), Table 3 summarizes the three continuous variables that are used to describe characteristics of schools. Two school characteristics have already been presented in Table 1, 153 schools are non-conventional, and 48 schools are both conventional and private. Two variables describing schools are derived from the 2006 and 2016 Canadian censuses. For schools operating between 2005-06 and 2015-16 the percentage of lone parent households in a school community and the percentage of adults with a completed university degree in a school community are calculated using lists of the number of students at each school in each year residing in each census dissemination area (DA) of the 2006 census.[25] Student locations by DA using the 2016 census are used for the same calculations for the years from 2015-16 to 2018-19.[26] The DA is the smallest census unit reporting data as averages of the unit, roughly 700 people normally live in a DA. The SES variables for the

[25] Counts of persons within a DA with completed university degrees is the measure of education that is consistent between the 2006 and 2016 census.
[26] These are all students enrolled at the school, not the students enrolled in the diploma courses. That information is not available from Alberta Education.

school in each year are weighted average of DA characteristics of adults or households who live in the DA, weighted by the number of school attendees from each DA. Then Table 3 presents averages, standard deviations and the number of observed units of the census-based variables in two ways. One measure uses each school as the unit of observation and presents statistics on school means and the standard deviations of the school-unit means. The second measure, instead of taking simple averages across the school unit observations, each school observation weighted by the number of students enrolled at schools. This calculation describes the population sending students to high schools. Two facts emerge from this analysis. First, larger schools have more highly educated parents. Second, the level of education increased substantially in Alberta between the 2006 and 2016 measure. This makes it necessary to construct the following variables to describe schools as follows. At each school in each year, the school's percent of adults with a completed university degree (percent of lone parent households) was subtracted from the population percent of adults with a university degree (percent of lone parent households) sending students to high school.[27] Thus the continuous variables in equation (2) are expressed in percentage points of adults with a completed university degree (percentage points of single parent households) as deviations from the means described above. Then these variables are averaged across whatever years the school is operating among the 14 years. A positive (negative) value is a school with more (less) educated parents or more (less) lone parent households. The final important note from Table 3 is that the percentage of parents with completed university degrees at private schools much higher than at public schools while the percent of lone parent households is only slightly lower at private schools. Table 3 also shows

[27] Nearly identical variables take the school mean of the variable, subtract the provincial mean and divide by the provincial standard deviation. Results using this method of joining the data measured across the two censuses yielded identical results.

that 66% of all schools are rural and 40% of private schools are rural. Finally, Table 3 shows that the average size of a diploma course varies widely. While the average size of a diploma course at all schools or at conventional schools is 39 to 46 students, the standard deviation of the average size of a diploma class is larger than the mean. There are schools with very large cohorts of diploma students.[28] In contrast, private schools have very small diploma class sizes and less variation in average cohort size compared to non-private schools.

Table 4 presents estimates of equation (2) for each of the school groupings. There are two parts to the table, the lower part restricts the samples by excluding the schools with less than 10 courses. All estimates describe similar relations between school characteristics and the estimated school fixed effects. When all schools are included, some schools are non-conventional as defined earlier. The coefficient on the non-conventional school indicator variable is positive and significant with a value of about two. At the same examination grades, the teachers award average grades two percentage points higher on all diploma courses at a non-conventional school than would be awarded at a conventional school. This effect can be seen clearly in the upper right panel of Figure 1 where the open bars count non-conventional schools. The mass of bars representing non-conventional schools clearly fall to the right of zero, schools with over-grading school cultures.

Average school awarded grades at private schools are 2 to 3 percentage points higher across all courses than at non-private schools when students have the same examination results. The smaller coefficient is found in the lower panel of Table 4 where the sample is restricted to include schools with 10 or more courses. The private school effect is readily observed in lower

---

[28] There are several very large non-conventional schools operated by public school boards in Calgary and Edmonton.

left panel of Figure 1 where private schools fixed effects are represented by the open bars. It is very clear that most private schools have large positive school fixed effects, that is, systematically award higher grades at the same examination grades. Private schools have low grading standards.

Finally, average school awarded grades at schools outside Alberta's two urban areas with more than one million in population, where the rural indicator takes a value of one, are about one percentage point higher than grades at the urban schools when two schools have the same grades on the externally graded examinations. This effect is illustrated in the lower right panel of Figure 1 where the mass of rural schools, the open bars, lie to the right of zero.

The three remaining estimated coefficients in Table 4 are the coefficients on continuous variables describing school characteristics. The first continuous variable is the average size of all diploma courses offered at the school over the years in each group.[29] If the average size of a diploma course offered at a school is larger, then the school fixed effects is predicted to be smaller. If the coefficient on average course size is - 0.006, then reducing the average cohort size at a school by 50 students (roughly one standard deviation in Table 3, predicts an increase in the school fixed effect of 0.3 percentage points. Schools with smaller diploma course cohorts are predicted to have lower grading standards as part of their grading culture. Teachers at schools where diploma cohorts are smaller give higher grades across all courses at the same examination grades.

[29] This is calculated over all the diploma courses offered at each school. It is partly a proxy for school size. However, if average total enrolment at the school is used in equation (2) in place of average enrolment in diploma courses, its estimated coefficient remains negative but is less frequently statistically significant. If both an enrolment variable and a diploma course variable are used, neither is significant. There is collinearity between the two variables.

The coefficient on the higher education measure is roughly 0.05. Using the 0.05 value as an example, if the percentage of parents with a university degree at one school is larger than at another school by 10 percentage points, roughly one standard deviation of the variation in the percent of adults with completed university degrees in the school population average, the predicted value of the estimated school fixed effect increases average school awarded grades by 0.5 percentage points. Schools with a larger percentage of well-educated parents are predicted to have lower grading standards.

The last continuous variable measures, in percentage points, the deviation of the school's average percentage of lone parent households from the population average. The coefficient on this variable is negative, a larger percentage of lone parent households predicts a reduction in school-awarded grades. The coefficient is larger when non-conventional schools are excluded from the sample and even larger when the schools with 9 or less diploma courses are excluded from the sample and even larger when only university courses are in the sample. The last entry in the column of coefficients is -0.13. Using a value of -0.10 as the example, if the percentage of lone parent households rises by 10 percentage points, the average school fixed effect is predicted to be roughly 1.0 percentage point lower. Schools with a higher percentage of lone parent households are predicted to have higher grading standards. The results of estimating equation (2) show that a reasonable amount of the variation in school fixed effects is associated with the five or six observable characteristics of schools used as right-hand side variables in equation (2); the R^2 values range from 0.23 to 0.31. The larger values are in the estimates that exclude the non-conventional schools. This makes sense since the census variables are likely more accurate representations of school characteristics for conventional schools where more students are part of the parental home.

