CAHIER DE RECHERCHE ♯1002E
Département de science économique
Faculté des sciences sociales
Université d'Ottawa

WORKING PAPER ♯1002E
Department of Economics
Faculty of Social Sciences
University of Ottawa

Estimating the Benefit of High School for College-Bound Students*

Louis-Philippe Morin†

January 2010

* A previous version of this paper was entitled "Do College-Bound High School Students Need an Extra Year? Evidence from Ontario's `Double Cohort'". I would like to thank Victor Aguirregabiria, Joshua Angrist, Michael Baker, Pierre Brochu, Mark Evans, Xavier Fazio, Robert McMillan, Phil Oreopoulos, Silvana Simone, Jeffrey Smith, and especially Aloysius Siow, for helpful comments. I am grateful to George Altmeyer, Khuong Doan, Susan Pfeiffer, Pekka Sinervo, and the Faculty of Arts and Science at the University of Toronto for making this project possible. All remaining errors are mine. Disclaimer: The views, opinions, findings, and conclusions expressed in this paper are strictly those of the author. For confidentiality reasons, the data used in this paper cannot be released by the author. All requests about these data should be directed to the Faculty of Arts and Science at the University of Toronto.

† IZA and Department of Economics, University of Ottawa, 55 Laurier E., Ottawa, Ontario, Canada, K1N 6N5; Email: lmorin@uottawa.ca.

Abstract

Studies based on instrumental variable techniques suggest that the value of a high school education is large for potential dropouts, yet we know much less about the size of the benefit for students who will go on to post-secondary education. To help fill this gap, I measure the value-added of a year of high-school mathematics for university-bound students using a recent Ontario secondary school reform. The subject specificity of this reform makes it possible to identify the benefit of an extra year of mathematics despite the presence of self-selection: one can use subjects unaffected by the reform to control for potential ability differences between control and treatment groups. Further, the richness of the data allows me to generalize the standard difference-in-differences estimator, correcting for heterogeneity in ability measurement across subjects. The estimated value-added to an extra year of mathematics is small for these students – of the order of 17 percent of a standard deviation in university grades. This evidence helps to explain why the literature finds only modest effects of taking more mathematics in high school on wages, the small monetary gain being due to a lack of subject-specific human capital accumulation. Within- and between-sample comparisons also suggest that the extra year of mathematics benefits lower-ability students more than higher-ability students.

Key words: *Human Capital, High School Curriculum, Education Reform, Mathematics, Factor Model.*

JEL Classification: I20, I21, I28.

Résumé

Les études utilisant des variables instrumentales suggèrent que le bénéfice d'une année supplémentaire de scolarité secondaire est large pour des décrocheurs potentiels. Par contre, nous en savons beaucoup moins à propos du bénéfice d'une même année de scolarité pour les étudiants qui poursuivront des études postsecondaires. Afin de combler ce vide, j'utilise une récente réforme du système d'éducation secondaire ontarien permettant de mesurer la valeur ajoutée d'une année supplémentaire de mathématiques de niveau secondaire pour des étudiants qui iront à l'université. Puisque que la réforme n'affecta que certains sujets, il est possible d'estimer le bénéfice d'une année supplémentaire de mathématiques malgré la présence d'auto-sélection; nous pouvons utiliser les sujets qui n'ont pas été affectés par la réforme afin de contrôler les différences potentielles entre les étudiants du groupe contrôle et ceux du groupe traité. De plus, la richesse des données disponibles nous permet de généraliser l'estimateur des différences en différences, permettant ainsi aux différents sujets de mesurer l'habileté des étudiants de manière hétérogène. La valeur ajoutée estimée d'une année supplémentaire de mathématiques est petite pour ces étudiants – l'effet représentant environ 17 % de l'écart-type des notes universitaires observées. Ces résultats aident à expliquer pourquoi la littérature scientifique trouve que prendre plus de cours de mathématiques à l'école secondaire n'a qu'un impact modeste sur le salaire des individus : le peu de capital humain accumulé dans ces cours expliquerait cet effet modeste. Des comparaisons à l'intérieur et à l'extérieur de l'échantillon suggèrent également qu'une année supplémentaire de mathématiques serait plus bénéfique aux étudiants à habileté 'faible' qu'aux étudiants à habileté 'élevée'.

Mots clés: *Capital humain, curriculum d'études secondaires, réforme éducationnelle, mathématiques, modèle à facteur.*

Classification JEL: I20, I21, I28.

1 Introduction

Following Angrist and Krueger's landmark 1991 study, a number of economists have found surprisingly high rates of return to an additional year of secondary schooling using instrumental variables techniques. These high returns are likely to be driven by lower-ability students, given the instruments used primarily affect potential school leavers (dropouts).[1] In turn, evidence of large potential gains for these students has prompted a number of programs intended to reduce high-school dropout rates.[2]

Despite the understandable policy interest in dropouts, college-bound students now represent a majority of students in United States and Canada, following concerted efforts to expand rates of college attendance.[3] Yet we know much less about the benefits of secondary schooling for these higher-ability students[4] – in particular, whether the benefits from attending high school may be altogether lower for college-bound students.

The benefits of high school are very relevant to an important policy question concerning the way that a high school education should be delivered. How many years should it take, and related, what specific curricula should be taught?[5] This high school design issue was prominent in shaping a recent radical reform in Ontario, Canada's most populous province. Motivated by a desire to conform with a majority of North American secondary school curricula and by the prospect of lowering costs in the educational system, the Ontario government compressed its secondary school curriculum starting in 1999. Under the new system, students were expected to graduate from high school after four years (i.e. after Grade 12) instead of five (after Grade 13), suggesting the 'reverse' of the typical compulsory schooling law. As a consequence of the reform, in 2003 the first cohort of students graduating from Grade 12 and the last cohort of Grade-13 graduates entered Ontario colleges simultaneously – the so-called 'double cohort' – affording a unique and useful comparison that helps shed light on

[1] For example, Angrist and Krueger (1991), Harmon and Walker (1995), Staiger and Stock (1997), Meghir and Palme (2005), and Oreopoulos (2006) use either reforms (e.g. changes in compulsory schooling laws) or variables (e.g. quarter of birth) affecting the minimum legal number of years of schooling as an instrument.

[2] Examples include *Talent Development High Schools*, the *Quantum Opportunity Program,* and *Graduation Really Achieves Dreams*. See Dynarsky and Gleason (1998) and Dynarsky *et al.* (2008) for a discussion of other dropout intervention programs.

[3] In United States, the averaged freshman graduation rate for public secondary schools was 73.9 percent in 2002-2003, and 72.3 percent of the 2002-2003 graduates were attending college in 2003-2004. In Canada, the typical-age (18 year old) graduation rate was 67 percent in 2002-2003; 52 percent of 19 year olds were enrolled in college or university in 2003-2004. (Sources: Tables C5.2 and E1.1 in Canadian Education Statistics Council (2006), and Tables 102 and 193 in T.D. Snyder *et al.* (2008).)

[4] Measuring the benefit of an *extra year* of high school for college-bound students has proved difficult to answer as, ideally, one would need to find an instrument affecting the number of years of secondary schooling for students who will complete high school anyway. In this case, instruments such as quarter of birth or changes in minimum school-leaving age leave the number of years of schooling unchanged for college-bound students.

[5] In practice, education systems across the world exhibit wide variation along both dimensions. The grouping of grades into high school differs across countries. For instance, in the United States and most of Canada, secondary schools span grades 9 through 12, while in Europe secondary schooling can cover more than 6 years (as in France and Germany). There are also significant differences in terms of high school curricula, with students in some countries specializing early and encountering what would elsewhere be college-level material (the UK being an example), while high school remains more general in North America.

the benefits of high school for college-bound students.

A key feature of the reform was the way it changed the high school curriculum non-uniformly. Of note, the difference between the high school curricula of Grade 12 and Grade 13 students does *not* correspond to a year of schooling but to subject-specific years of education. That is, some subjects were drastically affected while others were not. For instance, the length of the high school mathematics curriculum for college-bound students went from five years to four while the length of the biology curriculum for the same students remained unchanged at two years.

This paper uses the non-uniform changes in curriculum to identify the value-added of Grade 13 mathematics:[6] one can use academic performance in subjects that were not affected by the reform to control for potential ability differences between Grade 13 and Grade 12 students due to self-selection. Controlling for self-selection is especially important in the context of this reform. Knowing that 2003 would be a more competitive year for college admissions, students sorted themselves, with some students delaying their university application by a year while others "fast-tracked" high school, graduating a year early to avoid the double cohort. As consequence, we cannot assume that Grade 12 and Grade 13 students have the same levels of ability.

To estimate the value-added of Grade 13 mathematics, I construct a flexible factor model that has three appealing features. First, it takes into account the possibility that Grade 12 and Grade 13 students might differ in academic ability (as measured by three academic performance indicators). Second, the model allows for the possibility that subjects do not measure ability in the same way, the identified value-added from the factor model being a generalization of the standard difference-in-differences estimator, correcting for heterogeneity in ability measurement across subjects. Further, the framework makes it possible to test for other potentially important effects of the reform, such as the presence of high school grade inflation.[7]

The model is estimated using administrative data from the University of Toronto, the largest university in Canada. The size of the university and its classes make it possible to observe a large number of Grade-12 and Grade-13 graduates (close to 1,000 students) with similar backgrounds except for Grade-13 Mathematics, 'competing' in same first-year compulsory courses, one of them being a mathematics course. The administrative nature of the data guarantees virtually error-free performance measures and clear identification of the treatment and control groups,[8] and the sample size is large enough to give precise estimates of the value-added of Grade-13 mathematics.

[6] Since the outcome variable studied in this paper is academic performance and not earnings, I will refer to the benefit of schooling as "the value-added of" and not "the return to" schooling in order to avoid any potential confusion.

[7] In this paper, grade inflation will be taken to mean that one group has been graded more (or less) severely than the other. One type of grade inflation that cannot be identified here arises if both groups had their grades increased by a same amount.

[8] The data contain a G12/G13 indicator for every student. See Section 3 for more details.

The main finding of the paper is that, for these high-ability students, the estimated (human capital) benefit to an extra year of high school math is small: students coming out of Grade 13 have a 2.2 point advantage (on a 100 point scale) over students from Grade 12, representing 17 percent of a standard deviation (σ) in mathematics performance. My within-sample investigation also suggests that the extra year of mathematics benefits lower-ability students more than higher-ability students. Comparing my results to those from a related study by Krashinsky (2006), based on the impact of the same reform on university-bound students with lower high-school averages further indicates that there is substantial heterogeneity in the benefit to an additional year of high school mathematics.[9] The estimated effect of Grade 13 found in my paper (0.17σ) is far below the 0.5σ to 1.2σ range found in Krashinsky (2006).

Though the main identification strategy controls for ability differences between the control and treatment groups due to self-selection, other factors could potentially affect the identification of the value-added of Grade 13 mathematics. The results are precisely estimated and robust to changes in estimation technique or control group. In particular, all estimation strategies used in this paper (means comparison, differences-in-differences, OLS, or GMM) suggest that the value-added of Grade 13 is modest. The choice of the control subject is not important: I obtain very similar results when using chemistry instead of biology as control subject. More importantly, robustness checks suggest that the age difference between Grade 12 and Grade 13 students does not affect the estimate of the value-added of Grade 13. Controlling for age using age regressors or restricting the sample to students close in age does not affect the estimate of the value-added. One might also be concerned that Grade 12 students tried to compensate for their lack of mathematics preparation. Yet they did not drop out from mathematics courses more readily than Grade 13 students and they did not avoid mathematics-intensive programs – the program enrolment numbers are very similar. There remains a possibility that Grade 12 students reallocated their study effort to concentrate on mathematics; yet comparing Grade 12 students' relative performance across subjects does not suggest any effort substitution.[10]

1.1 Related Literature and Policy Relevance

Given the robustness of these findings, the analysis has broader relevance, apparent after placing the paper in the context of the prior literature. Several prior studies have looked at the impact of high school curriculum on wages, primarily to shed light on the "human capital/screening" debate. If schooling serves mainly as a screening device, we would expect

[9] The first drafts of this paper and the paper by Krashinsky were written simultaneously, both taking advantage of unique features of the reform. More details about Krashinsky's (2006) paper are found in Section 7.2.

[10] I investigated the possibility of effort substitution across subjects by verifying whether subjects for which students are not expected to take any mathematics courses (e.g. Humanities) favor Grade 12 students more than subjects for which students are expected to take mathematics (e.g. biology, chemistry, economics).

the return to a year of schooling to be larger than the return to a year-equivalent of courses.[11] On this theme, Altonji (1995), Levine and Zimmerman (1995), and Rose and Betts (2004) have looked at the impact of high school curriculum on wages, with a special emphasis on mathematics, using IV techniques. Altonji (1995) investigates the effect of curriculum on wages for a general population of high school graduates using data from the National Longitudinal Survey of High School Class of 1972 (NLS72); Levine and Zimmerman (1995) focus on gender differences in the effect of curriculum on wages, occupation, and education outcomes using the National Longitudinal Survey of Youth (NLSY) and the 1980 senior cohort of High School and Beyond (HSB); and Rose and Betts (2004) examine the effect of specific types of mathematics courses on earnings, using the 1980 sophomore cohort of HSB.

