Estimating the Benefit of High School for University-Bound Students:

Evidence of Subject-Specific Human Capital Accumulation*

Louis-Philippe Morin†

University of Ottawa

Abstract

Numerous studies suggest that the value of high school education is large for potential dropouts, yet we know much less about the benefits for students who will go on to post-secondary education. To help fill this gap, I measure the value-added (in terms of university grades) of an extra year of high-school mathematics for university-bound students using a recent Ontario secondary school reform. The estimated value-added is small for these students—2.3 points on a 100 point scale. This evidence helps to explain why the literature has found only modest effects of taking more mathematics in high school on wages.

Keywords: Human Capital Accumulation, High School Curriculum, Education Reform, Mathematics, Factor Model.


*I would like to thank Victor Aguirregabiria, Joshua Angrist, Michael Baker, Pierre Brochu, Mark Evans, Fernanda Estevan, Xavier Fazio, Rob McMillan, Phil Oreopoulos, Silvana Simone, Aloysius Siow and Jeffrey Smith 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.

†Contact info: Louis-Philippe Morin, Department of Economics, University of Ottawa, Ottawa, Ontario, Canada, K1N 6N5. Telephone: (613) 562-5800 ext. 4867. Email: lmorin@uottawa.ca.
1 Introduction

Following Angrist and Krueger’s landmark 1991 study, a number of economists have found that the benefit to an additional year of secondary schooling is large for potential school leavers (dropouts)\(^1\) in turn prompting a host of programs intended to reduce high-school dropout rates. Despite the understandable policy interest in dropouts, university-bound students now represent a majority of students in North America as well as a significant fraction in many European countries\(^2\). Yet, we know much less about the benefits of secondary schooling for these higher-ability students—whether the overall gains from attending high school are lower for university-bound students, and what the effects of specific curricula are for this numerically and economically significant group.

Quantifying the benefits of high school is relevant to an important policy question concerning the way that a high school education should be delivered: How many years should it take, and what specific curricula should be taught? Clearly, if high school could be shortened and existing curricula cut without significant adverse consequences for student learning, then the possibility arises that substantial cost savings could be achieved through curriculum reform.

This high school design issue was prominent in shaping what arguably became one of the most radical and controversial educational reforms in Canada over the past few decades. 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); and as a consequence of the reform, the first cohort of students graduating from Grade 12 and the last cohort of Grade-13 graduates entered Ontario universities simultaneously in 2003—the so-called

\(^1\)See for example, Harmon and Walker (1995), Staiger and Stock (1997), Meghir and Palme (2005), and Oreopoulos (2006). As these studies 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, the estimates presented in these papers are interpreted as the benefit of an extra year of high school for potential school leavers.

\(^2\)In the United States, for example, the average 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 university 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. In the UK, the Higher Education Initial Participation Rate (HEIPR)—an approximate probability that a 17 year old participates in higher education by age 30—was 40.3 percent in 2003-2004. Sources: Tables C5.2 and E1.1 in Canadian Education Statistics Council (2006), Tables 102 and 193 in Snyder et al. (2008), and Department for Education and Skills (2007).
'double cohort’—affording a unique and useful comparison that helps shed light on the benefits of high school for university-bound students.

At first blush, the reform might suggest the ‘reverse’ of the typical compulsory schooling law, with the Grade 12 and Grade 13 cohorts differing solely in terms of years of high school taken. However, two important features of the reform point to a more complex treatment. First, there was scope for differential selection: although students could not choose the program they would graduate from (i.e. Grade 12 or Grade 13), they could decide when to apply to and enroll in university. Knowing that 2003 would be a more competitive year for university admissions, some students from both Grade 12 and Grade 13 delayed their university applications by a year, while some Grade 13 students ‘fast-tracked’ high school, graduating a year early to avoid the double cohort. As consequence, we cannot assume that the Grade 12 and Grade 13 students who entered university in 2003 have the same levels of ability. Importantly, this type of selection cannot be dealt with using age as an instrument for a student’s secondary school program. Hence, one has to use a different strategy in order to learn from comparisons of these two groups of students.

The second crucial feature of the reform is that the change in high school curriculum was subject-specific: that is, some subjects were drastically affected while others were not. In particular, the length of the high school mathematics curriculum for university-bound students went from five years to four, affecting both the amount of material covered and the time spent on some topics, while the length of the biology curriculum for the same students remained unchanged at two years. Hence, rather than involving one more year of secondary education, the treatment consisted of one more year of subject-specific courses.

From an estimation perspective, the non-uniform changes in curriculum mean that one cannot estimate the benefit of an extra year of schooling from this reform. But this same non-uniformity can be used to estimate the benefit of a year of subject-specific education, even in the presence of self-selection. Particularly, one can use the non-uniform changes in curriculum to identify the value-added to university performance of Grade 13 mathematics: 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.

\[3\]

\[3\] Even in the absence of any ‘maturity’ effects, the age of the student is likely to be correlated with her academic ability given that ‘lower’ ability Grade 12 students, for example, might have selected out of the double cohort. See Section 4 for more details.
The primary contribution of the paper is to develop an estimation strategy and present new estimates shedding light on the subject-specific benefits of high school for a large group of students, about whom the existing literature has very little to say. To estimate the value-added of Grade 13 mathematics, I construct a flexible factor model that has four 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) due to the type of selection just mentioned. Second, it allows for the value-added to vary with student ability. Third, the identified value-added estimator from the factor model can be seen as a generalization of the standard difference-in-differences estimator, allowing (unlike the difference-in-differences estimator) for the possibility that subjects do not measure ability in the same way. Fourth, the framework makes it possible to test for other potentially important effects of the reform, such as the presence of high school grade inflation.

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 the same first-year compulsory courses, one of them being a university mathematics course.

The main empirical 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 2.3 point advantage (on a 100 point scale) over students from Grade 12, representing 17 percent of a standard deviation ($\sigma$) in mathematics performance. My within-sample investigation also suggests that, if anything, 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), who examines the impact of the same reform but on university-bound students with significantly lower high-school averages, further indicates that there is substantial heterogeneity in the benefit to an additional year of high school mathematics. The estimated effect of Grade 13 found in my paper ($0.17\sigma$) is far below the $0.5\sigma$ to $1.2\sigma$ range found in Krashinsky (2006).