4. Conclusions

This paper shows grading cultures exist in high schools in Alberta, Canada. Schools can be identified as over-grading (low standards) schools, under-grading (high standards) schools and schools that neither over grade nor under grade. In an over (under) grading school, the average school assigned grade in all courses is higher (lower) than in another school with the same average examination grades in those courses. This is an important result. Because student final grades are a weighted average of school assigned grades and examination grades, the school you attend affects your final grade in two ways. First, your school may raise your level of skills and thus both your examination grade and school-assigned grade that depend positively on the level of those skills are higher. Second, your school may be an over (or under) grading school and raise (lower) your school-assigned grade at the same examination grade and the same level of skills. The systematic variation in school-awarded grades at the same examination grades is large enough that it could affect both admission to high demand university programs as well as graduation from high school. Attending a school where school-assigned average grades across all courses are three percentage points higher at the same examination grades than another school would be a common occurrence. Then the average final grade across all courses at the high grading school, in the period where the school-awarded grade was 50% of the final grade would be 1.5 percentage points higher, a magnitude large enough to affect post-secondary admissions and high school graduation.[30] Individual student data would be needed to understand how school

[30] In the final four years studied, the weight on the examination grade fell to 30% and the weight on the school-awarded grade increased to 70%. School over and under grading increases in importance in the final four years.

over and under grading is distributed across the grade distribution of students. It is a limitation of this paper and the Alberta data to analyze only averages of grades by school, course, and year.

The characteristics of a school partially predict whether a school is an over-grading or an under-grading school, that is, has low or high grading standards. Non-conventional schools, schools addressing the needs of students who previously dropped out and then return, are more likely to over grade. Private schools, where parents pay fees, are more likely to over grade. Rural schools are more likely to over grade. Schools where the average cohort of all diploma courses is smaller are more likely to over grade. Schools with a larger percentage of parents with a university degree are more likely to over grade. And finally, schools with a large percentage of lone parent households are more likely to under grade. The finding that schools with a higher percentage of parents with a completed university degree and a lower percentage of lone parent households over grade by quite large amounts suggests that students who already come to high school with some advantages in life are more likely to attend schools that over grade, a possible equity issue.

The measurement of school over and under grading in a school grading culture in this paper uses a crucial maintained assumption: that the curriculum-based provincial examination and the school-assigned grade in each course measure a common set of skills with different course-specific weights. One advantage of the high school environment in Alberta is that teachers are supposed to be teaching and assessing the provincial curriculum in the 10 courses used in this paper to measure school grading cultures. The provincial examination is based on that curriculum. Teachers know the content of the examination and of the curriculum. The common skills assumption is more plausible in this setting than, for example, when the comparator for measuring school over-grading or under-grading is an aptitude test or a more general standardized test in literacy or numeracy.

It is possible that school awarded grades measure significantly different skills than examination assigned grades and these different skills are similar in different courses. Higher school awarded grades at the same examination grades could thus be systematically different across schools and reward non-examination skills that are systematically different across schools. It may or may not be simple to think of such skills. Suppose the skill in mind is punctuality. This may sound like a skill with no direct weight on an external examination grade. But if a student who is often late missed the initial content of many classes, the examination grade falls, and the lack of punctuality would affect the examination grade. It may be quite difficult to think of skills that raise only the school-awarded grade and have zero weight on the examination grade. If you believe important non-examination skills exist and are systematically different across schools, then the results suggest that the additional skills measured by the school-assigned grade and not measured by the examination grade are systematically higher in private schools, at non-conventional schools, at rural schools, at schools with a smaller percentage of lone parent households, at schools with smaller diploma class sizes and at schools with a larger percentage of parents with completed university degrees. The systematic differences in school assigned grades associated with those school characteristics are then legitimate, they reward higher non-examination skills. Whether this is a plausible hypothesis is left up to the reader.

The alternative hypotheses interpret relations between school-level characteristics and school fixed effects, the over and under grading measures, in non-skill-based ways. I suggest the following examples: perhaps when the class is smaller, teachers find it harder to given lower grades to students they know better; perhaps in rural settings, where teachers know the families directly, it is harder for teachers to award lower grades to their students; perhaps better educated parents and their children are more aware of the importance of high grades and lobby their teachers more effectively to raise school-awarded grades than less educated parents; perhaps teachers, as very well educated upper middle class persons, give higher grades to students more like themselves[31]; perhaps single parents have less time to lobby teachers for higher grades; perhaps parents are simply paying for higher school-assigned grades at private schools and have more bargaining power in discussions of school-assigned grades with private school teachers, principals and boards; perhaps the non-conventional schools focus on getting returning students to simply get their high school diploma and thus over grade to facilitate that process. These potential explanations are consistent with the school grading cultures observed.

Whether there are additional skills systematically linked to higher school-awarded grades is an important research question. There is potential to ask that question with individual observations linking the school assigned grades and the examination grades to other measures of life success. Better life success might then be interpreted as higher skill acquisition. Possible measures of life success could be post-secondary grades, post-secondary graduation, employment, income, or periods on social assistance. Such a project might help resolve whether the systematic differences between school-awarded grades and examination grades presented in this paper reward different skills or result from schools developing and maintaining a grading culture with higher or lower grading standards.

[31] Teachers in Alberta earned at the 80th percentile of all employment earnings in Alberta in 2013/14, see Johnson (2015). Virtually all teachers have a completed university degree in Alberta.

References

Abowd, John M., Francis Kramarz and David N. Margolis (1999) "High Wage Workers and High Wage Firms." Econometrica 67(2): 251-333.

Allensworth, Elaine M. and Kallie Clark (2020) "High School GPAs and ACT Scores as Predictors of College Completion: Examining Assumptions About Consistency Across High Schools" Education Research 49 (3): 198-211.

Betts, Julian R. (1998) "The Impact of Educational Standards on the Level and Distribution of Earnings" American Economic Review 88 (1): 266-275

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.

Botelho, Fernando, Ricardo A. Madeira and Marcos A. Rangel (2015) "Racial Discrimination in Grading: Evidence from Brazil." American Economic Journal: Applied Economics 7(4):37-52

Burgess, Simon and Ellen Greaves (2013) "Test Scores, Subjective Assessment and Stereotyping of Ethnic Minorities." Journal of Labor Economics 31(3): 535-576.

Dhuey, Elizabeth and Justin Smith (2019) "How School Principals Influence Student Learning?" Empirical Economics 54(2):851-882.