In order to estimate the effect of curriculum on earnings, these studies regress students' log-earnings a few years after graduation on the number of credits (or semester hours) in specific subjects (Altonji (1995), and Levine and Zimmerman (1995)) or types of mathematics acquired in high school (Rose and Betts (2004)), as well as control variables including educational attainment, ability measures (GPA or test scores), family characteristics, and demographic characteristics.[12] They instrument an individual's curriculum using the average curriculum at the respondent's high school to correct for potential course selection bias.[13] The IV estimates all suggest that taking more mathematics in high school does not increase earnings significantly when mathematics credits are not disaggregated into types of mathematics. However, Rose and Betts' IV estimates suggest that some specific types of mathematics courses (e.g. algebra/geometry) have a significant positive effect while others (e.g. calculus) do not.

Since the samples used by Altonji (1995) and Levine and Zimmerman (1995) contain relatively educated individuals, their results could be interpreted as showing that college-bound students do not gain much from extra mathematics.[14] On the one hand, the underlying rationale might be that students acquired little human capital from specific courses (or schooling in general) – the return to schooling would come primarily from signaling. In contrast, it is also possible that students might acquire significant course-specific human capital that was just not rewarded in the labor market; that is, the labor market might reward specific human capital (curriculum) and general human capital (schooling) differently.

This paper helps inform that literature by providing a new way of measuring the amount of human capital learned in specific courses. By looking at student academic performance, I can estimate how much subject-specific human capital college-bound students acquire directly.

[11] To borrow the apt phrase of Rose and Betts (2004), we should see that "the whole is greater than the sum of its parts."

[12] Altonji (1995) and Rose and Betts (2004) also include school characteristics in some specifications.

[13] Altonji (1995) and Levine and Zimmerman (1995) both use a vector of the average number of courses taken in each subject (e.g. English, fine arts, and mathematics) at the respondent's high school as the instrument. Rose and Betts (2004) use a vector of the average number of mathematics credits taken in each type of mathematics (algebra/geometry, calculus, and vocational) at the respondent's high school.

[14] Rose and Betts (2004) include high school dropouts in their analysis unlike Altonji (1995) and Levine and Zimmerman (1995).

The exogenous change in curriculum due to the Ontario Secondary School reform, combined with multiple measures of student academic ability make it possible to get clear identification of the value-added of Grade 13 mathematics, helping to shed light on the forces driving the results from these important earlier studies.

Finding that the benefit to an extra year of high school mathematics is small for high-ability students has two important implications. First, as rehearsed, it provides a possible explanation as to why the studies by Altonji (1995) and Levine and Zimmerman (1995) find only modest (or no) effects of taking more mathematics in high school on wages. The results in the current study suggest that high-ability students gain little curriculum-specific human capital from an extra year of high school. Hence, one should not expect large effects of taking more math on wages for these individuals.[15] Since Grade 13 mathematics was essentially Calculus, the results also help (using the same logic) to account for the finding that there is no significant effect of Calculus on wages (Rose and Betts, 2004). Policies compressing the high-school curriculum might not be harmful for high-ability students. In fact, if these students start working a year early as a consequence of the compression, these policies could be beneficial for them. As a second implication, within- and between-sample comparisons point to the presence of significant heterogeneity in the benefit to high school courses across ability levels. As Lang (1993) and Card (1995) suggest for schooling in general, the benefit to an extra year of high school mathematics could be larger for lower ability students.

The rest of this paper is organized as follows: In the next section, I present characteristics of the 1997 Ontario Secondary School Reform which allow for the identification of the value-added of Grade 13 mathematics. Data are described in Section 3. Section 4 presents results from estimating the value-added of Grade 13 math using popular estimation methods such as simple means comparison, difference-in-differences and OLS regression; I then discuss shortcomings of these methods briefly in the context of the Ontario Secondary School Reform briefly. A model which accounts for these shortcomings is introduced in Section 5, and parameter identification and estimation strategies are presented in Section 6. Key results are reported in Section 7, and the robustness of the results is discussed in Section 8. Section 9 concludes.

2 The Ontario Secondary School Reform

In 1997, the provincial government of Ontario announced that it would compress its secondary school curriculum from five to four years. This reform would bring Ontario into line with most surrounding provinces and potentially lower the costs of the educational system in a significant way. Thus, starting in 1999, students were expected to graduate from high school after four years (after Grade 12) instead of five. A few years later, in 2003, the first cohort

[15] I estimate the earnings difference between Grade 13 and Grade 12 students resulting from a 2.2 point difference in university math performance to be around 2 percent using findings by Jones and Jackson (1990) and Loury and Garman (1995), and back-of-the-envelope calculations. See Section 7 for more details.

of students from the new curriculum graduated from high school, and in the same year, Grade 13 was also abolished. Thus, in 2003, Ontario colleges had students with two different high school backgrounds in the same classes: some students had four years of high school (henceforth referred to as 'G12' students), while others had five ('G13' students).

The intensity of the treatment effect on university preparation should not be seen as being uniform across subjects: the reform did not simply force students to take one less year of schooling. Even though students were now expected to graduate after four years instead of five, they still had to complete the same number of credits (30) as their predecessors in order to satisfy the OSS Diploma requirements.[16] We might think that students from the two curricula (G12 and G13) learned the same material. But college-bound students also need to satisfy college admission requirements, which depend on the program they plan to attend.

An inspection of changes in two subject-specific high school curricula (biology and mathematics) illustrates the heterogeneity across subjects in the effects of the reform on the amount of material taught to university-bound students. Figure 1 shows the transition between the old and new biology course sequences required of typical university-bound life-science students.[17] Prior to taking a biology course, both groups should have successfully completed Grade 9 and Grade 10 Science courses. Despite the reform, the amount of biology material taught in high school is similar for both groups: G12 students have to take essentially the same two courses that were offered in the G13 program.[18] Conversations with professors at the Ontario Institute for Studies in Education at the University of Toronto, and comparison of covered-topics description of these two biology sequences confirm the similarity between the two sequences.

While the impacts of the reform on biology and on a majority of subjects were minimal, this is not true for mathematics or the English course sequences. For these subjects, obtaining the senior high school year credit requires a sequence of prerequisites starting in Grade 9 that changed under the reform. Figure 2 illustrates the transition from the G13 to G12 curriculum in the mathematics sequence followed by a typical university-bound high school student. The reform clearly affected the sequence of courses: under the new system, students were now expected to take four courses in mathematics instead of five. The amount of material covered in class was affected, with less material covered, and less time spent on some topics. Of note, it is common knowledge that essential information was purged from the G13 math curriculum, as illustrated by the Council of Ontario Universities (2002): "We recognize that students in the new curriculum will have less calculus preparation in high school." Comparison of the covered-topics description of these two mathematics sequences

[16] The Ontario Ministry of Education and Training (1999a) defines a credit as "a means of recognition of the successful completion of a course for which a minimum of 110 hours has been scheduled."

[17] All students interested in pursuing a life-science university education should complete a sequence of two biology courses prior to attending university. This was true for students enrolled in the G13 curriculum and it is still true today for G12 students. University preparation courses taken by G13 students were known as Ontario Academic Credit (OAC) courses.

[18] See Ontario Ministry of Education (1987, 2000b).

shows that some material which used to be covered in the later stages of the G13 sequence (e.g. integration and derivatives of trigonometric functions) were not covered in the G12 sequence at all, suggesting that the reform was not a simple repackaging of the material covered in the former sequence.[19,20] If one sees a year of schooling as the product of material covered and time spent on this material, then comparing the G12 and G13 sequences further emphasizes a possible loss in human capital accumulation.

The heterogeneity in the treatment intensity will be very useful in identifying the value-added of Grade 13 math in the presence of selection issues. By observing students' university performance in at least two subjects – biology and mathematics – one of which was not affected by the reform, it becomes possible to control for potential unobserved differences across groups and so achieve identification of the value-added of Grade 13 mathematics.[21]

3 Data

The student data used in this study are provided by the Faculty of Arts and Science of the University of Toronto, one of the largest universities in North America. These administrative data contain information about 2003 students' first-year university academic performance (e.g. grades, dropped courses, program[22]), and pre-admission academic history (e.g. high school average, identification of secondary school institutions attended, and an indicator of secondary school curriculum graduated from – G12/G13). The data also contain each student's date of birth, gender, and her/his student number.

One advantage of using administrative data for this type of study is that the observations are free of recall bias.[23] For example, since we have an indicator of the secondary school curriculum attended by each student (G12/G13 indicator), we do not have to rely on her date of birth to decide which curriculum the student graduated from. That said, the date of birth allows us to concentrate on the population we are most interested in, namely students born in 1984 and 1985.

I restrict the sample to students enrolled in the Life Sciences program. The advantages of doing so are numerous. First, this is a large program which allows the researcher to observe students taking both a course affected by the reform – mathematics – and another which was not – biology. Second, these subjects are likely to be largely "independent" in that knowledge

[19] See Ontario Ministry of Education (1985, 2000a).

[20] By looking at course-specific strands, I show in Appendix A.3 that less material is indeed covered and less time is spent on specific material (e.g. calculus) under the new program.

[21] I will be able to address other potential problems such as potential maturity and dropout issues as well by exploiting the richness of the data.

[22] In 2003, students interested in studying at the University of Toronto Faculty of Arts and Science had to apply to one of the following programs: Commerce, Computer Science, Humanities and Social Sciences, and Life Sciences.

[23] For related studies using college administrative data, see Sacerdote (2001) using Dartmouth College data, and Angrist, Lang and Oreopoulos (2006), and Hoffmann and Oreopoulos (2009) using data from a large Canadian university.

of biology should not affect a student's knowledge of mathematics and vice versa.[24]

The third advantage of focusing on Life Sciences is that students interested in a Life Sciences discipline have to complete a list of compulsory courses during their first year of university. This allows me to alleviate course selection issues. All first year students must take the same biology course (BIO150Y), and almost all programs require an introductory calculus course (MAT135Y). About 90 percent of students for whom we observe a grade for BIO150Y also had a grade for MAT135Y.[25]

As a fourth advantage, Life Sciences students' backgrounds, except for G12/G13 differences, are similar. Before joining the Life Sciences program, G12 students must have successfully completed (in high school) *Advanced Functions and Introductory Calculus* (MCB4U), while G13 students must have high school *Calculus* (MCA0A). These are the last courses of the standard university-preparation course sequence of their respective curricula, as seen in Figure 2. Hence, Life Sciences G13 students have one more year of high school mathematics than their G12 classmates, as described in Section 2 and Appendix A.3. Students should also have a senior high school biology credit.[26] Hence, students are expected to have completed both course sequences of their respective curriculum; these are shown in Figures 1 and 2.

Table 1 presents descriptive statistics on these double cohort students. The last two columns of the bottom panel present differences across groups of mean characteristics and their associated standard errors. Aside from the age difference,[27] the two groups of students seem very similar: they take roughly the same number of university courses[28] and are both composed of a majority of female students with excellent high school averages. Although discussed in more detail in the following sections, a quick inspection of university grades for BIO150Y and MAT135Y presented in Table 1 do not suggest large differences in university performance across the two groups.

4 Estimating the Value-Added of Grade 13

This section highlights potential problems in using standard techniques, such as means comparison, OLS, or difference-in-differences, to estimate the value-added of Grade 13, motivating the need for a more flexible estimator which will be presented in the next section.

[24] English was not analyzed in this paper for this reason. I could not find a program in which we observe students taking both English and another subject reasonably "independent" of it.

[25] This is true for both groups of students. It is not surprising to observe such a high proportion of students taking mathematics as well as biology since students may be uncertain about their exact preferences in terms of field of specialization and might simply insure against this uncertainty.

[26] Most fields in the Life Sciences (39 out of 43) require students to have a senior high school biology credit.

[27] I discuss potential identification issues due to the age difference in Section 8.2.2. The age difference does not seem to affect the estimated value-added of Grade 13 mathematics.

[28] Note that the difference in the number of university courses is statistically significant but economically very small. I discuss in more detail the number of university courses taken by students in Section 8.

4.1 Means Comparison

Consider the situation where two factors influence a student's average mathematics performance when comparing G12 and G13 – the curriculum taken and student ability. The expected difference in mathematics performance (Δ_M) could then be characterized by the sum of the value-added of G13 (Δ_V) and the difference in average initial level of ability between G12 and G13 (Δ_η).[29] Thus,

$$\Delta_M = \Delta_V + \Delta_\eta. \tag{1}$$

If the reform could be thought of as a random experiment, we might expect the difference in the average level of (initial) ability to be negligible ($\Delta_\eta \simeq 0$). Then the difference in mathematics performance would fully capture the effect of the reform (i.e. $\Delta_M = \Delta_V$).