In terms of its broader significance, this paper speaks to an important literature examining the impact of high school curriculum on wages. Several studies have focused on this linkage, primarily to shed light on the “human capital/screening” debate: if schooling serves mainly as a
screening device, we would expect the return to a year of education to be larger than the return to a year-equivalent of courses. Altonji (1995), Levine and Zimmerman (1995), and Rose and Betts (2004) have all looked at the impact of high school curriculum on wages, with a special emphasis on mathematics. Their instrumental-variable (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 university-bound students do not gain much from extra mathematics\(^4\). 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. Yet 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 (curriculum) and general (schooling) human capital differently.

This paper helps to 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 university-bound students acquire directly. 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 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. 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). As a further implication, within- and between-sample comparisons point to the presence of 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.

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 of students from the new curriculum graduated from high school, and in the same year, Grade 13 was also abolished. Thus, in 2003, Ontario universities 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 Ontario Secondary School Diploma requirements. We might think that students from the two curricula (G12 and G13) learned the same material. But university-bound students also need to satisfy university 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 life-science students. 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.

5 The Ontario Ministry of Education and Training (1999) 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.”
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 (Ontario Ministry of Education 1987, Ontario Ministry of Education and Training 2000b). Conversations with professors at the Ontario Institute for Studies in Education at the University of Toronto, and a 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 the 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. The reform clearly affected the sequence of mathematics 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 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 (Ontario Ministry of Education 1985, Ontario Ministry of Education and Training 2000a). 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.
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 performance (e.g. grades, dropped courses, program\(^6\)), and pre-admission academic history (e.g. high school average, attended secondary-school institutions, and an indicator of the secondary-school curriculum the student 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.\(^7\) Since I have an indicator of the secondary school curriculum attended by each student (G12/G13 indicator), I 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 me to concentrate on the population I am 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 of biology should not affect a student’s knowledge of mathematics and vice versa.\(^8\)

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). Close to 90 percent of students for whom I observe a grade for BIO150Y also had a grade for MAT135Y.\(^9\)

\(^6\)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.

\(^7\)For related studies using university administrative data, see Sacerdote (2001) using Dartmouth College data, and Angrist et al. (2009), and Lindo et al. (2010) using data from a large Canadian university.

\(^8\)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.

\(^9\)This is true for both groups of students: 88.1 percent of G12 students and 89.2 percent of G13 students (the difference not being statistically significant at any conventional confidence levels) for whom I observe a grade for BIO150Y also had a grade for MAT135Y. It is not surprising to observe such a high proportion of students taking
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. Hence, Life Sciences G13 students have one more year of high school mathematics than their G12 classmates. Students should also have a senior high school biology credit. Hence, students are expected to have completed both course sequences of their respective curriculum, as described in Section 2.

Table 1 presents descriptive statistics for these double cohort students who completed both BIO150Y and MAT135Y with grades greater or equal to 30 percent. I assume that students with grades below 30 percent (11 students in total) dropped the course. This restriction does not affect the results (see Section 8.1). 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, the two groups of students seem very similar: they take roughly the same number of university courses 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 Mathematics

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 mathematics, motivating the need for a more flexible estimator which will be presented in the next section.

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 ($\Delta_M$) could then be characterized by the sum of the value-added of

---

10 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.

11 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.
Table 1: Descriptive Statistics of 2003 Life Sciences Students

<table>
<thead>
<tr>
<th>A. G12 (n=502)</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>18.2</td>
<td>0.3</td>
<td>17.8</td>
<td>18.7</td>
</tr>
<tr>
<td>Female</td>
<td>0.64</td>
<td>0.5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>HS Average</td>
<td>90.8</td>
<td>3.4</td>
<td>83.0</td>
<td>98.8</td>
</tr>
<tr>
<td>Number of Courses</td>
<td>5.8</td>
<td>0.6</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>BIO150Y Grade</td>
<td>75.8</td>
<td>10.7</td>
<td>43</td>
<td>97</td>
</tr>
<tr>
<td>MAT135Y Grade</td>
<td>70.2</td>
<td>13.3</td>
<td>30</td>
<td>99</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. G13 (n=436)</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>19.2</td>
<td>0.3</td>
<td>18.8</td>
<td>19.7</td>
</tr>
<tr>
<td>Female</td>
<td>0.67</td>
<td>0.5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>HS Average</td>
<td>90.9</td>
<td>3.2</td>
<td>83.7</td>
<td>99.2</td>
</tr>
<tr>
<td>Number of Courses</td>
<td>5.7</td>
<td>0.5</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>BIO150Y Grade</td>
<td>74.3</td>
<td>10.7</td>
<td>41</td>
<td>95</td>
</tr>
<tr>
<td>MAT135Y Grade</td>
<td>70.6</td>
<td>13.0</td>
<td>30</td>
<td>98</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C. G13-G12 Difference</th>
<th>Mean</th>
<th>Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.0</td>
<td>0.02</td>
</tr>
<tr>
<td>Female</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>HS Average</td>
<td>0.01</td>
<td>0.22</td>
</tr>
<tr>
<td>Number of Courses</td>
<td>-0.10</td>
<td>0.04</td>
</tr>
<tr>
<td>BIO150Y Grade</td>
<td>-1.48</td>
<td>0.70</td>
</tr>
<tr>
<td>MAT135Y Grade</td>
<td>0.44</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Notes. This table presents the descriptive statistics of 2003 first-year Life Science students who enrolled in BIO150Y and MAT135Y and received a grade equal or greater than 30 percent. I assumed that students with grades below 30 percent dropped the course. 12 students were dropped from the sample for this reason. The sample size is 938. Panels A and B present descriptive statistics for G12 and G13 students, respectively. Panel C compares the descriptive statistics (and their associated standard errors) across G12 and G13 students.