David N. Figlio and Maurice E. Lucas (2004) "Do High Grading Standards Affect Student Performance?" Journal of Public Economics 88,no. 9–10:1815–34.

Finnie, Ross and Stephen Childs, Dejan Pavlic, Nemanjic Jevtovic (2019) "How Much do University Graduates Earn?" EPRI #Grad Earnings RESEARCH BRIEF #3 (VERSIONB ,14-11-21)
https://www.eqbank.ca/?utm_source=google&utm_medium=cpc&utm_campaign=*EQ_Brand&utm_term=Main&utm_content=ABetterWayToBank&gclid=EAIaIQobChMI5bbjzoOc5wIVCbSzCh1WiAPZEAAYASAAEgLotfD_BwE

Seth Gershenson (2018) "Grade Inflation in High Schools (2005–2016)" Washington, D.C.: Thomas B. Fordham Institute.

Gershenson, Seth (2020) "Great Expectations: The Impact of Rigorous Grading Practices on Student Achievement" Washington, DC: Thomas B. Fordham Institute (February 2020). https://fordhaminstitute.org/national/research/great-expectations-impact-rigorous-grading-practices-student-achievement

Gibbons, Stephen and Arnaud Chevalier (2008) "Assessment and age 16+ education participation." Research Papers in Education 23 (2):113-123.

Godfrey, Kelly (2011) "Investigating Grade Inflation and Non-Equivalence." College Board Research Report 2011-2.

Himmler, Oliver and Robert Schwager (2012) "Double Standards in Educational Standards – Do Schools with a Disadvantaged Student Body Grade More Leniently?" German Economic Review 14(2):166-189.

Hurwitz, Michael and Jason Lee (2017), "Grade Inflation and the Role of Standardized Testing," in Measuring Success: Testing, Grades, and the Future of College Admissions, ed. Jack Buckley, Lynn Letukas, and Ben Wildavsky, Baltimore, MD: Johns Hopkins University Press, 64–93.

Johnson, David R. (2005) "Signposts of Success: interpreting Ontario's Elementary School Test Scores." Toronto: C.D. Howe Institute.

Johnson, David R. (2015) "Value for Money? Teacher Compensation and Student Outcomes in Canada's Six largest Provinces" Toronto: C.D. Howe Institute Commentary No. 434. https://www.cdhowe.org/sites/default/files/attachments/research_papers/mixed/Commentary_434_0.pdf

Lavy, Victor (2018) "On the origins of gender gaps in human capital: Short- and long-term consequences of teachers' biases." Journal of Public Economics 167: 263-279.

Marcenaro-Gutierrez, Oscar and Anna Vignoles (2014) "A comparison of teacher and test-based assessment for Spanish primary and secondary students." Education Research 57(1): 1-21.

Murphy, Richard and Gill Wyness (2020) "Minority Report: the impact of predicted grades on university admissions of disadvantaged groups." Education Economics 28(4):333-350.

Protivinsky, Tomas and Daniel Munich (2018) "Gender bias in teachers' grading: What is in the grade?" Studies in Educational Evaluation 59: 141-149.

Rangvid, Beatrice Schindler (2015) "Systematic Differences across evaluation schemes and educational choice." Economics of Education Review 48: 41-55

Rangvid, Beatrice Schindler (2019) "Gender Discrimination in Exam Grading" Double Evidence from a Natural Experiment and a Field Experiment." B.E. Journal of Economic Analysis and Policy

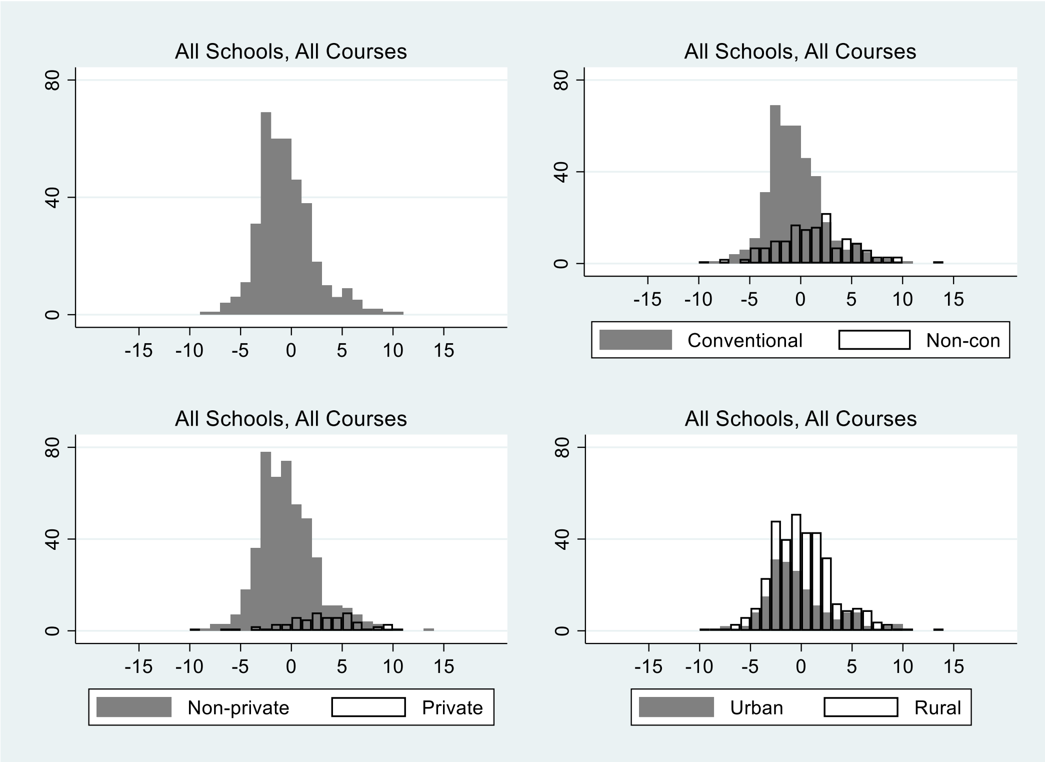
Rauschenberg, S. (2014) "How consistent are course grades? An examination of differential grading." Education Policy Analysis Archives, *22*(92) http://dx.doi.org/10.14507/epaa.v22n92.2014

Solon, Gary, Stephen J. Haider and Jeffrey M. Wooldridge (2015) "What are we weighting for?" Journal of Human Resources 50(2): 301-316.

Tyner, Adam and Seth Gershenson (2020) "Conceptualizing Grade Inflation" Economics of Education Review 78: 102037.