Table 1 presents students' performance in MAT135Y. The results suggest that the value-added of Grade 13 is very small (0.4 points on a 100 point scale), if not zero, since the difference in performance is not statistically significant.

Notice that in equation (1), if both Δ_V and Δ_η are different from zero, there is no way to disentangle the value-added from the difference in ability. In particular, if G12 students have a higher average level of ability than G13 students, then G12 students' ability could compensate for lack of knowledge usually acquired in Grade 13. There are reasons to think this may be the case. Since two cohorts of students were expected to graduate from secondary school simultaneously in June 2003, the double cohort created an expected surge of applicants for post-secondary institutions for September 2003. We can see the dramatic increase in the number of Ontario university applicants clearly in Figure 3. Between 2001 and 2003, the number of applicants (per year) increased from about 60,000 to close to 102,000. This increase has been expected by students and parents since the announcement of the reform (1997), and is likely to have given rise to behavioral effects.

The expected increase in the number of applicants for 2003 led some students to try to avoid the double cohort. For example, it was possible under the G13 curriculum to "fast-track" the program and graduate after four years,[30] with the fear of the double cohort probably encouraging some G13 students to try to fast-track and graduate in 2002 instead of 2003.

This idea is supported by Figure 3. The number of applicants rose by about 16% (from 60,000 to 69,000) between 2001 and 2002, which is much larger than the average increase prior to 2001, suggesting that some G13 students successfully escaped from the double cohort.[31] If, by fast-tracking, "high" ability G13 students disappeared from the 2003 cohort, the average ability of 2003 G13 students would probably be lower than the average ability

[29] The initial level of ability is the general level of academic ability acquired prior to secondary schooling.

[30] Even though it was possible to fast-track secondary school, this was far from being common practice. Prior to 2002, around 8% of Ontario university students had graduated from high school after four years.

[31] Demographics cannot explain such increase. The number of 19 year-olds in Ontario increased by 3.4% in 2002. Source: Statistics Canada, CANSIM Table 051-0001.

of 2003 G12 students. Furthermore, the number of applicants seems larger than expected in 2004, suggesting that some G12 students delayed their university applications. If we think that this behavior is more likely to occur among "low" ability students, then we have even more reason to think that the estimator of the value-added of Grade 13 would be biased downward when simply comparing mathematics performance.

4.2 OLS and Difference-in-Differences Estimation of the Value-Added

Observing more than one university outcome for each student can help to control for any difference in ability. We could use a student's biology grade (or high school average in the absence of grade inflation) as a proxy for ability and regress the university mathematics grade on the university biology grade and a dummy variable $G13_i$, equal to 1 for G13 students and 0 for G12 students:

$$M_i = \alpha + \Delta_V \times G13_i + \beta B_i + u_i \tag{2}$$

where M_i and B_i are the student's university mathematics and biology grades, respectively. Δ_V should measure the value-added of Grade 13 math if B_i correctly measures student ability.

Table 2 presents OLS regression results using biology or high school grades as an ability proxy. Results from using biology grades (column I) suggest that the value-added of Grade 13 is small (1.68 points) but statistically significant. To put this in perspective, the standard deviation of grades in MAT135Y is about 13 points. Results from using high school average as a proxy for ability (column II) suggest that the value-added is not statistically different from zero. A potential problem with the OLS interpretation is that we assume that biology measures ability perfectly. If not, the measure of the value-added of Grade 13 will be biased.[32]

If one assumes that mathematics and biology measure ability the same way (up to a constant), we can use the difference in average biology grades (Δ_B) as a measure of the difference in ability (Δ_η)

$$\Delta_B = \Delta_\eta \tag{3}$$

We can then construct a difference-in-differences estimator using equations (3) and (1) to identify the value-added of G13:

$$\Delta_{DD} \equiv \Delta_M - \Delta_B = \Delta_V \tag{4}$$

The difference between differences in average university mathematics grades (Δ_M) and in average biology grades (Δ_B) would give us the value-added of Grade 13 (Δ_V). The difference in average biology performance presented in Table 1 suggests that G12 students do significantly better in biology than G13 students, which also suggests that we are facing two different groups in terms of ability levels.

Table 3 presents the difference-in-differences estimate of the G13 value-added based on

[32] The sign of the bias will depend on the groups' relative performance in university biology. See Appendix A.2 for details.

equation (4). The difference-in-differences estimate is more than four times greater than the average difference in mathematics performance. It is also statistically significant. The estimate is precisely estimated but it is still small when compared to the students' mathematics average (70.4) and standard deviation (13.1). In particular, a G12 student would have had a 1.9 point increase in her mathematics performance in the absence of the reform. For many students, this difference would not affect their GPA. Only students close a grade cut-off (e.g. between an A and a B) might see their GPA suffer from missing Grade 13.[33]

The difference-in-differences estimator assumes that biology and mathematics measure students' ability in the same way (up to a constant). If biology does not measure ability in the same way that mathematics does, we can write

$$\Delta_B = \lambda^B \Delta_\eta \tag{5}$$

where $\lambda^B \neq 1$. Then

$$\begin{aligned} \Delta_{DD} &= \Delta_M - \lambda^B \Delta_\eta & (6) \\ &= \Delta_V + \frac{\left(1-\lambda^B\right)}{\lambda^B} \Delta_B & (7) \end{aligned}$$

where equation (7) is obtained using equations (1) and (5). Hence, if both $\Delta_B \neq 0$ and $\lambda^B \neq 1$ then the difference-in-differences estimator will also be biased. I have already shown that $\Delta_B \neq 0$; I now show that mathematics and biology might measure ability differently.

Intuitively, if mathematics and biology measure ability the same way, they should have the same relationship with a third measure of ability. But when we look at Table 4, we can see that the sample covariances between the high school average and biology and mathematics differ. The difference is consistent across groups. The covariance between biology and the high school average is between 15 and 20 percent smaller than the covariance between mathematics and high school (e.g. 16.7/20.5). Not only might the two groups differ in ability, but the two measures of ability used to capture the value-added of G13 might not do so in the same way.

Since we know that Δ_B is negative, the sign of the bias will depend on whether $\lambda^B > 1$ or $\lambda^B < 1$. If $\lambda^B < 1$, for instance, the difference-in-differences will be downward-biased. Hence, to correctly estimate the value-added of Grade 13, one needs an estimator that will take into account both the differences in ability between G12 and G13 students *and* the fact that biology and mathematics measure ability differently. This can be done using a factor model.

5 A Grading-Rule Model

In this section, I propose a grading-rule model that can account for the relationship between human capital accumulation and academic performance in the specific environment of the

[33] At the University of Toronto, Grade B covers scores from 70 to 79. Grade A covers scores 80 and above.

double cohort. In particular, the model shows how available observables – university grades, high school grades, and high school curricula – can be linked. It is constructed such that the estimator of the G13 value-added is a generalization of the standard difference-in-differences estimator presented above.

The grading-rule model accounts for differences in ability measurement between mathematics and biology. The model also allows for grade inflation at the high school level, different levels of average ability across groups, and heterogeneity in ability measurement across subjects. I begin by summarizing the main elements.

5.1 Factors Influencing Student Academic Performance

Assume that a student's grade can be thought of as the product of three factors: the student's academic ability, the grading rule of the academic institution (high school or university), and a curriculum effect. Other possible influences such as student effort and teacher quality will be not be explicitly modeled here since the data do not contain information on these factors.[34]

Students begin high school with an initial level of general academic ability. Instead of seeing this 'ability' as defined solely by the individual's innate characteristics (such as IQ), we will view it as the stock of human capital the student brings to the learning process – the joint product of the individual's own innate, acquired, and environmental characteristics. Academic ability, thought of in these terms, is assumed to be partially unobservable; neither the econometrician nor teachers can measure it perfectly.

A grading rule is a means that teachers and professors use to signal (via a grade) a student's human capital. Different subjects may call upon and foster different types of skill, in which case there could be different grading rules for different courses. The model will allow this.

As the third component, student performance is likely to be influenced by a group-specific curriculum effect when compared to students with different high school curriculum backgrounds. In the context of the Ontario reform, the difference in the curriculum effect between G13 and G12 students represents the value-added of G13. The curriculum determines how much human capital a group will acquire during high school. Its effect is not only group-specific but also subject-specific, reflecting the fact that a curriculum change can affect some subjects more than others.

5.2 Model

I now present the model more formally, first specifying the way that human capital is accumulated through high school and university, then specifying how this human capital is signaled

[34] One could imagine that student effort might interact with student's ability and that teacher quality could influence the curriculum effect. If student effort or teacher quality do not differ across G12 and G13 students, then abstracting from these factors should not affect my results. I nevertheless investigate the potential identification issues from not observing student effort in Section 8.2.3 as well as providing robustness evidence.

via a grade by teachers at each stage.

5.2.1 Human Capital Accumulation

There are two institutions, superscripted by uppercase $I = \{H, U\}$, at which students accumulate subject-specific human capital: high school ($I = H$) and university ($I = U$). Assume only two subjects $S = \{B, M\}$[35] that are taken both at high school and university. Finally, there are two groups of students who take different curricula $C = \{G12, G13\}$ while in high school, but who then take the same courses in university.

Student i is initially endowed with a level of general academic ability (η_i) and then accumulates subject-specific human capital as she attends high school and university. While in high school, $G13$ students receive a treatment which affects the amount of mathematics-specific human capital they acquire ($G12$ students do not receive the treatment). High school biology is not affected by the treatment, and hence both groups are assumed to acquire the same biology material in high school.

The form of the human capital accumulation processes which drive students' performance in university[36] in biology and mathematics are assumed to be as follows:

$$\eta_i^{U,B} = \eta_i + \tau^{H,B} + \tau^{U,B} \quad (8)$$

$$\eta_i^{U,M,C} = \eta_i + \tau^{H,M,C} + \tau^{U,M} \quad (9)$$

The components in equations (8) and (9) are not directly observed. The variables on the left-hand sides of (8) and (9) represent the underlying stocks of biology- and mathematics-specific human capital accumulated by the end the first year in university, respectively. The first term on the right-hand side of each equation is the student's initial level of general academic ability, the second term is the amount of subject-specific human capital acquired through high school (superscripted by H), and the third term is the amount of subject-specific human capital acquired during first year of university (superscripted by U).

To be clear about the notation, in equation (8), since students are assumed to acquire the same amount of biology-specific human capital in high school, there is no curriculum superscript for biology ($\tau^{H,B,G13} = \tau^{H,B,G12} = \tau^{H,B}$). G12 and G13 students also acquire the same amount of human capital in university ($\tau^{U,B}$ and $\tau^{U,M}$) since they are in the same classes. In contrast, the amount of mathematics-specific human capital accumulated in high school will depend on the curriculum attended by the student. Thus we might have $\tau^{H,M,G13} \neq \tau^{H,M,G12}$. For this reason, the level of mathematics-specific human capital by the end of the first year of university ($\eta_i^{U,M,C}$) will depend on the curriculum taken by student i.[37]

[35] B stands for Biology while M stands for Mathematics.

[36] Hence the U superscripts on both left-hand side variables.

[37] The accumulation of subject-specific human capital in high school and university ($\tau^{H,B}$, $\tau^{U,B}$, $\tau^{H,M,C}$, and $\tau^{U,M}$) should be seen as functions of the amount of material taught and the time spent on the material.

5.2.2 Grading Rules

For each student, we observe three grades: 1) a mathematics university grade, 2) a biology university grade and 3) an overall high school average. These grades signal the student's subject-specific levels of human capital. In particular, grades are assumed to be assigned according to a linear rule. Part of the grade is assigned on a relative basis, comparing the student's subject-specific human capital compared to her classmates. Since the two groups are separated prior to university, high school grades only represent performance with respect to the student's own group (i.e. G12 or G13). For example, in high school biology, a teacher would compare a student's level of human capital ($\eta_i + \tau^{H,B}$) to the group average ($E_C[\eta_i] + \tau^{H,B} \equiv \bar{\eta}^C + \tau^{H,B}$). Since students from the same group learn the same material by assumption, a student's grade will only depend on her initial general academic ability (η_i) compared to the group average ($\bar{\eta}^C$).