G13 (Δ_V) and the difference in average initial level of academic ability between G12 and G13 (Δ_η). Thus,

\[ \Delta_M = \Delta_V + \Delta_\eta. \] (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 (Δ_η ≃ 0). Then the difference in mathematics performance would fully capture the effect of the reform (i.e. Δ_M = Δ_V).

The main challenge in estimating the value-added of Grade 13 mathematics is that the reform cannot be thought as a clean random experiment and that G12 and G13 students might have

\[^{12}\text{The initial level of ability is the general level of academic ability acquired prior to secondary schooling.}\]
different average levels of academic ability. In particular, if G12 students have a higher average level of ability than G13 students (i.e. $\Delta_\eta < 0$), 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. 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 with the fear of the double cohort probably encouraging some G13 students to try to fast-track and graduate in 2002 instead of 2003.

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 annual increase between 1996 to 2001 (about 1.6%), suggesting that some G13 students successfully escaped from the double cohort. 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 of 2003 G12 students. Furthermore, the number of applicants in 2004 (72,000) also seems larger than expected, 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. For the same reason, an instrumental variable estimation strategy based on student age would not give a consistent estimator of the value-added. The age of the student would be correlated with the student curriculum, but it would also be correlated with her academic ability.

Observing a university outcome not affected by the reform can help controlling for any difference in ability, but it also introduces measurement issues. For example, 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

---

13 Source: Ontario Universities’ Application Centre (http://www.ouac.on.ca/statistics/archive/).
14 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 (King et al. 2002).
15 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.
university mathematics grade on the university biology grade and a dummy variable \(G_{13,i}\), equal to 1 for G13 students and 0 for G12 students. A potential problem with the interpretation of the OLS estimate is that it assumes that biology measures ability perfectly. If not, and if the average biology grade is larger for G12 students—as is the case here—, the OLS estimator for the value-added of Grade 13 will be downward biased (Morin 2010).

Alternatively, if we assume that mathematics and biology measure ability the same way (up to a constant), we can then construct a difference-in-differences estimator to identify the value-added of G13:

\[
\Delta_{DD} \equiv \Delta_M - \Delta_B = \Delta_V
\]

The difference between differences in average university mathematics grades \((\Delta_M)\) and in average biology grades \((\Delta_B)\) would give us the value-added of Grade 13 \((\Delta_V)\). The difference in average biology performance presented in Table I 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.

The main problem with the difference-in-differences estimator is that it 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

\[
\Delta_B = \rho_B \Delta_\eta
\]

where \(\rho_B \neq 1\). Then

\[
\Delta_{DD} = \Delta_M - \rho_B \Delta_\eta = \Delta_V + \frac{(1 - \rho_B)}{\rho_B} \Delta_B
\]

where equation 5 is obtained using equations 1 and 3. Hence, if both \(\Delta_B \neq 0\) and \(\rho_B \neq 1\) then the difference-in-differences estimator will also be biased. Table I shows that \(\Delta_B < 0\); 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 compare the covariances of biology and
mathematics grades, 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.

<table>
<thead>
<tr>
<th>Table 2: Means and Covariances of Students’ Grades</th>
</tr>
</thead>
<tbody>
<tr>
<td>G13 (n=436)</td>
</tr>
<tr>
<td>High School</td>
</tr>
<tr>
<td>Biology</td>
</tr>
<tr>
<td>Mathematics</td>
</tr>
<tr>
<td>G12 (n=502)</td>
</tr>
<tr>
<td>High School</td>
</tr>
<tr>
<td>Biology</td>
</tr>
<tr>
<td>Mathematics</td>
</tr>
</tbody>
</table>

Notes. The first column presents the students’ mean performance in High School, Biology (BIO150Y), and Mathematics (MAT135Y). These were computed separately for G12 and G13 students. Covariances and variances of the performance measures are presented in the last three columns.

In sum, student self-selection and subject heterogeneity in ability measurement could lead standard estimators (e.g. the difference-in-differences estimator) to underestimate the value-added of Grade 13 mathematics. I now present a model that can account for these two issues.

5 A Grading-Rule Model

In this section, I propose a simple grading-rule model that can account for the relationship between subject-specific human capital accumulation and academic performance in the specific environment of the double cohort. In particular, the model shows how available observables—university grades, high school grades, and high school curricula—can be linked. A nice feature of the proposed grading rule model is that its 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 across subjects. It also allows for potential grade inflation at the high school level and different levels of average ability
across groups. Furthermore, it allows me to test whether higher-ability students benefited more from G13 mathematics than their lower-ability counterparts. 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.¹⁶

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 be taught in 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.

¹⁶One 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 and provide robustness evidence.
5.2 Model

I now present the model more formally, first specifying the way that subject-specific human capital is accumulated through high school and university, then specifying how this human capital is measured by a grade.

5.2.1 Human Capital Accumulation

There are two institutions, superscripted by an uppercase $I = \{H, U\}$, at which students accumulate subject-specific human capital: high school ($I = H$) and university ($I = U$). Assume only two subjects $S = \{B, M\}$, say biology and mathematics, that are taken both at high school and university. Finally, there are two groups of students who take different curricula $C = \{G_{12}, G_{13}\}$ 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 ($\eta_i$) and then accumulates subject-specific human capital which depends on the amount of material ($\tau_{I,S,C}$) presented to her as she attends high school and university. While in high school, $G_{13}$ students receive a treatment which affects the amount of mathematics-specific human capital they acquire ($G_{12}$ students do not receive the treatment). High school biology is not affected by the treatment, and hence both groups are assumed to being taught the same biology material in high school.