Tyner, Adam and Matthew Larsen (2017) End-of-Course Exams and Student Outcomes. Washington, DC: Thomas B. Fordham Institute (August 27, 2019). ttps://fordhaminstitute.org/national/research/end-course-exams-and-student-outcomes.

Figure 1: Histograms of school fixed effects: changes in the average grade across all courses offered at all schools



Notes: Vertical axes measure the number of schools. Horizontal axes measure the school fixed effects estimated in the paper. A school fixed effect is the change in the average grade in all courses at a school controlling for the examination grade in the course. Conventional schools offer all three grades; non-conventional schools (non-con) have mostly Grade 12 students. Rural schools are outside the Calgary and Edmonton Census Metropolitan Areas. Private schools receive fees per student.

Table 1: School Assigned and Examination assigned grades 2005-06 to 2018-2019

Type of Schools	Number of Schools	Number of Diploma Courses	Total Enrolment in Diploma Courses	Examination Average Grades[1]	School Assigned Average Grade[1]	Number of schools that offer 10 or more courses
All Schools	533	40,934	2,389,923	65.8	72.0	461
Conventional Schools	380	35,734	2,116,788	66.4	72.2	357
Conventional Private Schools	48	2,469	51,333	70.7	78.1	39

Source: Author's calculations. 1. These are averages of grades in courses weighted by course enrolment.

Table 2: Estimates of an AKM model of the relation between school-assigned grades and examination grades

Type of schools Courses[1]	Slope parameters measuring the relationship between the school assigned average grade and the examination assigned average grade (standard error)										Fixed effects without number of courses restricted		Fixed effects from schools with more than ten examinations	
	English 30-1	Math 30-1	Social Studies 30-1	Phy-sics	Chem -istry	Biol-ogy	English 30-2	Math 30-2	Social Studies 30-2	Science 30	Standard deviation (number of schools)	% with t-stats > |2.0|	Standard deviation (number of schools)	% with t-stats > |2.0|
All, All	.29 (.01)	0.26 (.006)	.28 (.008)	.31 (.08)	.29 (.007)	.36 (.007)	.27 (.008)	.27 (.008)	.33 (.009)	.39 (.01)	3.2 (533)	66.4	2.69 (461)	71.8
Con , All	.24 (.01)	.25 (.006)	.25 (.009)	.30 (.008)	.29 (.007)	.38 (.008)	.24 (.01)	.28 (.009)	.34 (.01)	.39 (.01)	2.7 (381)	69.8	2.52 (358)	72.3
Con , Univ	.25 (.01)	.25 (.006)	.27 (.009)	.31 (.007)	.30 (.007)	.39 (.008)	NA	NA	NA	NA	2.9 (266)	70.3	2.61 (231)	65.8

1. The first "All" indicates all schools are included; Con indicates only conventional schools are included. The second "All" specifies that all non-French courses are included. "Univ" specifies that the 6 university-oriented courses are included. Non-university courses are the courses labelled NA.

Table 3: Characteristics of Schools

Characteristic	Unit of observation	All Schools Mean (S.D.) ((number of units))	Conventional Schools Mean (S.D.) ((number of units))	Private Conventional Schools Mean (S.D.)((number of units))
Percent of adults associated with school with a completed university degree 2005-2006 to 2014-2015 Measured using 2006 census	Schools	13.01 (9.27) ((493))	13.38 (10.00) ((365))	21.32 (13.08) ((44))
	Weighted by enrolment	16.35 (9.75) ((225,431)	16.44 (10.20) ((192,904)	22.87 (13.62) ((11,987))
Percent of adults associated with school with a completed university degree 2015-2016 to 2018-2019 Measured using 2016 census	Schools	16.78 (10.30) ((486)	17.19 (11.23) ((355))	28.32 (14.90) ((39))
	Weighted by enrolment	21.01 (10.7) ((228,160)	21.19 (11.43) ((184,291)	31.13 (15.30) ((12,157))
Percent of households that are lone parent 2005-2006 to 2014-2015 Measured using 2006 census	Schools	23.32 (9.21) (493)	22.62 (8.74) ((365))	22.11 (8.92) ((44))
	Weighted by enrolment	24.41 (5.77) ((228,160)	24.35 (7.39) ((192,904)	22.61 (7.66) ((11,987))
Percent of households that are lone parent 2015-2016 to 2018-2019 Measured using 2016 census	Schools	25.57 (6.83) ((486))	24.78 (6.92) ((355)	22.11 (8.92) ((44))
	Weighted by enrolment	24.47 (5.77) ((192.904))	24.25 (6.02) ((184,291))	19.07 (6.28) ((12,157))
Percent of schools that are rural	Schools	66.0	66.6	40.8
Size of diploma course	Schools	38.80 (57.21)	46.03 (58.02)	15.77 (13.74)

Source: Author’s calculations

Table 4: Relationships between School Over and Under Grading and School Characteristics

Grouping Schools, Courses	Number of Schools	Non-conventional school indicator	Private school indicator	Rural indicator	Average course size at school	Lone parent measure	Higher education measure	R^2
All, all	533	2.02 (.35)*	2.86 (.58)*	1.16 (.31)*	-0.007 (.002)*	-0.048 (.025)	0.054 (.019)*	0.23
Con, all	381	NA	2.72 (.53)*	0.98 (.31)*	-0.006 (.002)*	-0.095 (.02)*	0.052 (.017)*	0.28
Con, Univ	266	NA	2.78 (.61)*	1.01 (.38)*	-0.008 (.002)*	-0.089 (.034)*	0.045 (.020)*	0.28
	Estimates from samples restricted to schools with more than 10 courses taught at the school							
All, all	461	1.77 (.36)*	2.06 (.68)*	1.02 (.27)*	-0.006 (.002)*	-.083 (.022)*	0.063 (.019)*	0.25
Con, all	357	NA	2.21 (.49)*	0.85 (.27)*	-0.005 (.002)*	-0.117 (.016)*	0.059 (.015)*	0.31
Con, Univ	230	NA	2.05 (.54)*	0.88 (.35)*	-0.006 (.002)*	-0.130 (.024)*	0.048 (.019)*	0.31

Robust standard errors in parentheses. * indicates coefficient is statistically different from zero at 5%