The high school grading rules (for biology and mathematics) are assumed to be linear in the student's relative level of human capital ($\eta_i - \bar{\eta}^C$). As a consequence, the high school average (H_i^C) is also a linear function of the difference between the student's initial academic ability and the average initial ability of the group she belongs to:

$$H_i^C = \pi^{H,C} + \lambda^{H,C}(\eta_i - \bar{\eta}^C) + \varepsilon_i^{H,C}$$

The slope and the intercept coefficients (respectively $\lambda^{H,C}$ and $\pi^{H,C}$) represent averages of slope and intercept coefficients across high school subject grading rules. These coefficients are under the teacher's control and are allowed to vary across high school curricula. We can rewrite the high school average more simply as

$$H_i^C = \upsilon^{H,C} + \lambda^{H,C}\eta_i + \varepsilon_i^{H,C} \tag{10}$$

where

$$\upsilon^{H,C} = \pi^{H,C} - \lambda^{H,C}\bar{\eta}^C.$$

Figure 4 illustrates possible high school grading rules for G12 and G13 students. The slope coefficient ($\lambda^{H,C}$) represents the payoff to ability. The difference in $\upsilon^{H,G13}$ and $\upsilon^{H,G12}$ represents grade inflation.[38] Grade inflation has two potential sources: one source comes from the way teachers link ability to grades, and is measured by $\pi^{H,C}$ and $\lambda^{H,C}$. The other source comes from the student population within each group and is captured by $\lambda^{H,C}\bar{\eta}^C$. This term captures the notion that the weaker are the students from a group (low $\bar{\eta}^C$), the easier it is to achieve a high grade within this group. The error term ($\varepsilon_i^{H,C}$) represents shocks due

This model does not disentangle the effects of 'time spent on material' from the effect of 'more material'. One could imagine $\tau = f(time, material)$, where both time and material have a positive effect on $\tau^{H,M,C}$. The model assumes an additively separable production function, which is not trivial. A different model in which the terms enter the production function in a multiplicative way was also estimated with very similar results.

[38] Grade inflation is taken to mean that one group has been graded more (or less) severely than the other.

to measurement error and possible shocks to student performance (e.g. bad luck or illness). $\varepsilon_i^{H,C}$ is assumed to have mean 0 and is uncorrelated with the student's ability. Notice that only the left-hand-side variable of equation (10) is observed.

In the same fashion, the university biology grading rule is given by

$$B_i^C = \pi^B + \lambda^B(\eta_i^{U,B} - \bar{\eta}^{U,B}) + \varepsilon_i^{B,C} \tag{11}$$

University students are now compared to classmates from both groups. For this reason, we have $\bar{\eta}^{U,B}$ and not $\bar{\eta}^{U,B,C}$ in the grading rule equation. I assume that professors do not discriminate against students based on their high school background.[39] The constant and the slope coefficients are then assumed to be the same for both groups. Similar to the high school grading rule, the error terms represent shocks that can be due to simple measurement error but also to temporary shocks affecting students' performance. These error terms are assumed to be uncorrelated with a student's ability but also uncorrelated with each other ($E(\varepsilon_i^{B,C}, \varepsilon_i^{H,C}) = 0$). We can rewrite (11) more simply as a function of the student's initial level of general academic ability and the population average initial level of general academic ability

$$B_i^C = \upsilon^B + \lambda^B \eta_i + \varepsilon_i^{B,C} \tag{12}$$

where

$$\upsilon^B = \pi^B - \lambda^B \bar{\eta}.$$

The variable $\bar{\eta}$ represents the total population initial level of general academic ability. According to equation (12), the expected biology grades for G12 and G13 students are $\upsilon^B + \lambda^B \bar{\eta}^{G12}$ and $\upsilon^B + \lambda^B \bar{\eta}^{G13}$, respectively. Hence, the difference between the expected biology grades ($\bar{B}^{G13} - \bar{B}^{G12} = \Delta_B$) is given by equation (5), i.e. $\Delta_B = \lambda^B \Delta_\eta$.

The mathematics high school sequence was affected by the reform. Thus, we can imagine that both the student's initial level of human capital and her curriculum will affect her grade, yielding:

$$M_i^C = \pi^M + \lambda^M(\eta_i^{U,M,C} - \bar{\eta}^{U,M}) + \varepsilon_i^{M,C}$$

or

$$M_i^C = \upsilon^{M,C} + \lambda^M \eta_i + \varepsilon_i^{M,C} \tag{13}$$

where

$$\upsilon^{M,C} = \pi^M - \lambda^M \bar{\eta} + \lambda^M \left(\tau^{H,M,C} - \frac{N_{G13}}{N} \tau^{H,M,G13} - \frac{N_{G12}}{N} \tau^{H,M,G12} \right)$$

with N_{G13} being the number of G13 students and N is the total number of students ($N = N_{G13} + N_{G12}$). The last term of the constant, in parentheses, represents the effect of the curriculum on the students' performance. Note that this is not the value-added of Grade 13. The difference between $\upsilon^{M,G13}$ and $\upsilon^{M,G12}$ will represent the value-added of Grade 13

[39] In reality, most professors (and teaching assistants) had no idea who had which background.

$(\Delta_V = \tau^{H,M,G13} - \tau^{H,M,G12})$ if and only if $\lambda^M = 1$.[40]

The grading rule model consists of equations (10), (12) and (13). We can easily see the resemblance to a standard one-factor model where the driving factor is the initial level of general academic ability (η_i). One necessary condition for identification of the factor model parameters is that the latent variable (η_i) must be scaled to one observed variable. That is, the slope and the intercept coefficients of one equation should be predetermined. Usually, the choice of the benchmark is irrelevant, but since I am interested in the difference between $\upsilon^{M,G13}$ and $\upsilon^{M,G12}$, a convenient normalization is to set the constant and slope coefficient of the mathematics grading rule for G12 students ($\upsilon^{M,G12}$ and λ^M) equal to 0 and 1 respectively. This normalization implies that mathematics professors compensate students such that any direct grade inflation (the grade inflation under their control) cancels out the indirect grade inflation (due to the population average ability).

Overall, the grading rule model can be summarized in six equations:

$$\begin{aligned}
H_i^{G13} &= \upsilon^{H,G13} + \lambda^{H,G13}\eta_i + \varepsilon_i^{H,G13} && \text{(10a)}\\
H_i^{G12} &= \upsilon^{H,G12} + \lambda^{H,G12}\eta_i + \varepsilon_i^{H,G12} && \text{(10b)}\\
B_i^{G13} &= \upsilon^{B} + \lambda^{B}\eta_i + \varepsilon_i^{B,G13} && \text{(12a)}\\
B_i^{G12} &= \upsilon^{B} + \lambda^{B}\eta_i + \varepsilon_i^{B,G12} && \text{(12b)}\\
M_i^{G13} &= \Delta_V + \eta_i + \varepsilon_i^{M,G13} && \text{(13a)}\\
M_i^{G12} &= \eta_i + \varepsilon_i^{M,G12} && \text{(13b)}
\end{aligned}$$

where only the left-hand sides of each equation are observable. Notice that, after the normalization, $E(M_i^{G13} - M_i^{G12}) = \Delta_V + \Delta_\eta$, just as in equation (1).

6 Identification and Estimation

The grading-rule model is summarized by a system of equations in which the correlation between the observables (the left-hand-side variables) is due to a single common factor (η_i). In fact, the only difference between this model and a pure factor model is that I allow the equations to have constant terms. The identification strategy follows the approach used in the factor models literature closely.[41] In order to identify the G13 value-added, I will use the basic hypothesis of these models which stipulates that, if the model is correct, the covariance matrix (Σ^C) of curriculum C's observed grades should be exactly reproduced by the covariance matrix implied by the model ($\Sigma(\Theta)^C$), so that

$$\Sigma^C = \Sigma(\Theta)^C \tag{14}$$

[40] I later normalize λ^M to be equal to 1. Hence, we will obtain the value-added of Grade 13 by taking the difference between $\upsilon^{M,G13}$ and $\upsilon^{M,G12}$.

[41] See Abowd and Card (1987, 1989), Altonji and Segal (1996), Baker (1997), or Altonji, Martins and Siow (2002) for applications of this estimation technique in the labor economics literature.

where

$$\Sigma^C = \begin{bmatrix} var(H_i^C) & cov(H_i^C, B_i^C) & cov(H_i^C, M_i^C) \\ cov(H_i^C, B_i^C) & var(B_i^C) & cov(B_i^C, M_i^C) \\ cov(H_i^C, M_i^C) & cov(B_i^C, M_i^C) & var(M_i^C) \end{bmatrix}$$

and

$$\Sigma(\Theta)^C = \begin{bmatrix} [\lambda^{H,C}]^2\sigma^2_{\eta^C} + \sigma^2_{\varepsilon^{H,C}} & \lambda^{H,C}\lambda^B\sigma^2_{\eta^C} & \lambda^{H,C}\sigma^2_{\eta^C} \\ \lambda^{H,C}\lambda^B\sigma^2_{\eta^C} & \left[\lambda^B\right]^2\sigma^2_{\eta^C} + \sigma^2_{\varepsilon^{B,C}} & \lambda^B\sigma^2_{\eta^C} \\ \lambda^{H,C}\sigma^2_{\eta^C} & \lambda^B\sigma^2_{\eta^C} & \sigma^2_{\eta^C} + \sigma^2_{\varepsilon^{M,C}} \end{bmatrix}$$

where $\sigma^2_{\eta^C} = var(\eta^C)$ and $\sigma^2_{\varepsilon^{B,C}} = var(\varepsilon^{B,C})$. Intuitively, we should expect to have the same kind of relation between the observed-grades' first moments, μ^C, and the first moments implied by the model, $\mu(\Theta)^C$. Hence:

$$\mu^C = \mu(\Theta)^C \tag{15}$$

where

$$\mu^C = \begin{bmatrix} E(H_i^C) \\ E(B_i^C) \\ E(M_i^C) \end{bmatrix} \quad \text{and} \quad \mu(\Theta)^C = \begin{bmatrix} \upsilon^{H,C} + \lambda^{H,C}(E_C[\eta_i]) \\ \upsilon^B + \lambda^B(E_C[\eta_i]) \\ I_{G13}\Delta_V + (E_C[\eta_i]) \end{bmatrix},$$

where I_{G13} is an indicator variable equal to 1 if the student is a G13 student and $E_C[\eta_i]$ is the average level of initial general academic ability of group C. The model has a total of 18 measured moments and there are 17 coefficients to be estimated (Δ_V, $E_{G13}[\eta_i]$, $E_{G12}[\eta_i]$, $\sigma^2_{\eta^{G13}}$, $\sigma^2_{\eta^{G12}}$, υ^B, λ^B, $\upsilon^{H,G13}$, $\upsilon^{H,G12}$, $\lambda^{H,G13}$, $\lambda^{H,G12}$, $\sigma^2_{\varepsilon^{H,G13}}$, $\sigma^2_{\varepsilon^{H,G12}}$, $\sigma^2_{\varepsilon^{B,G13}}$, $\sigma^2_{\varepsilon^{B,G12}}$, $\sigma^2_{\varepsilon^{M,G13}}$, and $\sigma^2_{\varepsilon^{M,G12}}$). The plausible 'no-discrimination' (based on high school curriculum) assumption about the university grading rules and the normalization of the G12 university mathematics grading rule allow for the identification of the model parameters. The simplicity of the model makes it easy to write the parameters of interest as functions of population moments.[42] The slope coefficient of the biology grading rule is

$$\lambda^B = \frac{cov(H_i^C, B_i^C)}{cov(H_i^C, M_i^C)}. \tag{16}$$

Looking ahead, a testable restriction of the model is that the ratio of covariances in equation (16) should be the same for both groups of students. Finding that the ratio of covariances in equation (16) differ across groups would signal that G13 and G12 students were graded differently in university biology (more specifically, $\lambda^{B,G13} \neq \lambda^{B,G12}$).

The value-added of G13 is

$$\Delta_V = \Delta_M - \frac{cov(H_i^C, M_i^C)}{cov(H_i^C, B_i^C)}\Delta_B \tag{17}$$

[42]See Appendix A.1 for more details about the identification of the parameters.

where $\Delta_M = E(M_i^{G13}) - E(M_i^{G12})$ and $\Delta_B = E(B_i^{G13}) - E(B_i^{G12})$. Equation (17) is equation (6) written differently, since the ratio of the covariances is simply λ^B. The estimator of the value-added is thus a modified difference-in-differences estimator allowing for different measures of ability across courses.

We can link (17) to the naive estimators presented in Section 4 above. If biology and mathematics were to measure ability in exactly the same way, then equation (17) would become the standard difference-in-differences estimator. If the two groups of students were identical, then equation (17) would become a simple means comparison. Finally, if students' high school and biology grades were identical ($H_i^C = B_i^C$), then equation (17) would give the OLS estimator.

The groups' average initial ability levels ($E_C(\eta_i)$) are defined by:

$$E_{G12}(\eta_i) = E(M_i^{G12}) \tag{18}$$

$$E_{G13}(\eta_i) = \frac{cov(H_i^C, M_i^C)}{cov(H_i^C, B_i^C)}\Delta_B + E(M_i^{G12}). \tag{19}$$

From equations (18) and (19), we can see that the student performance in the university biology course will sign the difference in average ability ($\Delta_\eta \equiv E_{G13}[\eta_i] - E_{G12}[\eta_i]$), since $1/\lambda^B$ is positive.