The subject-specific human capital accumulation will depend on the amount of material presented to the student and on her ability, for both mathematics and biology. More specifically the accumulation process which drive students’ performance in university in biology is as follows. First, the student accumulates biology-specific human capital through high-school:

$$\theta_{i}^{H,B} = \eta_i + \tau_{H,B}^{H,B}(1 + \phi \eta_i)$$

where $\phi$ is the heterogeneity coefficient. If higher ability students benefit more from the material presented in class, we would expect $\phi$ to be positive.\(^{17}\) Once in university, students are presented more biology material, such that by the end of their first year in university they will have accumu-

\(^{17}\)The form of the human-capital accumulation is kept relatively simple as is done in many papers studying skills formation (e.g. Todd and Wolpin (2007), and Cunha and Heckman (2008)). Allowing for a more complex subject-specific human capital accumulation function (e.g. involving nonlinear functions of ability) would introduce non-trivial identification and estimation issues that are beyond the scope of this paper. See Schennach (2004), Cunha et al. (2010), and Chen et al. (2011) for a discussion on models with nonlinear functions of latent variables.
lated \( \theta_i^{U,B} \) biology-specific human capital:

\[
\theta_i^{U,B} = \theta_i^{H,B} + \tau_i^{U,B}(1 + \phi \eta_i)
\]

\[
= \tau_i^{H,B} + \tau_i^{U,B} + (1 + \phi(\tau_i^{H,B} + \tau_i^{U,B}))\eta_i.
\]

(6)

To be clear about the notation, in equation (6), since students are taught the same material in high school biology, there is no curriculum superscript for biology \( \tau^{H,B,G,13} = \tau^{H,B,G,12} = \tau^{H,B} \). G12 and G13 students are also taught the same biology material in university \( \tau^{U,B} \) since they are in the same classes.

Similarly, a student with ability \( \eta_i \) will have accumulated \( \theta_i^{U,M,C} \) mathematics-specific human capital by the end of her first year of university:

\[
\theta_i^{U,M,C} = \tau_i^{H,M,C} + \tau_i^{U,M} + (1 + \phi(\tau_i^{H,M,C} + \tau_i^{U,M}))\eta_i.
\]

(7)

In contrast to biology, the amount of mathematics-specific material taught in high school will depend on the curriculum attended by the student. Thus we might have \( \tau^{H,M,G,13} \neq \tau^{H,M,G,12} \). For this reason, the level of mathematics-specific human capital by the end of the first year of university \( (\theta_i^{U,M,C}) \) will depend on the curriculum taken by student \( i \)\(^{18}\). Finally, since G12 and G13 students are in the same classes, they are taught the same mathematics material in university \( (\tau^{U,M}) \).

\[
5.2.2 \text{ Grading Rules}
\]

For each student, I observe three measures of performance: 1) a mathematics university grade, 2) a biology university grade and 3) an overall high school average. The university biology and mathematics grades signal the student’s subject-specific levels of human capital. I allow G12 and G13 students to be graded differently in high school—there could be grade inflation across curricula. In this case, the high school average cannot be used to measure the value-added of Grade 13. On the other hand, it can be used as a measure of general academic ability \( (\eta_i) \). Since students from

\[^{18}\text{Subject-specific material in high school and university (} \tau^{H,B}, \tau^{U,B}, \tau^{H,M,C}, \text{ and } \tau^{U,M} \text{) should be seen as functions of the amount of material taught and the time spent on the material. This model does not disentangle the effects of ‘time spent on material’ from the effect of ‘more material’. One could imagine } \tau = f(time, material), \text{ where both time and material have a positive effect on } \tau^{H,M,C}.\]
the same curriculum are taught the same material, high-school average differences among students from the same curriculum will be driven by differences in ability.

Grades are assumed to be assigned according to a linear rule. Since the main driving factor differentiating students high-school averages is assumed to be ability, the high school average ($H^C_i$) can be written as

$$H^C_i = \gamma^{H,C} + \rho^{H,C} \eta_i + \varepsilon^{H,C}_i$$

The slope coefficient ($\rho^{H,C}$) represents the payoff to ability. Differences in $\gamma^{H,G13}$ and $\gamma^{H,G12}$ could be due to grade inflation and/or to average ability differences between G12 and G13 students (if high school grades on a bell curve). The error term ($\varepsilon^{H,C}_i$) represents shocks due to measurement error and possible shocks to student performance (e.g. bad luck or illness). $\varepsilon^{H,C}_i$ is assumed to have mean 0 and is uncorrelated with the student’s ability. Notice that only the left-hand-side variable of equation (8) is observed.

In a similar fashion, the university biology grading rule will link the student’s level of biology-specific level of human capital to a biology grade:

$$B^C_i = \pi^B + \lambda^B \theta_i^U,B + \varepsilon^{B,C}_i$$

I assume that professors do not discriminate against students based on their high school background. The constant and the slope coefficients are then assumed to be the same for both groups. We can rewrite the biology grading as a linear function of a student’s general academic ability using equation (6):

$$B^C_i = \gamma^B + \rho^B \eta_i + \varepsilon^{B,C}_i$$

where

$$\gamma^B = \pi^B + \lambda^B (\tau^{H,B} + \tau^{U,B})$$

$$\rho^B = \lambda^B (1 + \phi (\tau^{H,B} + \tau^{U,B}))$$

19 I could assume that high-schools and university grade their students based on relative performance. It is the case in Morin (2010). Doing so only complicates the notation of the model without adding significantly more to it.

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

17
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 error term of the high-school grading rule ($E(\varepsilon_{i}^{B,C},\varepsilon_{i}^{H,C}) = 0$).

According to equation (9), the expected biology grades for G12 and G13 students are $\gamma^B + \rho^B \eta_{G12}^B$ and $\gamma^B + \rho^B \eta_{G13}^B$, respectively. Hence, the difference between the expected biology grades ($\bar{B}_{G13} - \bar{B}_{G12} = \Delta_B$) is given by equation (3), i.e. $\Delta_B = \rho^B \Delta_\eta$.