Appendix A Table 1: High school grades at entry in three Alberta universities

Fall of Year	Average Entering Grade "University of			Percent of entrants with average grades less than 80 percent			Percent of Entrants with average grades less than 70%		
	Alberta	Calgary	Lethbridge	Alberta	Calgary	Lethbridge	Alberta	Calgary	Lethbridge
2019	88.7	86.7	79.9	4.3	11.0	50.8	0.0	0.0	9.1
2018	88.3	86.7	80.1	7.5	10.6	50.2	0.0	0.3	9.3
2017	88.8	86.0	79.0	4.7	13.4	50.5	0.0	0.3	8.7
2016	87.4	85.7	81.2	10.0	15.3	40.6	0.0	0.1	5.2
2015	87.1	85.8	81.5	10.7	14.2	40.5	0.0	0.2	4.9
2014	87.3	86.5	80.7	8.7	10.7	42.4	0.0	0.1	5.7
2013	86.5	86.2	80.7	10.3	12.6	53.4	0.0	0.0	6.9
2012	86.6	84.1	80.7	12.7	22.8	44.5	0.2	0.1	5.2
2011	86.2	82.7	80.7	14.2	32.1	45.0	0.1	0.4	5.6
2010	85.5	82.2	80.2	18.1	37.5	48.2	0.1	0.3	6.5
2009	85.7	82.2	80.2	15.1	36.1	49.2	0.1	0.5	6.5
2008	86.0	82.9	80.1	16.4	33.2	47.4	0.2	0.1	7.0
2007	86.1	83.1	79.6	15.3	29.9	53.0	0.4	0.0	9.1
2006	NA	NA	78.0	NA	NA	50.0	NA	NA	6.2
2005	86.2	84.0	79.7	NA	NA	NA	NA	NA	NA

Source: Macleans magazine annual guide to Canadian universities reports entering grades from the previous fall. The average entering grade and lowest entry grades are affected by the program mix within each university. In the missing years, universities did not participate in the survey.

Notes on specific programs entry grades in 2021.

The University of Alberta (January 2022) website reports admission average grade ranges: Arts, mid-70s to low 80s; Education, mid 70s to high 80s; Engineering, Low to high 80s; Kinesiology, High 80s to low 90s; Nursing, High 80s to low 90s; Science, High 80s to low 90s. https://www.ualberta.ca/admissions/undergraduate/admission/admission-requirements/competitive-requirements.html. The University of Calgary website reports admission grade ranges: Business, Mid 80s; Civil Engineering, low 90s; Computer Science, mid 80s; Economics, mid 80s; History, high 70s; Kinesiology, high 80s; Nursing, low 90s. https://www.ucalgary.ca/future-students/undergraduate/requirements. The University of Lethbridge actual admission averages in fall 2021 (provided from Lethbridge internal sources) by program: Bachelor of Arts, 80; Bachelor of Science, 85; Bachelor of Management, 81; Transfer Program, Engineering, 85; pre-Bachelor of Education/Science 85; pre-Bachelor of Education, Arts 83. The last three programs provide subsequent access to engineering or education. The last two universities MacEwan University (opened 2009) and Mount Royal University (opened 2011) both were conversions from non-university to university post-secondary institutions. They have much lower average entry grades and do not participate in the Macleans' survey. Mount Royal https://www.mtroyal.ca/Admission/_pdfs/ssdata_admission_requirements.pdf reports required averages on the Diploma courses of 85-90% in Education, 90-100% in midwifery and 95-100% in nursing. MacEwan University https://www.macewan.ca/academics/programs/bachelor-of-science-in-nursing/admissions/requirements/ has similar admission averages in nursing but seems to have lower requirements in other programs.

Appendix B: Average Diploma Course Grades 2005-06 to 2018-19

Year	Average Examination Grade	Average School Assigned Grade
2005-06	66.7	70.6
2006-07	66.9	70.7
2007-08	66.6	70.8
2008-09	65.5	70.9
2009-10	65.0	71.1
2010-11	64.6	71.2
2011-12	65.1	71.6
2012-13	65.5	71.7
2013-14	65.7	72.2
2014-15	65.8	72.6
2015-16	65.4	73.2
2016-17	65.8	73.5
2017-18	66.5	73.8
2018-19	66.6	73.9

Source: Calculations by author. Averages are across students from school-level data weighted by number of enrolments. Excluding grades in French courses.

Appendix C: A derivation of the equation relating school average grades and examination average grades

1 Skills

Skills acquired by student i are divided into two mutually exclusive groups measured by indexes K_{ei} and K_{oi} (examination skills and other skills respectively). Skills K_{ei} improve grades on both the externally graded examination and on the grades awarded by teachers at schools, school awarded grades. Skills denoted K_{oi}, “other” skills, increase only school awarded grades. An example of this type of skill could be the ability to make an oral presentation with a loud clear voice, a skill that might or might not be useful in raising school awarded grades but is likely not useful in raising examination grades. The derivation in this appendix shows that a critical assumption in the literature and in this paper is that all skills with positive weights in school-awarded grades also have some positive weight in examination grades, a common skills assumption. If this assumption is made, then it is possible to derive equation (1) in the paper, a regression of the average school awarded grade on the average examination awarded grade with school and course fixed effects where the estimated school fixed effects are estimates of over and under grading across all courses offered at the school.

The level of social and economic status (SES) of student i’s household is measured by index variable(s) Z_i. The skill levels of the i^{th} student attending the kth school at the time of assessment in a diploma course are

$$K_{eki} = \gamma_0 + \gamma_1 Z_i + Q_{ek} + \varepsilon_{1ki} \quad (1)$$

and

$$K_{oki} = \alpha_0 + \alpha_1 Z_i + Q_{ok} + \varepsilon_{2ki} \quad (2)$$

Skills are not course specific but, as will be made clear below, a given skill is not equally weighted in constructing the school awarded or examination grade in different courses.[32] Student skills are partly acquired from their home environment and increase in the SES variable Z_i.[33] Q_{ek} and Q_{ok} represent exam skills and other skills respectively as accumulated by a student enrolled at school k from their past sequence of schools including school k (and including any effect of possible peers at the past sequence of schools they may have attended) up to time t when we observe a school awarded grade and an examination grade. Q_{ek} and Q_{ok} take the same value for all students at school k and are time invariant.[34]

There are two mean zero random components of skills across all students at all schools that are uncorrelated with Z_i denoted ε_{1i} or ε_{2i}. The average value of ε_{1ki} or ε_{2ki} across students' i in course j in a school k need not be zero. In the most obvious case, students with a high value of ε_{1i} or ε_{2i} (there is no school subscript here) may find themselves directed into specific schools or specific courses. Parents of a high ε_{1i} or ε_{2i} student may seek out a specific school, perhaps even schools with higher values of Q_{ek} or Q_{ok}. A positive or high epsilon student has a positive value of ε_{1i} (or ε_{2i}) and a higher level of skills conditional on his/her value of Z_i than the skill level predicted for the average student at each level of Z regardless of school of enrolment. The school quality terms Q_{ok} or Q_{ek} then add (or reduce) a student's skills from the student's skill