The empirical strategy is to fit the sample moments to the moments implied by the model.[43] For each group, we have a fit function defined by

$$F(\boldsymbol{\theta})_C = (\mathbf{s}^C - \boldsymbol{\sigma}(\boldsymbol{\theta})^C)' \mathbf{W}_C^{-1} (\mathbf{s}^C - \boldsymbol{\sigma}(\boldsymbol{\theta})^C) + (\mathbf{x}^C - \boldsymbol{\mu}(\boldsymbol{\theta})^C)' \mathbf{S}_C^{-1} (\mathbf{x}^C - \boldsymbol{\mu}(\boldsymbol{\theta})^C) \tag{20}$$

where x^C and s^C are vectors of sample first and second moments respectively, while $\mu(\theta)^C$ and $\sigma(\theta)^C$ are vectors of first and second moments implied by the model.[44] θ is the vector of parameters I wish to estimate, S_C is the sample covariance matrix and W_C^{-1} is a weight matrix to be defined.

The first part of the equation is the standard GMM fit function used in the analysis-of-covariance literature. The second part is the fit function for first moments which is necessary to estimate the coefficient of G13's value-added. The results are obtained using as the weight matrix W_C^{-1} estimates of the fourth-order moments. This estimator is the Optimal Minimal Distance (OMD) estimator. Following concerns about the use of the OMD estimator voiced by Altonji and Segal (1996), I also used different weight matrices W_C^{-1} to check for any disparities in the parameters estimates due to the choice of the weight matrix. The use of the OMD weight matrix, the identity matrix, or a diagonal weight matrix using fourth-order moments as weight matrices all give very similar results.

The global fit function used in the minimization problem is a weighted average of the

[43] I choose GMM over maximum likelihood (ML) since the normality assumption required for the validity of the ML is rejected for all six outcomes. Nevertheless, the results are not sensitive to the choice of estimation strategy.

[44] See Jöreskog and Sörbom (1996).

groups' fit functions

$$F(\boldsymbol{\theta}) = \frac{N_{G13}}{N} F(\boldsymbol{\theta})_{G13} + \frac{N_{G12}}{N} F(\boldsymbol{\theta})_{G12}$$

and the parameter estimates are given by $\theta_{OMD} = ArgMin_{\boldsymbol{\theta}} F(\theta)$. These are discussed in the following section.

7 Results

Table 5 presents results of the GMM estimation using the OMD estimator weight matrix. Each point estimate comes with two measures of standard errors since those from the OMD estimator may be unreliable (see Altonji and Segal 1996, and Horowitz 1998). Usual OMD standard errors (obtained from asymptotic distribution theory) are presented in parentheses, while bootstrap standard errors are in square brackets. Both measures of standard errors suggest that the value-added of G13 is positive and precisely estimated. Note that bootstrap standard errors are significantly larger for intercept coefficient estimates, and for the ability measure estimates. Controlling for ability, Grade 13 increases a student's mathematics performance by 2.2 percentage points. Comparing the value of G13 to the mathematics average and standard deviation,[45] the benefit to Grade 13 is modest in terms of human capital accumulation. We can get a ballpark estimate of the 'return' (or earnings growth rate due) to G13 mathematics using existing literature. Loury and Garman (1995) and Jones and Jackson (1990), for example, found that a one-point in GPA would lead to about 9-10 percent earnings increase. Using these estimates, and assuming that the 2.2 points would translate directly to the GPA, the return to Grade 13 would be around 2 percent.[46]

All three subject grading rules are different. The difference between the mathematics and the biology slope coefficients is 0.17 (1-0.83) and is statistically significant. The interpretation of this difference is that students' relative proficiency is more easily signaled in mathematics than in biology.[47] The high school grading rule slope coefficients are much smaller than 1. The admission standards, combined with bell-shaped university grading, can explain the difference in university and high school slope coefficients.[48] The difference between the university mathematics and the high school intercepts is about 77 points. This difference captures the greater difficulty of university courses and the more intense competition in university classrooms.

[45] See Table 1.

[46] In order to get this estimate I also assumed that an increase in 2.2 points would lead to a 0.22 increase in GPA since letter grades contain 10 points.

[47] Many factors could explain this difference. For example, the test formats are different: biology test questions are all multiple-choice questions while mathematics uses a mixture of question types. Because of the nature of the multiple-choice questions, luck might play a bigger role, relative to ability, in biology than in mathematics for lower ability students.

[48] Students admitted to the university have high school averages above 80%. At the university level, we usually observe grades varying between 30 and 100%. So, for accepted students, the span of grades is increased between high school and university while the span of ability is fixed. As a consequence, the payoff of an extra unit of ability has to be more important at the university level to cover the new span of grades.

The average levels of initial academic ability seem to differ across groups. This finding, combined with the results suggesting an ability measurement discrepancy across subjects, favors the use of the grading rule model over difference-in-differences estimation. I test the significance of this difference by re-parametrizing the model. The estimated difference $\hat{\Delta}_\eta$ is -1.77, which is just the difference between estimated levels of initial academic ability ($\bar{\eta}^{G13}$ and $\bar{\eta}^{G12}$) obtained in Table 5. Given that the OMD and bootstrap standard errors of $\hat{\Delta}_\eta$ are 0.08 and 0.23 respectively, I can reject the hypothesis of equal average ability. G12 students look brighter than G13 students, which is consistent with the selection story in which more able G13 students escaped from the double cohort (and from the sample).

Results from high school grading rules do not reveal any clear pattern in the way teachers graded students in high school. Even though the intercept coefficient of the high school grading rule for G13 students is more important than for G12 students, the opposite is true for the slope coefficient. Using LR tests, I successively test for the equality of slope coefficients and the equality of the intercept coefficients. Table 6 presents the results.

I test the restrictions of equal slope and intercept coefficients by comparing the fit of the restricted models to the fit of the model used to measure the value-added of G13 (labeled as the baseline model). The first step in doing so is to test whether the model presented in Table 5 fits the data well. If the model is valid, then N times the fit function evaluated at the estimated coefficient values ($d = NF(\hat{\theta})$) is asymptotically χ^2 distributed. The model used to measure the value-added of G13 fits the data well, given to the low value of d. In the case of this model, the test simply looks at whether the ratio of covariances in equation (16) is the same for both groups since there is only one overidentifying restriction. The second step is to compare the fit of the restricted models I want to test with the fit of the baseline model. If the restrictions imposed on the model are valid, then the difference in fit (between the restricted and unrestricted models) measured by Δd is also asymptotically χ^2 distributed.[49] The p-values of these tests are presented in the last column of Table 6. I cannot reject the hypothesis that both high school grading rules have the same slope or intercept coefficient, but I do reject the hypothesis that the grading rules are the same (equal slope and intercept coefficients). The results from the tests might seem surprising but we have to remember that the variation in high school marks is small and that no mark is close to zero. As a consequence, it is almost impossible to disentangle a small shift in intercept from a small shift in slope coefficients.

7.1 Summary of Results

The results all suggest that the value-added to Grade 13 is modest for high-ability students. Estimates of the value-added are similar whether I use means comparison (0.45), OLS estimation (1.68), difference-in-differences (1.92), or the grading rule model (2.21) as way of

[49] The number of degrees of freedom of Δd is given by the difference in degrees of freedom of the compared models. See Chamberlain (1984) for details.

capturing the value-added of Grade 13 mathematics.

That said, the factor model proves to be useful in capturing effects which the other methods presented in the paper do not account for. The results from the factor model show that difference-in-differences estimation would lead to biased estimates of the value-added of Grade 13 if biology and mathematics do not measure ability in the same way. In the present case, the factor model estimate is 15% above the difference-in-differences estimate and 32% above the OLS estimate. The factor model estimate is above the OLS estimate because of the correlation between the amount of schooling and the average level of ability of students. It is also close to five times the means difference estimate, which shows the importance controlling for heterogeneity in average ability level across the two groups.

7.2 Heterogeneity

The estimates found in this paper are far below those found in Krashinsky (2006), who looks at the impact of the same reform on students with lower high school averages than students studied in this paper. The difference in performance between G12 and G13 students found in Krashinsky (2006) ranges between 0.5 and 1.2 standard deviations, while the difference is about 0.17 (= 2.2/13) standard deviations in the present paper.[50] This difference suggests the possibility of heterogeneity in the treatment effect. Students studied in Krashinsky (2006) are from University of Toronto's Scarborough Campus and have a high school average around 84 percent while students studied in this paper have a 91 percent high school average. This difference is considerable: the 'average' student found in Krashinsky (2006) has a high school average close to the minimum average found in this study (83 percent) and about two standard deviations below this group's average.[51] In order to investigate whether the difference in academic ability could explain the difference in estimates, I look for the presence of heterogeneity in the value-added of Grade 13 within the present sample.

I first separated the sample in two groups based on their academic ability. I formed a higher-ability group and a lower-ability group using the median university biology grade as a cutoff point. Estimating the value-added separately for each group, I find that the value-added for lower-ability students is 1.4 points greater than for the higher-ability students. Although modest, the difference has the expected sign − lower-ability students gain more from an extra year of high school. Alternatively, I can introduce an extra parameter in the grading rule model since there is one degree of freedom in the baseline model. For example, it is easy to rewrite equations (8) and (9) to allow for the value-added to be a linear function

[50] Note that Krashinsky's methodology differs from the one used in this paper, which could explain part of the difference in the results. As a check, I estimated the benefit to Grade 13 for MAT135 using OLS and students that are closer and closer in age – as does Krashinsky. Estimates obtained from this strategy are smaller than 2 points, and not statistically significant, suggesting that the methodology used here is not driving the difference between Krashinsky's and my results.

[51] The fact that applications to the University of Toronto are processed separately across campuses, and that students differ in terms of enrolled programs (Life Science in this study vs. Management in Krashinsky) could explain such a difference.

of students' ability:

$$\eta_i^{U,B} = \eta_i(1+\phi(\tau^{H,B}+\tau^{U,B}))+\tau^{H,B}+\tau^{U,B} \tag{21}$$
$$\eta_i^{U,M,C} = \eta_i(1+\phi(\tau^{H,M,C}+\tau^{U,M}))+\tau^{H,M,C}+\tau^{U,M} \tag{22}$$

where ϕ is the heterogeneity coefficient. If the value-added is decreasing with ability, we would expect ϕ to be negative. Note that equations (21) and (22) are identical to (8) and (9) if there is no heterogeneity (i.e. $\phi = 0$).

I re-estimated the model allowing for heterogeneity in the value-added (as specified by equations (21) and (22)) and found that $\hat{\phi} = -0.03$. The estimate is not statistically significant, as we could expect from the estimate of d for the baseline model in Table 6. Overall, it is still surprising to find some evidence of heterogeneity in the value-added of an extra year of high school for such a homogeneous group of individuals.[52]

8 Robustness

8.1 Control and Treatment Groups

The results presented in Table 5 assumed that female and male students receive the same benefit from Grade 13. This might not be the case. In order to investigate the possibility of heterogeneity across gender in the value-added of Grade 13 mathematics, I estimated the grading rule model separately for females and males.

Table 7 presents the results by gender. The estimated value-added of Grade 13 is 2.4 points for female students and 2.0 for male students. Both male and female G12 students have higher initial ability levels than their G13 counterparts. The ability difference is only slightly more important for males students – ability differences are 1.45 and 2.27 points for females and males respectively. Again, OMD standard errors look seriously downward-biased for the ability measure estimator.

Table 7 shows that the results presented in Table 5 are not gender-driven. Both estimation results suggest that the value-added of Grade 13 is modest and that G12 and G13 students are different in terms of initial levels of ability.

I replicated the experiment using chemistry instead of biology. Chemistry is another course that life science students must take which was not affected by the reform and for which a student's performance should not be influenced by her mathematics knowledge. The results are similar to the ones presented here (the estimated value-added of G13 is 1.7 points). I also replicated the estimates using chemistry instead of mathematics. In this case, any evidence of value-added for Grade 13 would be problematic. The estimated value of

[52] In addition, while high ability students do not seem to have been severely affected by the reform, King *et al.* (2002, 2004, 2005) report that lower ability students were adversely affected by the curriculum compression. King *et al.* (2004) note when talking about workplace-bound students' credit accumulation toward high school graduation: "These data suggest there are serious problems with the progress of students taking Applied courses."

Grade 13 ($\hat{\Delta}_V$) in this case is very small (0.25). This evidence supports the hypothesis that biology and chemistry were not affected by the reform.