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:

$$M_i^C = \pi^M + \lambda^M \theta_i^{U,M,C} + \varepsilon_i^{M,C}.$$  

Using equation (7) we get

$$M_i^C = \gamma^{M,C} + \rho^{M,C} \eta_i + \varepsilon_i^{M,C} \tag{10}$$

where

$$\gamma^{M,C} = \pi^M + \lambda^M (\tau^{H,M,C} + \tau^{U,M})$$

$$\rho^{M,C} = \lambda^M (1 + \phi (\tau^{H,M,C} + \tau^{U,M}))$$

Then the value-added of Grade 13 mathematics (expressed in university mathematics grades) will be given by the difference in conditional expectations of equation (10) across student groups:

$$\Delta V_i = \lambda^M (\tau^{H,M,G13} - \tau^{H,M,G12}) + \lambda^M \phi (\tau^{H,M,G13} - \tau^{H,M,G12}) \eta_i$$

$$\equiv \gamma^{M,\Delta} + \rho^{M,\Delta} \eta_i. \tag{11}$$

The first term in equation (11) is the difference in intercept coefficients in equation (10) between G13 and G12 students (i.e. $\gamma^{M,\Delta} = \gamma^{M,G13} - \gamma^{M,G12}$) and the term multiplying the student ability is the difference in slope coefficients (i.e. $\rho^{M,\Delta} = \rho^{M,G13} - \rho^{M,G12}$). Notice that the value-added of Grade 13 will be a linear function of the student academic ability. It will be larger for better students if $\phi > 0$. If we can identify and estimate both $\gamma^{M,C}$ and $\rho^{M,C}$, we can estimate $\Delta V_i$. It
will also allow me to test whether the value-added of Grade 13 is larger for better students since:

\[
\phi = \frac{\rho_{M,G13} - \rho_{M,G12}}{\gamma_{M,G13} - \gamma_{M,G12}}.
\]  

(12)

Overall, The grading rule model consists of six equations: three equations (equations (8), (9), and (10)) per student group. We can easily see the resemblance to a standard one-factor model where the driving factor is the initial level of general academic ability (\(\eta_i\)). Note that only the left-hand sides of these equations are observable.

### 6 Identification and Estimation

The simplicity of the model makes the identification of the parameters of interest (i.e. \(\Delta V_i\) and \(\phi\)) relatively straightforward. One necessary condition for identification of the factor model parameters is that the latent variable (\(\eta_i\)) must be scaled to one observed variable. That is, the slope and the intercept coefficients of one equation should be predetermined. Since the biology intercept and slope parameters are assumed to be the same across student group, a convenient normalization is to set these two parameters (\(\gamma^B\) and \(\rho^B\)) equal to 0 and 1, respectively.\(^{21}\) Note that this normalization will not play a role in testing for heterogeneity or estimating the value-added of Grade 13.

The identification of the average academic ability levels is straightforward given the normalization I have made:

\[
E_C(\eta_i) = E(B^C_i)
\]  

(13)

The mathematics slope parameters are identified using:

\[
\rho_{M,C} = \frac{\text{cov}(H^C_i, M^C_i)}{\text{cov}(H^C_i, B^C_i)}.
\]  

(14)

Now, equation (14) is key for the identification of \(\Delta V_i\) and \(\phi\), and so deserves more attention. The

---

\(^{21}\) After normalization, the model has a total of 18 measured moments and there are 18 coefficients to be estimated: \((E_{G13}[\eta_i], E_{G12}[\eta_i], \sigma^2_{\eta,G13}, \sigma^2_{\eta,G12}, \gamma^{M,G13}, \gamma^{M,G12}, \rho^{M,G13}, \rho^{M,G12}, \gamma^{H,G13}, \gamma^{H,G12}, \rho^{H,G13}, \rho^{H,G12}, \sigma^2_{H,G13}, \sigma^2_{H,G12}, \sigma^2_{M,G13}, \sigma^2_{M,G12})\). \(E_{G13}[\eta_i]\) and \(E_{G12}[\eta_i]\) are the average levels of initial general academic ability of G13 and G12 students, respectively. Also, \(\sigma^2_{\eta,C} = \text{var}(\eta^C)\) and \(\sigma^2_{\eta,B,C} = \text{var}(\epsilon^{B,C})\). Note that some of the model parameters (e.g. \(\tau^{H,B}\) and \(\tau^{U,B}\)) are not identifiable.
sample analog of equation (14) is the slope of an instrumental variable estimator when regressing mathematics grades on biology grades, using high school averages as instrument. This should not come as a surprise as there is a clear link between Structural Equation Model (SEM) parameters—the type of model estimated here—and the IV estimator (Angrist and Pischke 2009). Given the form of equation (14), we should expect the mathematics slope parameter to be greater than 1 (i.e. the biology slope parameter) if the covariance between mathematics and high school grades is larger than the covariance between biology and high school grades. Table 2 suggests that this is indeed what we should expect. Once \( \rho^{M,C} \) and \( E_C(\eta_i) \) are identified, we can identify \( \gamma^{M,C} \) using equation (10) as:

\[
\gamma^{M,C} = E(M^C_i) - \frac{\text{cov}(H^C_i, M^C_i)}{\text{cov}(H^C_i, B^C_i)} E(B^C_i)
\]

(15)

Finally, since \( \rho^{M,C} \) and \( \gamma^{M,C} \) are identified, so are \( \Delta_V \) and \( \phi \) (using equations (11) and (12)).

Looking ahead, it is worthwhile presenting what happens if we do not allow for heterogeneity in the value-added of Grade 13 mathematics (i.e. if \( \phi = 0 \)). Doing so, will shed light on the link between the estimator coming out of the grading-rule model and the difference-in-differences estimator.

First, notice that the mathematics slope parameters in equation (10) (\( \rho^{M,C} \)) would be equal to \( \lambda^M \) for both G12 and G13 students. This imposes a testable restriction on equation (14): the ratio of covariances in equation (14) should be equal across student groups. The value-added would then collapse to (see equation (11)):

\[
\Delta V_i \bigg|_{\phi=0} = \lambda^M (\tau_{H,M,G13} - \tau_{H,M,G12}) = \gamma^{M,G13} - \gamma^{M,G12}
\]

(16)

which is the difference in the intercept coefficients found in equation (10).

We can easily link equation (16) to the difference-in-differences estimator using equation (15):

\[
\Delta V_i \bigg|_{\phi=0} = [E(M_{iG13}^M) - E(M_{iG12}^M)] - \frac{\text{cov}(H_i^M, M_i^M)}{\text{cov}(H_i^M, B_i^M)} [E(B_{iG13}^M) - E(B_{iG12}^M)]
\]

(17)

\(^{22}\)Details about the identification of the other parameters are found in the Appendix.
where the $C$ superscript are omitted from the covariance terms as their ratios should be equal when $\phi = 0$. If biology and mathematics were to measure ability in the same way (i.e. $\text{cov}(H_i, M_i) = \text{cov}(H_i, B_i)$), then the sample analog of equation (17) would become the standard difference-in-differences estimator. Furthermore, if the two groups of students were identical in terms of ability (i.e. $E(B_{13}^C) = E(B_{12}^C)$), then equation (17) would become a simple means comparison.