[32] Literacy could have a higher weight in an English course and mathematical skills a higher weight in physics but both courses would likely require some literacy and some mathematics.
[33] Z_i could be, for example, total years of education of both parents or the number of years of education of the most educated parent.
[34] Forcing Q_{ok} and Q_{ek} to be identical for all students at a school may seem like a strong assumption. Since the empirical work uses average grades by school and course, if at this stage there was some random variation within students at the school in these skills acquired from schools, it would be averaged out later. One source of variation could be that individual students enrolled at school k at time t had experienced a different sequence of schools in the period before time t. The notation is already sufficiently dense.

level conditional on the student's Z_i and that student's realization of ε_{1i} or $\varepsilon_{2i.}$ A student can have a high level of skill because they come from a high SES background; because they come from a "good" school with a high Q_{ok} or Q_{ek}; or because they have a large value of ε_{1i} or ε_{2i}.[35]

2 Grades

$E_{ijk(t)}$ is the examination grade in course j at time t for a student enrolled at school k. Skills at school k are transformed into examination grades $E_{ijk(t)}$ using

$$E_{ijk(t)} = \delta_{0j} + \delta_{1j} K_{ei} + D_{j(t)} + \varepsilon_{3ijk(t)} \qquad (3)$$

δ_{0j} and δ_{1j} are course specific parameters. $\varepsilon_{3ijk(t)}$ is the random component in (3). Any random component that hits all students in course j at all schools in the entire province in time t is absorbed into the parameter $D_{j(t)}$. There could also be a school-specific shock to all courses at school k in time t, a forest fire in school k's area. There could be a specific shock to a course that affected all students in that specific course at that school - a student vomiting in the exam room in that course but not in other courses at that school. There could be a student specific random component, student "i" just has a bad or good day. If (3) were to be estimated using individual observations across courses and years, the error structure implies clusters at the school-course-year as well as the school-year level. There are no direct measures of individual

[35] This broad set for the sources of student ability is important to describe variation in average school results on external examinations. Some schools where children come from disadvantaged backgrounds outperform schools where children come from advantaged backgrounds indicating high school quality when schools are not selecting on ability. Schools have a substantial range of examination results across schools attended by students of the same average SES background. Johnson (2005) documents these effects in elementary schools in Ontario. In some cases, some schools have entrance policies that select on high overall ability, some combination of SES and high epsilon values. These schools, in a very uninteresting way, invariably have very strong examination results conditional on the SES background of their students. The Calgary charter schools fall into this category in Alberta.

skills to enable (3) to be estimated and, in the Alberta data, no individual observations. Instead there are only data on school average grades.

The term $D_{j(t)}$ in the grading equation is controversial. This formulation clarifies that a change in the value of $D_{j(t)}$ is a change in the difficulty of the examination; the same level of skill translates to a different grade.[36] Provincial examinations are carefully constructed with the intention of being equally difficult across years and, if that goal were achieved, then $D_{j(t)}$ equals zero in all courses in all years. Then all variation in the average examination outcome across different years is variation in average examination-related skills of the participating students.

Both skills for the same student in the same course are transformed into school awarded grades $S_{ijk(t)}$ at school k using

$$S_{ijk(t)} = \beta_{0j} + \beta_{1j} K_{ei} + \beta_{2j} K_{oi} + P_{jk(t)} + \varepsilon_{4ijk(t)} \qquad (4)$$

$\varepsilon_{4ijk(t)}$ is a random component similar to $\varepsilon_{3ijk(t)}$ clustered with individual observations at both the school-year level and the school-course-year level. $P_{jk(t)}$, when positive, indicates that teachers at school k in course j award higher grades to all students in course j than are consistent with the actual skill level of their students in that course. For convenience, $P_{jk(t)}$ is called the over-grading term, if negative teachers at school k in course j are tough graders, if positive, the school has

[36] The economics literature typically adds a time fixed effect and or standardizes individual grades to capture or remove possible variation in the difficulty of the examination over time. If such an adjustment is not made, there is an implicit or explicit assumption that assigned grades across different years are comparable. This notation makes the adjustment for variation in difficulty explicit so that the grades remain in the reported scale from 0 to 100. One advantage of that choice is that admissions to post-secondary are usually set in terms of cut off grades measured from 0 to 100 and for the measurement of school fixed effects on grades in this paper, this is a natural scaling.

generous graders in that course. $P_{jk(t)}$ is zero when there is neither over nor under grading in course j at school k.[37]

This model of grades states that within a school, at the same level of either ε_{1i} or ε_{2i}, higher SES students have higher grades within a school. Although it is likely schools with a larger proportion of high SES students will have both higher average school awarded grades and higher average examination grades than schools with a lower proportion of high SES students, this may not happen if a school is either a higher quality school (a high value of Q_{ek} or Q_{ok}) or if the school has a large proportion of high epsilon students.[38]

$D_{j(t)}$ measures variation in examination difficulty over time in the simplest case. $P_{jk(t)}$ measures over grading or under grading by school and course.[39] The goal is to a measure that is the average value of $P_{jk(t)}$, averages across courses within a school at time t. This is possible by creating a relationship between school-awarded grades and examination grades.

3 Expressing the school awarded grade as a function of the examination grade

Equations (1), (2), (3) and (4) are manipulated to parameterize a relation between the school awarded grade and the examination grade, first in the most general case and then in two special cases.

---

[37] Equations like (5) are typically not estimated since skills are not directly observed. One exception is Cornwall, Mustard and Van Parys (2013) who add measures of specific non-academic skills to successfully explain gaps between external and teacher grades in kindergarten.

[38] League tables of school average examination results are often presented without any such nuance. In many but not all cases, schools with high average examination results are schools where more students come from stronger SES backgrounds. It is the schools with strong results and large proportion of students from lower SES backgrounds that are the most interesting schools. That is the focus of Johnson (2005) and subsequent studies of the same type.

[39] Bonesronning (2008) estimates a parameter similar to $P_{jk(t)}$ as the average difference between internal grades and external grades within a classroom of students.