Because covariances are sensitive to outliers, I excluded 12 students with grades below 30% in either biology or mathematics and assumed that these students dropped out. Including these students does not change the results ($\hat{\Delta}_V = 2.17$ as compared to 2.21). Also, students only get a grade if they complete the course they are enrolled in. If a disproportionate fraction of G12 students drop out of mathematics, then the G13 value-added estimator would be biased. Interestingly, there are no students who officially dropped out of mathematics but who completed biology. This could be due to the fact that these courses are compulsory for admission into life sciences specialization fields. When we look at the unconditional drop-out rates in these two courses, it is clear that they are similar, and for both courses relatively low (5% for mathematics and 2% for biology).

G12 students could take fewer courses if they felt less well prepared than G13 students to face university challenges. This is not the case. We can see in Table 1 that G12 students take an average of 5.8 courses over the first year while G13 take 5.7 from the Faculty of Arts and Science. The difference is very small.

Students also select the program they want to attend. G12 students, perhaps knowing that their preparation in mathematics is not as good as G13 students, might have avoided applying to programs involving mathematics. But students do not differ significantly in terms of the program they chose (within the Faculty of Arts and Science). In fact, there is a slightly larger proportion of G13 students who chose a humanities over a life science program than G12, which again supports the hypothesis that G12 students did not try to compensate for their lack of mathematics preparation.[53] Furthermore, application numbers from 2001 suggest that the proportion of applications to Life Sciences (or any other program) did not change, consistent with the view that students did not select into different programs due to the double cohort.

8.2 Identification Issues

In order to make clear what can (and cannot) be identified using the factor-model estimation strategy and the available data, it is useful to step back and consider a more general conceptual framework, making clear the main differences between G12 and G13 students that could affect their university academic performance. One could easily argue that the factors presented in Section 5.1 do not cover all the inputs into the human-capital production function. Aside from the amount of subject-specific human capital (from the curriculum and years of schooling) and ability differences, G12 and G13 students might also differ in terms of general human capital and effort levels, two important ingredients in the human-capital production function. Not being able to fully capture differences across student groups in

[53] Humanities represent 38% of G12, and 41% of G13 student applications. Life Sciences represent 39% of G12, and 34% of G13 student applications.

these two factors may result in underestimating the benefit of Grade 13 mathematics. The next paragraphs explore these possibilities.

8.2.1 General Human Capital

If Grade 13 mathematics gives students general human capital that affects all subjects similarly, then the estimation methods presented in this paper would fail to capture the full extent of the benefit of this extra year. This would be true if, for some reason, the reform affected a student's university biology performance as well as mathematics performance. It is true that a year of mathematics might bring more to students than just math-specific knowledge – for instance, abstract reasoning capacity or better study habits. If so, G13 students would be expected to do better than G12 students in every course. Table 8 suggests otherwise:[54] G12 students do not do significantly worse than G13. They actually do better in a majority of courses (except mathematics). I cannot totally rule out the possibility of such an effect since the higher average ability level of G12 could compensate for the lack of abstract reasoning, for example: the abstract reasoning effect could be confounded with academic ability as defined in this paper.

One sign of such a missing variable could be the presence of group-heteroskedasticity. For example, if all G13 students have abstraction capacity while only some G12 students have, then this capacity should not play a role in the within-group grades variation for G13 grades but should play a role for G12 students. We would expect the two groups to have different variability in the error terms (since the variability in abstraction capacity would be included in the variance of the error term). The null hypothesis of homoskedasticity cannot be rejected using a similar test used for the equality of high school grading rules.[55] Without ruling out the possibility of such a general effect on students, it is hard to find support for such an effect with the data I analyzed. If we think that high school teachers are grading students similarly across groups (and remember that I cannot reject the hypothesis of equal slope or intercept high-school coefficients), it would be even harder to support such a possibility. Note that it is also possible that mathematics involve forms of general human capital that other subject do not, suggesting that my Grade 13 value-added estimator might be capturing more than just subject-specific human capital.

8.2.2 Age

G13 students not only have one more year of mathematics than G12, but they are also (on average) one year older. I could then be facing a similar problem as studies trying to estimate the effect of school start age on test scores. A well known problem with these studies is that we cannot separate the start age effect from the maturity (or age) effect (Angrist and Pischke,

[54] The subjects analyzed in Table 8 are anthropology, biology, chemistry, economics, history, mathematics (for business and life sciences), philosophy, psychology, and sociology.

[55] $\Delta d = 1.93$. The p-value is 0.38.

2009). If the age difference between G13 and G12 students affects biology and mathematics performances differently – which is very likely – all estimators presented so far in this paper will be biased. In order to investigate the potential sign and magnitude of the bias due to age difference, I first looked for age effect within each student group. The estimated age effect (not presented here) from regressing (separately for G12 and G13 students) university performance (e.g. MAT135Y or BIO150Y) on high school average and age (in months) is always small, negative, and statistically insignificant. Age does not play a role in within-group student performance. I also estimated the value-added using G12 and G13 students closer and closer in age (without imposing G12 students to be born in 1985 and G13 students to be born in 1984 to have clearer identification). Whether I looked at G13 and G12 students in the first half of 1985 and second half of 1984, or first quarter of 1985 and last quarter of 1984, the estimated value-added is always small and statistically insignificant.[56] Overall, the age difference between G12 and G13 does not affect the estimated value-added of Grade 13 mathematics.

8.2.3 Effort

Effort may be a factor influencing students' performance. If the amount of effort is the same in both groups (G12 and G13) or if it is constant across courses for the same group then effort should not affect the validity of my results. In the first case, it would not affect the groups' relative performance, while in the second case, the difference in effort level would be captured by the difference in the average ability measure. But students can use effort to compensate for their lack of preparation in mathematics; G12 students might put more effort into studying mathematics than G13 students. If there is an important substitution effect between study time for mathematics and study time for biology, then the effect of Grade 13 would be diluted by the extra effort exerted by G12 students in mathematics, and the estimate of the value-added of G13 math would then be downward-biased. The substitution effect would influence both the mathematics and biology grades. This means that the difference in ability would also be downward-biased (since the performance of G12 students in biology would be negatively affected).

The absence of information about students' study habits makes it impossible to formally test for the presence of effort substitution. But Table 8 does not suggest the presence of such behavior on the part of G12 students. If G12 students substituted effort from biology or chemistry to mathematics, then we would expect to see the difference in performance between the two groups being more important for courses in which students are not expected to take any advanced mathematics. Humanities subjects should favor G12 students more than biology, chemistry, or economics. This is not the case. Anthropology, history, philosophy, and sociology, as a whole, do not favor G12 more than biology, chemistry and economics. Overall, there is no strong evidence that the factor model measure of the value-added of Grade 13

[56] The results are not presented here, but are available upon request.

mathematics is downward-biased.

9 Conclusion

Despite representing a majority of high school students in Canada and United States, we know little about the benefit of secondary schooling for college-bound students. In particular, the literature is silent as to how much subject-specific knowledge (human capital) these students acquire during a year of high school. The 1999 Ontario Secondary School reform provides researchers with a valuable opportunity to shed light on this important issue. The reform allows me to compare the university performance of two groups of students, with one group having one more year of high-school mathematics than the other. As a result, I can directly measure the value-added of an extra year of high-school mathematics for college-bound students.

The results obtained in this study suggest that the benefit to an extra year of mathematics for college-bound students is modest. I find that students coming out of Grade 13 only have a 2.2 point advantage (on a 100 point scale) over students from Grade 12, once I control for ability differences. Furthermore, within-sample investigation and comparison to Krashinsky's (2006) findings point to the presence of heterogeneity in the benefit to an additional year of mathematics across ability levels.

These results have implications for the previous literature. First, the lack of human capital accumulation found in this paper can explain why previous studies (e.g. Altonji 1995) only found modest or no monetary benefits to an extra year of mathematics. Second, the presence of heterogeneity supports the idea that the benefit to an extra year of mathematics could be larger for lower ability students – as suggested by Lang (1993) and Card (1995) for schooling in general.

The results also raise questions for public policy. The finding that high-ability students do not gain much from an extra year of mathematics raises an obvious question: why is there so little value-added? It is possible that high school teachers direct most of their effort toward lower-ability students, leaving high-ability students with fewer resources to acquire additional knowledge. Another possibility is that high-ability students, once in university, can make up for the missing year of mathematics 'effortlessly.' Understanding why high-ability students do not benefit much from an extra year of mathematics can lead to more informed decisions regarding the allocation of (scarce) high school resources. This issue warrants further investigation.

References

[1] ABOWD, J. M., AND D. CARD (1987): "Intertemporal Labor Supply and Long Term Employment Contracts," *American Economic Review*, 77 (1), 50-68.

[2] ABOWD, J. M., AND D. CARD (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57 (2), 411-445.

[3] ALTONJI, J. A. (1995): "The Effects of High School Curriculum on Education and Labor Market Outcomes," *Journal of Human Resources*, 30 (3), 409-438.

[4] ALTONJI, J. G., AND L. M. SEGAL (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business & Economic Statistics*, 14 (3), 353-366.

[5] ALTONJI, J. G., A. P. MARTINS AND A. SIOW. (2002): "Dynamics Factor Models of Consumption, Hours and Income," *Research in Economics*, 56 (1), 3-59.

[6] ANGRIST, J. D. AND A. B. KRUEGER (1991): "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, 106 (4), 979-1014.

[7] ANGRIST, J. D., D. LANG, AND P. OREOPOULOS (2006): "Lead Them to Water and Pay Them to Drink: An Experiment with Services and Incentives for College Achievement," NBER Working Paper No. 12790.

[8] ANGRIST, J. D. AND J.-S. PISCHKE (2009): Mostly Harmless Econometrics: An Empiricist's Companion, Princeton University Press, Princeton.

[9] BAKER, M. (1997): "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," *Journal of Labor Economics*, 15 (2), 338-375.

[10] CANADIAN EDUCATION STATISTICS COUNCIL (2006): "Education Indicators in Canada: Report of the Pan-Canadian Educational Indicators Program 2005," Catalogue no. 81-582-XPE, Statistics Canada.

[11] CARD, D. (1995): "Earnings, Schooling and Ability Revisited," in Solomon W. Polachek, ed., *Research in Labor Economics*, Vol. 14, 23-48.

[12] CHAMBERLAIN, G. (1984): "Panel Data," in Z. Griliches and M. Intrilligator Eds., Handbook of Econometrics, Elsevier.

[13] COUNCIL OF ONTARIO UNIVERSITIES (2002): "The Knowledge Track," September.

[14] DYNARSKI, M., AND P. GLEASON (1998): "How Can We Help? What We Have Learned from Evaluations of Federal Dropout-Prevention Programs," Mathematica Policy Research.

[15] DYNARSKI, M., L. CLARKE, B. COBB, J. FINN, R. RUMBERGER, AND J. SMINK (2008): "Dropout Prevention: A Practice Guide," National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

[16] HARMON, C., AND I. WALKER (1995): "Estimates of the Economic Return to Schooling for the United Kingdom," *American Economic Review*, 49 (5), 1377-1398.

[17] HOFFMANN, F., AND P. OREOPOULOS (2009): "Professor Qualities and Student Achievement," *Review of Economics and Statistics,* 91 (1), 83-92.

[18] HOROWITZ, J. L., (1998): "Bootstrap Methods for Covariance Structures," *Journal of Human Resources*, 33 (1), 39-61.

[19] IMBENS, G. W., AND J. D. ANGRIST (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62 (2), 467-475.

[20] JONES E. B., AND J. D. JACKSON (1990): "College Grades and Labor Market Rewards," *Journal of Human Resources*, 25, 253-266.

[21] JÖRESKOG, K. AND D. SÖRBOM (1996): *LISREL 8: User's Reference Guide* (Lincolnwood, IL: Scientific Software International).

[22] KING, A. J. C., W. WARREN, W. BOYCE, P. CHIN, M. KING, AND J.-C. BOYER (2002): "Double Cohort Study: Phase 2 Report," Social Program Evaluation Group, Queen's University.

[23] KING, A. J. C., W. WARREN, W. BOYCE, P. CHIN, M. KING, J.-C. BOYER, AND B. O'CONNOR (2004): "Double Cohort Study: Phase 3 Report," Social Program Evaluation Group, Queen's University.

[24] KING, A. J. C., W. WARREN, W. BOYCE, P. CHIN, M. KING, J.-C. BOYER, AND B. O'CONNOR (2005): "Double Cohort Study: Phase 4 Report," Social Program Evaluation Group, Queen's University.

[25] KRASHINSKY, C. (2006): "How Would One Extra Year of High School Affect Academic Performance in University? Evidence from a Unique Policy Change," Working Paper.

[26] LANG, K. (1993): "Ability Bias, Discount Rate Bias and the Return to Education," Mimeo. (Department of Economics, Boston University).