The empirical strategy is to fit the sample moments (i.e. sample means, variances and covariances) to their corresponding population moments implied by the model (by equations (8), (9), and (10)). Let $\xi$ be the vector of the 18 parameters to be estimated. Let $S^C$ be the vector of the sample first and second moments and $\Pi(\xi)^C$ be the vector of population moments implied by the model for students group $C$. Then, for each group, we have a fit function defined by

$$F(\xi)_C = (S^C - \Pi(\xi)^C)'W_{C}^{-1}(S^C - \Pi(\xi)^C)$$

(18)

where $W_{C}$ is the estimated asymptotic variance-covariance matrix of $S^C$. The global fit function used in the minimization problem is a weighted average of the groups' fit functions

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

(19)

where $N_{G12}$ and $N_{G13}$ are the number of G12 and G13 students, respectively (and $N_{G12} + N_{G13} = N$). The parameter estimates presented in the next section minimize $F(\xi)$.

7 Results

Table 3 presents the results from minimizing equation (19). None of the mathematics valued-added parameter estimates (estimates for $\gamma_{M,\Delta}$ and $\rho_{M,\Delta}$) are statistically significant. Nevertheless, the value-added for a student with average ability (say $\eta_i = 75$) is estimated at 2.2 and is statistically significant at 1%. That is, controlling for ability, Grade 13 increases this student’s mathematics...
Table 3: Estimation Results

<table>
<thead>
<tr>
<th></th>
<th>G13</th>
<th>G12</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Intercept, γ</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biology</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(        )</td>
<td>(     )</td>
<td></td>
</tr>
<tr>
<td>Mathematics</td>
<td>-20.8</td>
<td>-19.3</td>
</tr>
<tr>
<td>(7.38)</td>
<td>(5.33)</td>
<td></td>
</tr>
<tr>
<td>[7.31]</td>
<td>[5.27]</td>
<td></td>
</tr>
<tr>
<td>High School</td>
<td>74.2</td>
<td>72.5</td>
</tr>
<tr>
<td>(1.61)</td>
<td>(1.36)</td>
<td></td>
</tr>
<tr>
<td>[1.65]</td>
<td>[1.38]</td>
<td></td>
</tr>
<tr>
<td><strong>B. Slope, ρ</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biology</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>(        )</td>
<td>(     )</td>
<td></td>
</tr>
<tr>
<td>Mathematics</td>
<td>1.23</td>
<td>1.18</td>
</tr>
<tr>
<td>(0.10)</td>
<td>(0.07)</td>
<td></td>
</tr>
<tr>
<td>[0.10]</td>
<td>[0.07]</td>
<td></td>
</tr>
<tr>
<td>High School</td>
<td>0.22</td>
<td>0.24</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td>[0.02]</td>
<td>[0.02]</td>
<td></td>
</tr>
<tr>
<td><strong>C. Ability, E[C[η]]</strong></td>
<td>74.3</td>
<td>75.8</td>
</tr>
<tr>
<td></td>
<td>(0.76)</td>
<td>(0.59)</td>
</tr>
<tr>
<td></td>
<td>[0.50]</td>
<td>[0.48]</td>
</tr>
</tbody>
</table>

**D. Value-Added and Heterogeneity**

|                      |       |       |
| γ^{M,Δ}              | -1.56 |       |
| (9.10)              | [9.14]|       |
| ρ^{M,Δ}              | 0.05  |       |
| (0.12)              | [0.12]|       |
| φ                    | -0.03 |       |
| (0.11)              | [0.12]|       |

Notes. Panels A, B, and C present the results from estimating equations (8), (9), and (10) simultaneously while Panel D present the derived estimates for the value-added parameters and the heterogeneity coefficient. OMD standard errors are in parentheses, while bootstrap standard errors (based on 1,000 replications) are in square brackets. n=938.

Performance by 2.2 percentage points. Comparing the value of G13 to the mathematics average and standard deviation (See Table 1), the benefit to Grade 13 is modest in terms of human capital accumulation. A ballpark estimate (based on Loury and Garman (1995) and Jones and Jackson
of the ‘return’ (or earnings growth rate due) to G13 mathematics would be around 2.2 percent.

Interestingly, the heterogeneity parameter estimate is not statistically significant and very close to zero, suggesting that higher ability students did not gain more from Grade 13 mathematics their lower ability counterparts. I further investigate this finding in Section 7.2. The results from Table 4 obtained after imposing homogeneity in the value-added of Grade 13, are very close to the results from Table 3. In particular, the value-added is estimated at 2.3 percentage points, which is about the estimated value-added for a student with average ability when allowing for heterogeneity (2.2 percentage points). Overall, allowing for (linear) heterogeneity does affect my results.

Tables 3 and 4 both suggest that the link between academic and grades differs across all three subjects. The difference between the mathematics and the biology slope coefficients is 0.20 in Table 4 and is statistically significant. This difference is nevertheless modest: all else equal, a 10 point difference in ability would lead a 10 point difference in biology and a 12 point difference in mathematics. The interpretation of this difference is that students’ relative proficiency is more easily signaled in mathematics than in biology. 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. The difference between the university mathematics and the high school intercepts is about 73 points. This difference captures the greater difficulty of university courses and the more intense competition in university classrooms.

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 -1.48, which is just the difference between estimated levels of initial academic ability ($\bar{\eta}^{G_{13}}$ and $\bar{\eta}^{G_{12}}$) in Tables 3 and 4.

---

25 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.