Combining expression (1) and expression (3) as

$$Z_{ijk(t)} = \frac{E_{ijk(t)}}{\delta_{1j}\gamma_1} - \frac{\delta_{0j}}{\delta_{1j}\gamma_1} - \frac{\gamma_0}{\gamma_1} - \frac{1}{\gamma_1} Q_{ek} - \frac{1}{\delta_{1j}\gamma_1} D_{j(t)} - \frac{1}{\gamma_1} \varepsilon_{1ijk(t)} - \frac{1}{\delta_{1j}\gamma_1} \varepsilon_{3ijk(t)} \quad (5)$$

shows that a student's examination grade in course j at school k; adjusted for the quality of the school k attended by student i; the difficulty of the examination; and the value of $\varepsilon_{1i;}$ can be written as a measure of the unobserved SES level of student i enrolled in course j at school k at time t.[40]

Substitution of equations (1) and (2) into (4) creates

$$S_{ijk(t)} = \beta_{0j} + \beta_{1j} (\gamma_0 + \gamma_1 Z_{ijk(t)} + Q_{ek} + \varepsilon_{1ijk(t)}) + \beta_{2j} (\alpha_0 + \alpha_1 Z_{ijk(t)} + Q_{ok} + \varepsilon_{2ijk(t)})$$

$$+ P_{jk(t)} + \varepsilon_{4ijk(t)} \quad (6)$$

and shows that school awarded grades increase in student skills which in turn increase with SES, the student's level of ε_1 and ε_2 and the quality of the student's school, Q_{ek} and Q_{ok}.

Substitution of (5) into (6) and simplifying yields

$$S_{ijk(t)} = \beta_{0j} + \beta_{2j}\alpha_0 - \frac{\beta_{2j}\delta_{0j}}{\delta_{1j}} - \frac{\beta_{1j}\delta_{0j}}{\delta_{1j}} - \frac{\beta_{2j}\gamma_0\alpha_1}{\gamma_1} - \left(\frac{\beta_{1j}\gamma_1 + \beta_{2j}\alpha_1}{\delta_{1j}\gamma_1}\right) D_{j(t)}$$

$$+ \left(\frac{\beta_{1j}\gamma_1 + \beta_{2j}\alpha_1}{\delta_{1j}\gamma_1}\right) \left[E_{ijk(t)} \right]$$

$$+ P_{jk(t)} - \left(\frac{\beta_{2j}\alpha_1}{\gamma_1}\right) \left[\varepsilon_{1ijk(t)} \right] + \beta_{2j}\, \varepsilon_{2ijk(t)} - \left(\frac{\beta_{2j}\alpha_1}{\delta\gamma_1}\right) Q_{ek} + \beta_{2j} Q_{ok} -$$

$$\left(\frac{\beta_{1j}\gamma_1 + \beta_{2j}\alpha_1}{\delta_{1j}\gamma_1}\right) \varepsilon_{3ijk(t)} + \varepsilon_{4ijk(t)} \qquad (7)$$

[40] Burgess and Greaves (2013) make a similar observation in their analysis of examination results in English schools.

Expression (7) is a useful positive relation between a student's school grade and a student's examination grade. Student i's school grade is predicted, up to the student level random component, from the student i's s examination grade adjusted for the difficulty of the examination, the quality of the school, school over grading and values of ε_1 and ε_2 of the student. If the examination was too easy, then $D_{j(t)}$ is positive and the prediction of the skill-consistent school grade from the examination grade in time t is adjusted downwards. Similarly, if the school attended by the student is very good at producing examination results, the prediction of the skill-consistent school grade is adjusted downward for large values of Q_{ek}. If the student is a high ε_1 student, the prediction of the skill-consistent school grade from the examination grade is adjusted downwards. These parameters affect exam specific skills. The skill-consistent school grade is adjusted upwards for a high ε_2 student or a high-quality school, a high value of Q_{ok}. These parameters affect skills that raise only school grades. Equation (7) is aggregated to a school-course average observation for estimation.

4. Aggregation

The average over grading parameters across all j courses offered at school k at time (t), $\widehat{P}_{k(t)}$, (notice j disappears in the subscript) are defined by:

$$P_{jk(t)} \equiv \widehat{P}_{k(t)} + \left(P_{jk(t)} - \widehat{P}_{k(t)} \right) \qquad (8)$$

A positive value of $\widehat{P}_k$ in time t describes a school with culture of over- grading, a negative value, an under-grading school. At the over-grading school students are awarded grades in all courses that are higher than the skill-consistent grade for that student.

In the most general case, (7) is aggregated to school-course-time period observations by averaging across student observations in course j at school k at time (t). Averages are denoted with a bar for course j at school k.[41] The relation between average school awarded grades and average examination grades by school and course is

$$\bar{\mathrm{S}}_{jk(t)} = \beta_{0j} + \beta_{2j}\alpha_0 - \frac{\beta_{2j}\delta_{0j}}{\delta_{1j}} - \frac{\beta_{1j}\delta_{0j}}{\delta_{1j}} - \frac{\beta_{2j}\gamma_0\alpha_1}{\gamma_1} - \left(\frac{\beta_{1j}\gamma_1 + \beta_{2j}\alpha_1}{\delta_{1j}\gamma_1}\right)\mathrm{D}_{j\,(t)}$$

$$+ \left(\frac{\beta_{1j}\gamma_1 + \beta_{2j}\alpha_1}{\delta_{1j}\gamma_1}\right)\left[\bar{\mathrm{E}}_{jk(t)}\right]$$

$$+ \hat{\mathrm{P}}_{k(t)} - \left(\frac{\beta_{2j}\alpha_1}{\delta\gamma_1}\right)\mathrm{Q}_{ek} + \beta_{2j}\mathrm{Q}_{ok}$$

$$- \left(\frac{\beta_{2j}\alpha_1}{\gamma_1}\right)\left[\bar{\varepsilon}_{1jk(t)}\right] + \beta_{2j}\,\bar{\varepsilon}_{2jk(t)}$$

$$+ \left(\mathrm{P}_{jk(t)} - \hat{\mathrm{P}}_{k(t)}\right) - \left(\frac{\beta_{1j}\gamma_1 + \beta_{2j}\alpha_1}{\delta_{1j}\gamma_1}\right)\bar{\varepsilon}_{3jk(t)} + \bar{\varepsilon}_{4jk(t)} \quad (9)$$

Equation (9) leads directly to the equation estimated in the paper. Variation in the average grade awarded at the school in a course at a point in time is modelled as a time fixed effect specific to each course in that time period; a linear function of the school's average grade on the examination in that course with a coefficient that varies by course; and an average school fixed effect across all courses at the school, as well as an error term.