[27] LEVINE, P. B., AND D. J. ZIMMERMAN (1995): "The Benefit of Additional High-School Math and Science Classes for Young Men and Women," *Journal of Business & Economic Statistics*, 13 (2), 137-149.

[28] LOURY, L. D., AND D. GARMAN (1995): "College Selectivity and Earnings," *Journal of Labor Economics*, 13, 289-308.

[29] MEGHIR, C., AND M. PALME (2005): "Educational Reform, Ability, and Family Background," *American Economic Review*, 95 (1), 414-424.

[30] ONTARIO MINISTRY OF EDUCATION (1985), "Mathematics, Intermediate and Senior Divisions, Curriculum Guideline. Part 3: Mathematics."

[31] ONTARIO MINISTRY OF EDUCATION (1987), "Science, Intermediate and Senior Divisions, Curriculum Guideline. Part 12: Biology."

[32] ONTARIO MINISTRY OF EDUCATION AND TRAINING (1999a): "Ontario Secondary Schools, Grades 9 to 12: Program and Diploma Requirements."

[33] ONTARIO MINISTRY OF EDUCATION AND TRAINING (1999b): "The Ontario Curriculum Grades 9 and 10: Mathematics."

[34] ONTARIO MINISTRY OF EDUCATION AND TRAINING (2000a): "The Ontario Curriculum Grades 11 and 12: Mathematics."

[35] ONTARIO MINISTRY OF EDUCATION AND TRAINING (2000b): "The Ontario Curriculum Grades 11 and 12: Science."

[36] ONTARIO UNIVERSITIES' APPLICATION CENTRE Website (2006): "Undergraduate Application Statistics," accessed June 1^{st} 2006. http://www.ouac.on.ca/news/uapp/underappstats.html

[37] OREOPOULOS, P. (2006): "Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter," *American Economic Review*, 96 (1), 152-175.

[38] ROSE, H. AND J. R. BETTS (2004): "The Effect of high School Courses on Earnings," Review of Economics and Statistics, 86 (2), 497–513.

[39] SACERDOTE, B. (2001): "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *Quarterly Journal of Economics*, 116 (2), 681-704.

[40] SNYDER, T.D., S.A. DILLOW, AND C.M. HOFMANN (2008): "Digest of Education Statistics 2007," National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

[41] STAIGER, D., AND J. H. STOCK (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65 (3), 557-586.

A Appendix

A.1 Coefficients Identification

Here I simply show one of the different possible strategies. I start by expanding the basic hypotheses of general structural equation models ($\Sigma^{(g)} = \Sigma(\Theta)^{(g)}$, and $\mu^{(g)} = \mu(\Theta)^{(g)}$)

$$\begin{aligned}
var(H_i^C) &= \left[\lambda^{HC}\right]^2 \sigma^2_{\eta^C} + \sigma^2_{\varepsilon^{H,C}} & (23)\\
var(B_i^C) &= \left[\lambda^{B}\right]^2 \sigma^2_{\eta^C} + \sigma^2_{\varepsilon^{B,C}} & (24)\\
var(M_i^C) &= \sigma^2_{\eta^C} + \sigma^2_{\varepsilon^{M,C}} & (25)\\
cov(H_i^C, B_i^C) &= \lambda^{HC}\lambda^{B}\sigma^2_{\eta^C} & (26)\\
cov(H_i^C, M_i^C) &= \lambda^{HC}\sigma^2_{\eta^C} & (27)\\
cov(B_i^C, M_i^C) &= \lambda^{B}\sigma^2_{\eta^C} & (28)\\
E(H_i^C) &= \upsilon^{HC} + \lambda^{HC} E_C\left[\eta_i\right] & (29)\\
E(B_i^C) &= \upsilon^{B} + \lambda^{B} E_C\left[\eta_i\right] & (30)\\
E(M_i^{G13}) &= \Delta_V + E_{G13}\left[\eta_i\right] & (31)\\
E(M_i^{G12}) &= E_{G12}\left[\eta_i\right] & (32)
\end{aligned}$$

The identification of the average academic ability for the G12 students ($E_{G12}\left[\eta_i\right]$) is trivial from the normalization I made. Next, we can isolate λ^B using (26) and (27)

$$\lambda^B = \frac{cov(H_i^C, B_i^C)}{cov(H_i^C, M_i^C)} \tag{33}$$

Dividing (26) by (28) we get

$$\lambda^{HC} = \frac{cov(H_i^C, B_i^C)}{cov(B_i^C, M_i^C)} \tag{34}$$

Equations (33), (32), and (30) give us an expression for the constant term of the biology grading rule

$$\begin{aligned}
\upsilon^B &= E(B_i^{G12}) - \lambda^B E_{G12}\left[\eta_i\right]\\
&= E(B_i^{G12}) - E(M_i^{G12})\frac{cov(H_i^C, B_i^C)}{cov(H_i^C, M_i^C)} & (35)
\end{aligned}$$

Plugging (35) and (33) in (30) will give us a measure of the average academic ability of G13 students

$$\begin{aligned}
E_{G13}\left[\eta_i\right] &= \frac{E(B_i^{G13}) - \upsilon^B}{\lambda^B}\\
&= \frac{cov(H_i^C, M_i^C)}{cov(H_i^C, B_i^C)}\left[E(B_i^{G13}) - E(B_i^{G12})\right] + E(M_i^{G12}) & (36)
\end{aligned}$$

Having isolated the average academic ability of G13 students, I am able to identify the value-added

$$\begin{aligned}\Delta_V &= E(M_i^{G13}) - E_{G13}[\eta_i] \\ &= E(M_i^{G13}) - E(M_i^{G12}) - \frac{cov(H_i^C, M_i^C)}{cov(H_i^C, B_i^C)}\left[E(B_i^{G13}) - E(B_i^{G12})\right] \end{aligned} \quad (37)$$

The constant term of G13 students high school grading rule can be found using (36), (34), and (29)

$$\begin{aligned}\upsilon^{HG13} &= E(H_i^{G13}) - \lambda^{HG13}E_{G13}[\eta_i] \\ &= E(H_i^{G13}) \\ &\quad -\frac{cov(H_i^{G13}, B_i^{G13})}{cov(B_i^{G13}, M_i^{G13})}\left\{E(M_i^{G12}) + \frac{cov(H_i^C, M_i^C)}{cov(H_i^C, B_i^C)}\left[E(B_i^{G13}) - E(B_i^{G12})\right]\right\}\end{aligned}$$

while the same constant for G12 students looks like

$$\begin{aligned}\upsilon^{HG12} &= E(H_i^{G12}) - \lambda^{HG12}E_{G12}[\eta_i] \\ &= E(H_i^{G12}) - \frac{cov(H_i^{G12}, B_i^{G12})}{cov(B_i^{G12}, M_i^{G12})}E(M_i^{G12})\end{aligned}$$

A.2 OLS Estimation Bias

This section shows conditions under which the OLS regression estimation of mathematics grades on biology grades gives a biased estimator of Δ_V.

Assume that the relationship between ability and university performance is characterized by equations (12) and (13). For notational simplicity, also assume that the error terms for G12 and G13 students are drawn from the same distribution (e.g. $\varepsilon_i^{M,G13} = \varepsilon_i^{M,G12} = \varepsilon_i^M$), and normalize λ^M to unity. We can then drop the c superscript on M_i and B_i and combine equations (12) and (13) to rewrite equation (2)

$$M_i = \alpha + \Delta_V \times G13_i + \beta B_i + u_i$$

where $\alpha = -\upsilon^B/\lambda^B$, $\beta = 1/\lambda^B$, and $u_i = \varepsilon_i^M - \varepsilon_i^B/\lambda^B$. I now transform the data in deviations from mean and work with matrix notation for clarity reasons

$$\tilde{M}_i = \Delta_V \times \tilde{G}13_i + \beta\tilde{B}_i + u_i$$

or

$$\tilde{\mathbf{M}} = \tilde{\mathbf{X}}\boldsymbol{\beta} + \tilde{\mathbf{u}}$$

where $\tilde{M}$ is a vector of demeaned mathematics grades (i.e. $\tilde{M} = M - \bar{M}$), and $\tilde{X}$ is the matrix of the demeaned $\tilde{G}13_i$ and $\tilde{B}_i$ variables, while

$$\boldsymbol{\beta} = \begin{bmatrix} \Delta_V \\ \beta \end{bmatrix}$$

The OLS estimator of β is

$$\hat{\boldsymbol{\beta}} = (\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{M}}$$

Hence, $plim\ \hat{\beta}$ is

$$plim\ \hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + plim(\frac{\tilde{\mathbf{X}}'\tilde{\mathbf{X}}}{N})^{-1} plim \left(\frac{\tilde{\mathbf{X}}'\tilde{\mathbf{u}}}{N}\right)$$

The dummy variable indicating if a student is a member of the G13 group is not correlated with the error – the first term of $plim \left(\frac{\tilde{\mathbf{X}}'\tilde{\mathbf{u}}}{N}\right)$ will be 0. But since $\tilde{B}_i$ is correlated with the error term (because of measurement error) the second term of the vector will be non-zero. This gives

$$plim \left(\frac{\tilde{\mathbf{X}}'\tilde{\mathbf{u}}}{N}\right) = \begin{bmatrix} 0 \\ -\frac{var(\varepsilon_i^B)}{\lambda^B} \end{bmatrix}$$

Unless we are facing a $plim(\frac{\tilde{\mathbf{X}}'\tilde{\mathbf{X}}}{N})^{-1}$ with 0 off-diagonal elements (e.g. the regressors are uncorrelated) we should get a biased estimator of the treatment effect.

$$plim(\frac{\tilde{\mathbf{X}}'\tilde{\mathbf{X}}}{N})^{-1} = plim \begin{bmatrix} \frac{N_{G12}N_{G13}}{N^2} & \sum\limits_{i \in G13} \frac{\tilde{B}_i}{N} \\ \sum\limits_{i \in G13} \frac{\tilde{B}_i}{N} & \sum\limits_{i=1}^{N} \frac{\tilde{B}_i^2}{N} \end{bmatrix}^{-1}$$

Let

$$\begin{aligned} plim(\frac{\tilde{\mathbf{X}}'\tilde{\mathbf{X}}}{N})^{-1} &\equiv plim\,(\mathbf{A})^{-1} \\ &= plim \frac{1}{\det \mathbf{A}} plim \begin{bmatrix} \sum\limits_{i=1}^{N} \frac{\tilde{B}_i^2}{N} & -\sum\limits_{i \in G13} \frac{\tilde{B}_i}{N} \\ -\sum\limits_{i \in G13} \frac{\tilde{B}_i}{N} & \frac{N_{G12}N_{G13}}{N^2} \end{bmatrix} \\ &\equiv plim \frac{1}{\det \mathbf{A}} plim \begin{bmatrix} B & -C \\ -C & D \end{bmatrix} \end{aligned}$$

Hence

$$plim\ \hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = plim \frac{1}{\det \mathbf{A}} plim \begin{bmatrix} B & -C \\ -C & D \end{bmatrix} \begin{bmatrix} 0 \\ -\frac{var(\varepsilon_i^B)}{\lambda^B} \end{bmatrix}$$

As long as $C \neq 0$, the OLS estimator of Δ_V will be biased. C could be equal to 0 if the average biology grades are equal across groups which could happen if both groups had the same level of academic ability. Then the $G13_i$ dummy would be uncorrelated with biology grades. The sign of the bias will depend on whether the average biology grade is higher for

G13, or for G12 students. If the average is higher for the latter group, we should expect a downward bias.

A.3 The Grade 12 and Grade 13 Mathematics Curricula

In this appendix, I highlight differences and similarities between the Grade 12 and Grade 13 mathematics curricula by looking at their specific strands (topics covered in class). Table A-1 lists the strands by grade level as described in the Ontario Ministry of Education official documents (see Ontario Ministry of Education (1985) and Ontario Ministry of Education and Training (1999b, 2000a)). Course codes are presented in parentheses.

An inspection of Table A-1 shows that the most important difference between the two curricula is with regards to calculus – both in terms of content and time spent on the subject matter. Integral calculus (antidifferentiation) is absent from the Grade 12 curriculum. In addition, Grade 13 students had a full year devoted to calculus, whereas Grade 12 split their last year of high school studying calculus and advanced functions (e.g. logarithmic functions), a topic covered in grade 12 (in the "Algebraic Operations" and "Relations and Functions") for Grade 13 students.

A detailed examination of official documents reveals important compression within strands, but mainly for upper years and for calculus related topics. For example, the "Limits and Derivatives" and "Underlying Concepts of Calculus" strands may sound like similar strands, but more material is covered in the former. While "Limits and Derivatives" covers limits, derivatives, fundamental properties of derivatives, derivatives of trigonometric functions, and derivatives of exponential and logarithmic functions, "Underlying Concepts of Calculus" only covers a subset of these topics (i.e. limits, and graphical definition of the derivative). Finally, the "Derivatives and Applications" strand covers topics that used to be covered in "Limits and Derivatives" (e.g. derivatives, fundamental properties of derivatives, and derivatives of exponential and logarithmic functions) and in "Applications of Derivatives" (e.g. curve sketching, solving for rates of change and extreme values).