26 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.
### Table 4: Estimation Results—Imposing $\phi = 0$

<table>
<thead>
<tr>
<th></th>
<th>G13</th>
<th>G12</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Intercept, $\gamma$</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biology</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Math</td>
<td>-18.4</td>
<td>-20.6</td>
</tr>
<tr>
<td>High School</td>
<td>74.4</td>
<td>72.4</td>
</tr>
<tr>
<td></td>
<td>(-)</td>
<td>(-)</td>
</tr>
<tr>
<td></td>
<td>(4.32)</td>
<td>(4.36)</td>
</tr>
<tr>
<td></td>
<td>[4.14]</td>
<td>[4.18]</td>
</tr>
<tr>
<td><strong>B. Slope, $\rho$</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biology</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Math</td>
<td>1.20</td>
<td>1.20</td>
</tr>
<tr>
<td>High School</td>
<td>0.22</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>(-)</td>
<td>(-)</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.06)</td>
</tr>
<tr>
<td></td>
<td>[0.05]</td>
<td>[0.05]</td>
</tr>
<tr>
<td><strong>C. Ability, $E_C [\eta_i]$</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>74.3</td>
<td>75.8</td>
</tr>
<tr>
<td></td>
<td>(0.51)</td>
<td>(0.48)</td>
</tr>
<tr>
<td></td>
<td>[0.49]</td>
<td>[0.46]</td>
</tr>
<tr>
<td><strong>D. Value-Added Imposing $\phi = 0$</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta V$</td>
<td>2.27</td>
<td>(0.66)</td>
</tr>
<tr>
<td></td>
<td>[0.67]</td>
<td></td>
</tr>
</tbody>
</table>

Notes. Panels A, B, and C present the results from estimating equations [8], [9], and [10] simultaneously while Panel D present the derived estimates for the value-added parameters and the heterogeneity coefficient. OMD standard errors are in parentheses, while bootstrap standard errors (based on 1,000 replications) are in square brackets. n=938.

This difference is statistically significant. 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.

Finally, by looking at the high school grading policy parameters, I can test for potential grade inflation across student groups. The 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. In fact, when I test for the equality of high school parameters across
G12 and G13 students, 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. Overall, there is no sign of systematic grade inflation for one group of student relative to the other. This last finding will be useful when investigating the robustness of my results.

7.1 Means Comparison, Difference-in-Difference, and OLS Estimates

In Section 4, I presented the potential problems with using standard estimation techniques (e.g. means comparison or difference-in-difference) for estimating the value-added of Grade 13 mathematics. These problems motivated the use of more flexible estimator. One of the nice features of this estimator is that it can be seen as a generalization of difference-in-difference estimator. I now compare the value-added estimates from the grading-rule model to the estimates from more standard estimators.

We can use Table 1 to get the value-added estimates for the means comparison and the difference-in-difference estimator. These are 0.44 and 1.92 (0.44 + 1.48), respectively. The OLS estimates vary significantly depending on whether I use the biology grade or the high school average as a measure of ability: 1.69 when I use the biology grade and 0.30 when I use the high school. The means comparison estimate and OLS estimate based on the high school average are not statistically significant (at any common confidence level). The results all suggest that the value-added to Grade 13 mathematics is modest for high-ability students. Estimates of the value-added are similar whether I use means comparison, OLS estimation, difference-in-differences, or the grading rule model as way of 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. It
is also close to five times the means difference estimate, which shows the importance of 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 standard deviations in the present paper. 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: it represents a difference of about two standard deviations.

The results presented in Table 3 do not suggest the presence of significant heterogeneity in the value-added of Grade 13. But, imposing that the value-added is linear in ability might be too restrictive. As an alternative to estimating the grading model allowing for the value-added to be a linear function of ability, I estimated the model (imposing $\phi = 0$) separately for two groups of students formed 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, this difference suggests that, if anything, lower-ability students gain more (not less) from an extra year of high school mathematics.

27Krashinsky’s main econometric approach is to use OLS and matching estimation on students that are closer and closer in age. In an attempt to reconcile our findings, I tried the following: 1) I used his econometric approach on the students in my main sample, and 2) I used his econometric approach on a set of students enrolled in an introductory commerce course at the main (St. George) campus of the University. In both cases, the estimates suggest that the value-added is very small (close to 2 percentage points) and not statistically different from zero. Finally, the raw data suggest that Grade 12 students do slightly better than Grade 13 students at the St. George Campus while they do significantly worse in Krashinsky’s sample. This further supports the idea that the discrepancy in our findings comes from differences in the two student populations.

28In 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.”
8 Robustness

8.1 Control and Treatment Groups

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 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 11 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 2.27). 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.\(^{29}\) 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.

I focused on Life Sciences since it is a large program which allows the researcher to observe a large number of students taking two ‘independent’ subjects; one subject being affected by the reform and the other not. Commerce students also have to take a calculus course (MAT133Y) during their first year. In order to estimate the value-added of Grade 13 mathematics for these students, I estimated the grading-rule model using Introduction to Commerce (COM110H) as the control subject. This course introduces students to management, and the amount of mathematics involved in this course is minimal. Table 5 suggests the value-added of Grade 13 mathematics is small for these commerce students, and not statistically significant. This finding is in line with estimated value-added for life-science students. Note that the results presented in Table 5 are less precise than for life-science students as the sample size is significantly smaller (228 versus 938 students). One difference between the results presented and those presented in Table 4 is that, here, one cannot reject the hypothesis that G12 and G13 students have the same level of academic ability.\(^{30}\)

### 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

\(^{29}\)Humanities represent 38% of G12, and 41% of G13 student applications. Life Sciences represent 39% of G12, and 34% of G13 student applications.