One useful way to understand (9) is that $\bar{\mathrm{E}}_{jk(t)}$ is a control variable that captures the otherwise unobservable average SES value of students in the course and school as well as part of the selection of high $\bar{\varepsilon}_{1jk(t)}$ into the school and part of the school quality effect denoted Q_{ek}. The

[41] Larger schools would have more than one class group in a course group. Alberta reports average grades by school.

course-specific time fixed effect captures a series of uninteresting constants in the conversion of skills to grades across courses as well as variation over time in the difficulty of the examination.

The parameters of the most interest are the values of $\widehat{\mathrm{P}}_{k(t)}$, the measure of over grading at school k. In equation (9) that parameter is grouped with the schools' ability to increase student skills, both skills measured on examinations Q_{ek} and those not measured on examinations, Q_{ok} in the school fixed effect. There is no separation in the general case between a school that over (under) grades, a school that produces higher (lower) skills assessed through the school-awarded grade (Q_{ok}) and a school whose higher (lower) quality of skill production in the examination dimension (Q_{ek}) reduces (increases) the skill-consistent school awarded grade for a given examination grade.

The fourth line shows an additional complication in the use of the average examination grade to predict the average school awarded in the general case with the isolation of the school fixed effect as the measure of over or under-grading. Part of the error term, $\left(\frac{\beta_{2j}\alpha_1}{\gamma_1}\right)\left[\bar{\varepsilon}_{1jk(t)}\right] + \beta_{2j}\,\bar{\varepsilon}_{2jk(t)}$ is the average values of ε_1 and ε_2 of students enrolled in each course at that school in that time period. These components are certainly correlated with the school's average examination grade. This would bias the coefficient estimate on the average examination grade and more importantly bias the estimates of the school fixed effect coefficient. If parents of higher ability students, even after conditioning on social and economic characteristics, choose higher quality schools, then the school fixed effects will over-estimate the over grading parameter at that school. What could be interpreted as school over grading, a high value of $\widehat{\mathrm{P}}_{k(t)}$, is partly higher ability students selecting into the school. Teachers are not over grading; they are rewarding the higher skills. One route to an unbiased estimate of $\widehat{\mathrm{P}}_{k(t)}$ is to place further restrictions on the model.

The error term in (9) is clustered at the school-time period level across the courses in that year if common shocks hit all students at a school in all courses in a time period. The example earlier was a forest fire that closed an entire school temporarily for part of a year.[42]

5 The special case of the common skills assumption

If there is a common set of skills measured by school awarded grades and examination grades then β_{2j} is zero.[43] Equation (9) then rewrites as

$$\bar{\mathrm{S}}_{jk(t)} = \beta_{0j} - \frac{\beta_{1j}\delta_{0j}}{\delta_{1j}} - \left(\frac{\beta_{1j}}{\delta_{1j}}\right)\mathrm{D}_{j\,(t)}$$

$$+ \left(\frac{\beta_{1j}}{\delta_{1j}}\right)\left[\,\bar{\mathrm{E}}_{jk(t)}\right]$$

$$+ \widehat{\mathrm{P}}_{k(t)}$$

$$+ \left(\mathrm{P}_{jk(t)} - \widehat{\mathrm{P}}_{k(t)}\right) - \left(\frac{\beta_{1j}}{\delta_{1j}}\right)\bar{\varepsilon}_{3jk(t)} + \bar{\varepsilon}_{4jk(t)} \qquad (10)$$

Under the common skills assumption, the estimates of the variation time-period course-specific fixed effects vary with the difficulty of the examination within a course and across courses. The slope coefficients on the average examination score vary by course. The school fixed effects are

[42] A clear discussion of clustering is found in Cameron and Miller (2015).

[43] That type of language is mentioned in passing in some, but not all, of the literature reviewed in Section 2. The restriction does not mean, as emphasized below, that skills are rewarded identically by teacher grading and examination grading.

measures of school over and under grading across all courses. The error term is clustered at the school across courses within a time period. This is the equation is estimated in the text.

5. A very strong assumption in the literature

A very strong assumption is that identical skills are transformed into identical grades (up to a random component) by teachers and examination graders. This assumption requires that β_{2j}=0; $\beta_{0j}= \delta_{0j}$; and $\beta_{1j} = \delta_{1j}$. In that case, equation (9) or equation (10) reduce to

$$\bar{\mathrm{S}}_{jk(t)} = \bar{\mathrm{E}}_{jk(t)} - \mathrm{D}_{j\,(t)} + \hat{\mathrm{P}}_{k(t)} + \left(\mathrm{P}_{jk(t)} - \hat{\mathrm{P}}_{k(t)} \right) - \bar{\varepsilon}_{3jk(t)} + \bar{\varepsilon}_{4jk(t)} \quad (11)$$

which is more commonly written as

$$\bar{\mathrm{S}}_{jk(t)} - \bar{\mathrm{E}}_{jk(t)} = -\mathrm{D}_{j\,(t)} + \hat{\mathrm{P}}_{k(t)} + \left(\mathrm{P}_{jk(t)} - \hat{\mathrm{P}}_{k(t)} \right) - \bar{\varepsilon}_{3jk(t)} + \bar{\varepsilon}_{4jk(t)} \quad (12)$$

The simple difference between school awarded grades and examination grades is the dependent variable in Gibbons and Chevalier (2008), Falch and Napier (2013) and Berg, Palmgren and Trufores (2020). The restriction that the coefficient on the average examination grade is unity in moving from (10) to (11) is strongly rejected in the empirical results. The paper presents estimates of equation (10) rather than equation (11).

Additional References for Appendix C

Berg, Petter, Ola Palmgren and Bjorn Tyrefors (2020) “Gender grading bias in junior high school mathematics.” Applied Economics Letters 27 (11): 915-919.

Burgess, Simon and Ellen Greaves (2013) “Test Scores, Subjective Assessment and Stereotyping of Ethnic Minorities.” Journal of Labor Economics 31(3): 535-576.

Cameron, A. Colin and Douglas L. Miller (2015) “A Practitioner’s Guide to Cluster-Robust Inference.” Journal of Human Resources 50(2): 317-372.

Cornwell, Christopher, David B. Mustard and Jessica Van Parys (2013) “Non-cognitive Skills and the Gender Disparities in Test Scores and Teacher Assessments: Evidence from Primary School.” Journal of Human Resources 48 (1): 236-264.

Falch, Torberg and Linn Renee Naper (2013) “Educational evaluation schemes and gender gaps in student achievement.” Economics of Education Review 36: 12-25

Gibbons, Stephen and Arnaud Chevalier (2008) “Assessment and age 16+ education participation.” Research Papers in Education 23 (2):113-123.