Changes in earlier grades are less important. Despite having been aggregated in different strands, the material covered in grade 9 is similar across curricula. Only a few topics previously covered in grade 9 are now covered in grade 10. For example, the study of slope of a line, previously under grade 10 "Geometry", is now under grade 9 "Analytic Geometry", and "distinguishing the exact and approximate representation of the same quantity", previously under grade 10 "Numerical Methods", is now under grade 9 "Numerical Sense and Algebra."[57] The properties of line segments ("Analytic Geometry") are now introduced in 10 instead of grade 11. "Investment Mathematics" was renamed "Financial Applications of Sequences and Series." Finally, trigonometry and trigonometric functions are studied starting

[57] Another minor difference is that some activities like working with transparent mirrors, compasses, and protractors seem to have been abandoned for activities using computers.

in grade 10 instead of grade 12 (under "Geometry" and "Relations and Functions"). "Geometry" and "Relations and Functions" were split into "Trigonometric Functions", "Tools for Operating and Communicating with Functions" and "Investigations of Loci and Conics."

Table 1: Descriptive Statistics of 2003 Entering Life Sciences Students

A. G12 (N=502)	Mean	Std. Dev.	Min	Max
Age	18.2	0.3	17.8	18.7
Female	0.64	0.5	0	1
HS Average	90.8	3.4	83.0	98.8
Number of Courses	5.8	0.6	3	8
BIO150Y Grade	75.8	10.7	43	97
MAT135Y Grade	70.2	13.3	30	99
B. G13 (N=436)	**Mean**	**Std. Dev.**	**Min**	**Max**
Age	19.2	0.3	18.8	19.7
Female	0.67	0.5	0	1
HS Average	90.9	3.2	83.7	99.2
Number of Courses	5.7	0.5	4	7
BIO150Y Grade	74.3	10.7	41	95
MAT135Y Grade	70.6	13.0	30	98
C. G13-G12 Difference	**Mean**	**Std Err.**		
Age	1.0	0.02		
Female	0.03	0.03		
HS Average	0.01	0.22		
Number of Courses	-0.1	0.04		
BIO150Y Grade	-1.5	0.70		
MAT135Y Grade	0.4	0.86		

Table 2: OLS Regression Results

Dependent Variable:	Proxy for Ability	
University Math. Grade	(I)	(II)
	Biology	High School
β	0.84	2.05
	(0.03)	(0.11)
Δ_V	**1.68**	**0.30**
	(0.63)	**(0.74)**
α	6.40	-115.9
	(2.25)	(9.66)
R^2	0.47	0.27
N	938	938

Robust standard errors are in parentheses.

Table 3: Difference-in-Differences Estimator

Math - Bio	Mean	Standard Errors
G13 (n=436)	-3.70	0.48
G12 (n=502)	-5.62	0.42
G13-G12$\equiv \Delta_{DD}$	1.92	0.64

Table 4: Means and Covariances of Students' Grades

G13 (n=436)	Mean	High School	Biology	Mathematics
High School	90.86	10.1		
	(0.2)	(0.6)		
Biology	74.31	16.7	114.1	
	(0.5)	(1.6)	(7.6)	
Mathematics	70.61	20.5	91.5	167.8
	(0.6)	(2.1)	(7.5)	(11.3)
G12 (n=502)		High School	Biology	Mathematics
High School	90.79	11.8		
	(0.2)	(0.6)		
Biology	75.79	20.6	115.3	
	(0.5)	(1.7)	(6.3)	
Mathematics	70.16	24.3	101.0	176.3
	(0.6)	(2.1)	(6.9)	(10.1)

Note: Covariances and variances are presented in the last three columns. Standard errors are in parentheses.

Table 5: Parameter Estimates (GMM Estimation)

	G13	G12
υ (intercept)		
MATH	2.21	0.00
	(0.09)	(-)
	[0.17]	[-]
BIO	17.21	17.21
	(2.76)	(2.76)
	[3.64]	[3.64]
HS	78.23	76.55
	(1.15)	(0.99)
	[1.11]	[1.02]
λ (slope)		
MATH	1.00	1.00
	(-)	(-)
BIO	0.83	0.83
	(0.04)	(0.04)
	[0.05]	[0.05]
HS	0.18	0.20
	(0.02)	(0.01)
	[0.02]	[0.01]
$\bar{\eta}$ (ability)	68.39	70.16
	(0.08)	(0.01)
	[0.63]	[0.61]

OMD standard errors are in parentheses, while bootstrap standard errors are in square brackets.

Table 6: Testing for the Equality of High School Grading Policies

	d	df	p-value	Δd	df	p-value
Baseline model*	0.18	1	0.67	-	-	-
Imposing:						
Same slope coefficients	0.92	2	0.63	0.74	1	0.39
Same intercept coefficients	1.48	2	0.48	1.30	1	0.25
Same slope & intercept coefficients	177.5	3	0.00	177.3	2	0.00

* The baseline model is the model used to present the results in Table 5.

Table 7: Parameter Estimates by Gender

	Females		Males	
	G13	G12	G13	G12
υ (intercept)				
MATH	2.43	0.00	1.97	0.00
	(0.08)	(-)	(0.20)	(-)
	[0.18]	[-]	[0.37]	[-]
BIO	15.85	15.85	19.36	19.36
	(3.21)	(3.21)	(5.15)	(5.15)
	[4.59]	[4.59]	[6.46]	[6.46]
HS	78.38	77.02	76.98	75.18
	(1.31)	(1.25)	(2.45)	(1.65)
	[1.34]	[1.28]	[1.92]	[1.79]
λ (slope)				
MATH	1.00	1.00	1.00	1.00
	(-)	(-)	(-)	(-)
BIO	0.86	0.86	0.79	0.79
	(0.05)	(0.05)	(0.07)	(0.07)
	[0.07]	[0.07]	[0.09]	[0.09]
HS	0.19	0.20	0.19	0.22
	(0.02)	(0.02)	(0.03)	(0.02)
	[0.02]	[0.02]	[0.03]	[0.02]
$\bar{\eta}$ (ability)	66.97	68.42	71.04	73.31
	(0.08)	(0.01)	(0.20)	(0.01)
	[0.72]	[0.71]	[1.03]	[1.00]
N	291	323	145	179

OMD standard errors are in parentheses, while bootstrap standard errors are in square brackets.

Table 8: Students Average Marks in 2003

	# Obs.	Marks		H_0 : G13 = G12
		G13	G12	p-values
ANT100	501	**66.8**	65.7	0.34
BIO150	1161	73.2	**74.4**	0.08
CHM138	1016	76.2	**76.8**	0.47
ECO100	629	68.1	**68.2**	0.91
HIS109	293	67.2	**69.0**	0.27
MAT133	281	**69.8**	67.3	0.14
MAT135	1092	**68.5**	68.5	0.97
PHL100	448	71.4	**72.0**	0.49
PSY100	883	70.0	**70.3**	0.73
SOC101	791	**65.5**	64.7	0.30

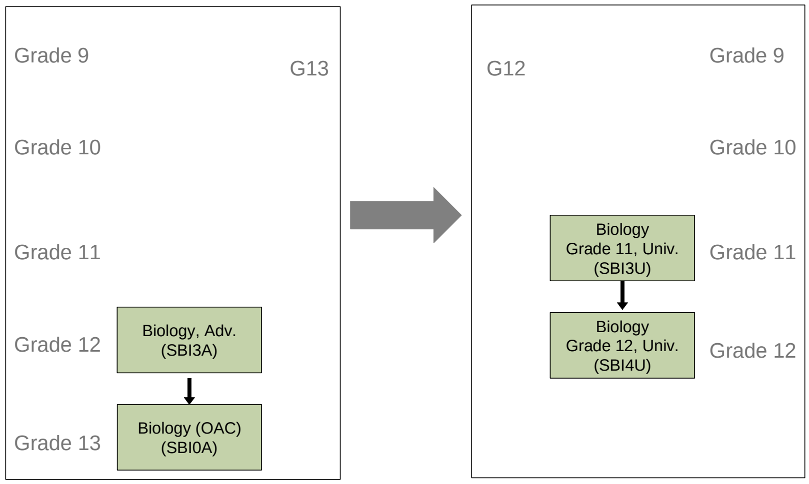

Figure 1: Reform of the Biology Sequence for University-Bound Students

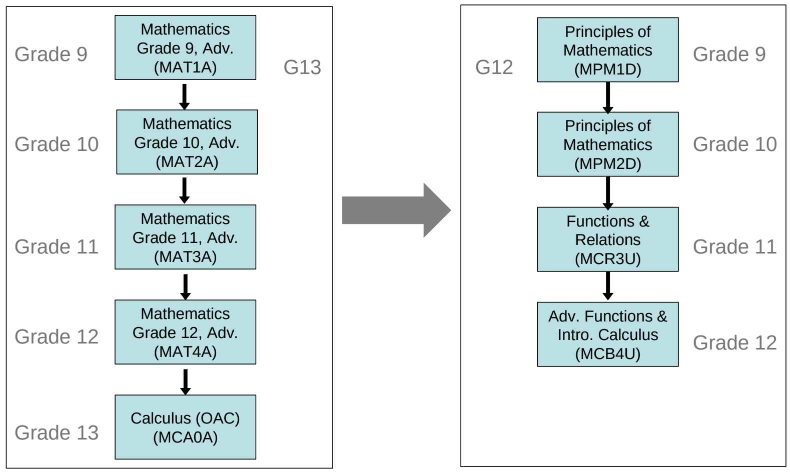

Figure 2: Reform of the Mathematics Sequence for University-Bound Students

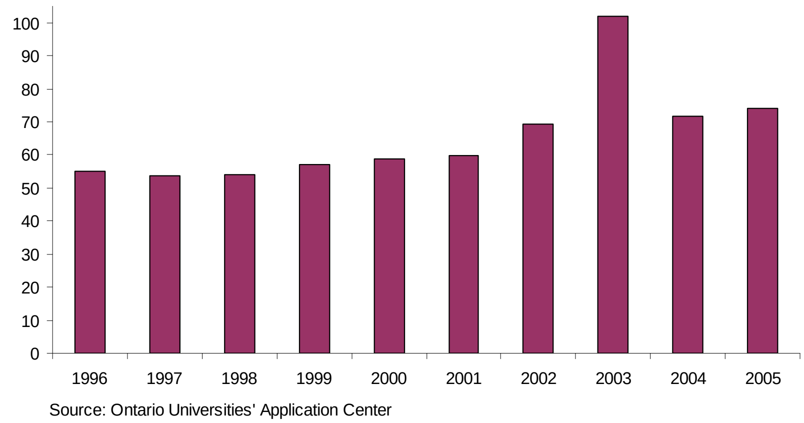

Figure 3: Number of Ontario University Applicants (in thousands)

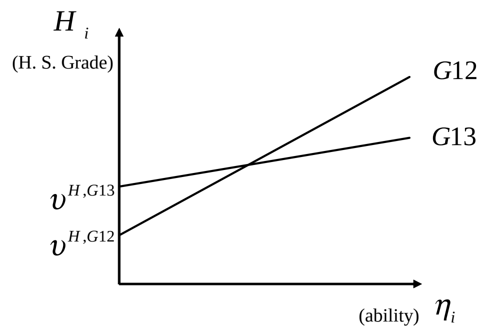

Figure 4: Possible High School Grading Policies

Table A-1: Specific Strands of the Grade 13 and Grade 12 Programs

	Grade 13		Grade 12
Grade 9 (MAT1A)	Numerical Methods	Grade 9 (MPM1D)	Number Sense and Algebra
	Algebra		Relationships
	Measurement and Geometry		Analytic Geometry
			Measurement and Geometry
Grade 10 (MAT2A)	Numerical Methods	Grade 10 (MPM2D)	Quadratic Functions
	Algebra		Analytic Geometry
	Geometry		Trigonometry
Grade 11 (MAT3A)	Algebraic Operations	Grade 11 (MCR3U)	Financial Applications of Sequences and Series
	Analytic Geometry		Trigonometric Functions
	Functions and Transformations		Tools for Operating and Communicating with Functions
	Investment Mathematics		Investigations of Loci and Conics
Grade 12 (MAT4A)	Geometry	Grade 12 (MCB3U)	Advanced Functions
	Relations and Functions		Underlying Concepts of Calculus
	Algebraic Operations		Derivatives and Applications
Grade 13 (OAC)	Limits and Derivatives		
	Applications of Derivatives		
	Antidifferentiation		

Sources: Ontario Ministry of Education (1985) and Ontario Ministry of Education and Training (1999b, 2000a).