\(^{30}\)I further investigated the potential value-added of Grade 13 mathematics for Commerce students by estimating the same model using two other introductory courses that do not involve calculus (Introduction to Financial Accounting and Introduction to Economics) with similar results. In both cases, the value-added estimate was small and not statistically significant. One should be cautious when interpreting the results for economics and financial accounting as mathematics skills could affect the performance in these subjects.
important ingredients in the human-capital production function. Not being able to fully capture differences across student groups in 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

---

**Table 5: Estimation Results for Commerce Students**

<table>
<thead>
<tr>
<th></th>
<th>G13</th>
<th>G12</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Intercept, ( \gamma )</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Introduction to Commerce</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Mathematics</td>
<td>-45.2</td>
<td>-46.0</td>
</tr>
<tr>
<td>High School</td>
<td>57.6</td>
<td>71.8</td>
</tr>
<tr>
<td><strong>B. Slope, ( \rho )</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Introduction to Commerce</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Mathematics</td>
<td>1.61</td>
<td>1.61</td>
</tr>
<tr>
<td>High School</td>
<td>0.44</td>
<td>0.24</td>
</tr>
<tr>
<td><strong>C. Ability, ( E_C[\eta_i] )</strong></td>
<td>73.0</td>
<td>72.1</td>
</tr>
<tr>
<td><strong>D. Value-Added Imposing ( \phi = 0 )</strong></td>
<td>( \Delta V )</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.07)</td>
</tr>
</tbody>
</table>

Notes. Panels A, B, and C present the results from estimating equations (8), (9), and (10) simultaneously while Panel D present the derived estimates for the value-added parameters and the heterogeneity coefficient. OMD standard errors are in parentheses, while bootstrap standard errors (based on 1,000 replications) are in square brackets. \( n = 228 \).
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 6 suggests otherwise: 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.

Table 6: Students Average Marks in 2003

<table>
<thead>
<tr>
<th></th>
<th>Marks</th>
<th>H₀ : G13 = G12</th>
<th>p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>G13</td>
<td>G12</td>
</tr>
<tr>
<td>Anthropology</td>
<td>501</td>
<td>66.8</td>
<td>65.7</td>
</tr>
<tr>
<td>Biology</td>
<td>1,161</td>
<td>73.2</td>
<td>74.4</td>
</tr>
<tr>
<td>Chemistry</td>
<td>1,016</td>
<td>76.2</td>
<td>76.8</td>
</tr>
<tr>
<td>Economics</td>
<td>629</td>
<td>68.1</td>
<td>68.2</td>
</tr>
<tr>
<td>History</td>
<td>293</td>
<td>67.2</td>
<td>69.0</td>
</tr>
<tr>
<td>Math (Commerce)</td>
<td>281</td>
<td>69.8</td>
<td>67.3</td>
</tr>
<tr>
<td>Math (Life Science)</td>
<td>1,092</td>
<td>68.5</td>
<td>68.5</td>
</tr>
<tr>
<td>Philosophy</td>
<td>448</td>
<td>71.4</td>
<td>72.0</td>
</tr>
<tr>
<td>Psychology</td>
<td>883</td>
<td>70.0</td>
<td>70.3</td>
</tr>
<tr>
<td>Sociology</td>
<td>791</td>
<td>65.5</td>
<td>64.7</td>
</tr>
</tbody>
</table>

Notes. The specific codes of the courses analyzed in this table are, in order, ANT100, BIO150, CHM138, ECO100, HIS109, MAT133, MAT135, PHL100, PSY100, and SOC101. The first column presents the number of observations. The last column presents the p-values associated with (two-sided) tests of equality for the average course marks between Grade 12 and Grade 13 students.

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) and that high school grades are good measures of general academic ability, it would be even harder to support the possibility of such a general effect on students. Running a regression of biology grades on high school grades suggests a small and negative value-added of Grade 13. 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 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 differences, 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—in fact the estimated value-added is smaller than when looking at the full sample. 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

---

31 The results are not presented here, but are available upon request.
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 6 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 mathematics is downward-biased.

9 Conclusion

Despite representing a majority of high school students in Canada and the United States and a very high proportion in much of Europe, we know little about the benefit of secondary schooling for university-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 university-bound students, accounting for ability differences due to self-selection.

The results obtained in this study suggest that the benefit to an extra year of mathematics for university-bound students is modest. I find that students coming out of Grade 13 only have a 2.3 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 helps 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.

In an area in which there is virtually no prior empirical evidence, the paper prompts further research: the finding that high-ability students do not gain much from an extra year of mathematics raises the 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.’ Either way, understanding in more detail why high-ability students do not benefit much from an extra year of mathematics should lead to more informed decisions regarding the allocation of (scarce) high school resources.

References


A Appendix

A.1 Identifications of the Model Parameters

In this subsection, I present details about the identification the parameters that are not presented in the main text. First, the slope parameters of the high school grading policies are identified using:

$$\rho_{H,C} = \frac{\text{cov}(H_i^C, M_i^C)}{\text{cov}(M_i^C, B_i^C)}.$$  \hspace{1cm} (20)
Once we have identified the high-school slope parameters and the average ability of G12 and G13 students (through equation (13)), we identify the high-school constant parameters as:

\[
\gamma^{H,C} = E(C_i^H) - \frac{\text{cov}(H_i^C, M_i^C)}{\text{cov}(M_i^C, B_i^C)} E(B_i^C)
\]  

We can use equations (9) and (10) to identify the variance of the G12 and G13 ability distributions:

\[
\text{cov}(M_i^C, B_i^C) = \rho M,C \sigma^2 \eta C
\]  

Using equations (14) and (22) we get:

\[
\sigma^2 \eta C = \text{cov}(M_i^C, B_i^C) \frac{\text{cov}(H_i^C, B_i^C)}{\text{cov}(H_i^C, M_i^C)}
\]  

Using equations (9) and (23), we can identify \(\sigma^2 \varepsilon_{B,C}:\)

\[
\sigma^2 \varepsilon_{B,C} = \text{var}(B_i^C) - \text{cov}(M_i^C, B_i^C) \frac{\text{cov}(H_i^C, B_i^C)}{\text{cov}(H_i^C, M_i^C)}
\]  

In a similar fashion, we get:

\[
\sigma^2 \varepsilon_{M,C} = \text{var}(M_i^C) - \text{cov}(H_i^C, M_i^C) \frac{\text{cov}(M_i^C, B_i^C)}{\text{cov}(H_i^C, B_i^C)}
\]

\[
\sigma^2 \varepsilon_{H,C} = \text{var}(H_i^C) - \text{cov}(H_i^C, M_i^C) \frac{\text{cov}(H_i^C, B_i^C)}{\text{cov}(M_i^C, B_i^C)}
\]