

# Evidence on Grades and Grade Inflation at Ontario's Universities

PAUL M. ANGLIN

RONALD MENG

*Department of Economics*

*University of Windsor*

*Windsor, Ontario*

En utilisant les notes obtenues durant la première année à l'université en Ontario nous examinons la possibilité d'une hausse des notes dans différentes disciplines. Dans un sondage de sept universités durant les périodes allant de 1973-74 à 1993-94, nous trouvons une hausse significative des notes dans divers programmes dans les arts et les sciences. Le taux de hausse n'est pas uniforme. Par exemple, dans les mathématiques, il n'y a eu pas ou peu de changement dans les notes moyennes dans la plupart des universités tandis qu'en anglais et en biologie, la hausse des notes était significative.

Using information on first-year university grades from across Ontario, we examine whether or not there has been grade inflation by discipline. In a survey of seven universities for the periods 1973-74 and 1993-94, we find significant grade inflation in various Arts and Science programs. The rate of inflation is not uniform. Some subjects, such as Mathematics experienced little or no change in average grades at most universities, while English and Biology experienced significant grade inflation.

---

## INTRODUCTION

Universities are being examined in more detail than they have been in recent memory. Tax-payers, governments, and students are becoming increasingly concerned about getting their money's worth. While this increased scrutiny makes some university administrators and professors uneasy (Bruneau 1994), the principles of accountability and transparency are slowly being accepted in the education industry. There have been recent studies on the relative efficiency of Canadian universities (McMillan and Datta 1998), the private economic return in obtaining a degree (Lavoie and Finnie 1999; Dooley 1986; Freeman 1999), the social returns to a degree (Vaillancourt 1995) and the future of university financing (Carmichael 1999; West

1988). *MacLean's* magazine's annual rankings of universities are often disputed but few disagree about the popularity, or the profit produced by its publication (both in English and Chinese, although there is no equivalent in French). These changes can be expected to continue because of the fundamental shift in education funding when governments reduce their contribution and deregulate fees (CAUT 1996).

Each of these and many other issues have been extensively and deservedly studied in this journal and elsewhere. Still, an important dimension of the modern Canadian university has been overlooked: grade inflation. Grade inflation is generally unmeasured and is therefore omitted from the list of indicators studied by *MacLean's* or by governments. The fact that our data was difficult to obtain

motivates one of the lessons that we wish to convey. We offer our results to stimulate an informed discussion about grading policies.

Why should researchers be interested in grade inflation? First, grading offers an unobserved margin of adjustment that can be exploited. In times of government cutbacks, this margin represents an important disadvantage to using performance indicators, such as student-professor ratios and research funding, as proxies for the cost of educating a student or the quality of that education. Simply put, professors could manipulate grades in order to attract more students and cushion the negative effects of any government cutbacks. Second, our data should help to refine the debate on the returns to increasing the quantity of education since those studies often ignore marginal changes in its quality. Third, if grade inflation varies between departments then the signals and incentives that students receive are distorted and students are less able to judge where their comparative advantage lies. If grades do not indicate a student's strengths and weaknesses then the expectation of a good grade in a certain discipline will influence their choice of courses (Sabot and Wakeman-Linn 1991). By lowering the relative price of some subjects compared to others, students may choose the "wrong" field of study in terms of their own comparative abilities. For example, to increase the number of computer scientists, maybe universities should be encouraged to offer higher grades in such courses and offer lower grades in disciplines with an excess supply of students. Finally, grades and the acquisition of a degree are signals to potential employers. Grade inflation makes these signals less clear. If an "A" student today is not the same as an "A" student 20 years ago, or the contribution of an "A" student in English cannot be easily compared to that of an "A" student in Mathematics, then what criteria does an employer use to distinguish differences in quality? Few people question that costs due to mistakes exist.

The issue of grade inflation has been discussed for many years (Durm 1993; Juarez 1996; McKenzie

1979) but debates on the issue are usually restricted by a lack of data. This paper helps to fill this void by reporting on grading policies at seven Ontario universities: Brock, Guelph, McMaster, Ottawa, Trent, Wilfrid Laurier, and Windsor in 12 traditional introductory (first-year) courses. Our reference years are for the 1973-74 and 1993-94 academic years.<sup>1</sup>

The next two sections look at two aspects of grades. The first section considers the mean grade point average (GPA), the second section focuses on changes in the distribution of grades, and this is followed by our conclusion. The Appendix shows how the data were constructed and how we accounted for changes in the way grades are recorded over time. In order to compare grades over time, we use a common index. Most institutions in our sample report final grades as a letter. For comparison purposes, we quantify letter grades in the following standard format: an A is 4, B is 3, C is 2, D is 1, F is 0, and we ignore + and - distinctions. Our data show that the rate of inflation has not been equal across courses and we give an indication of the changes across the distribution from A to F.

## GRADE INFLATION

Table 1 compares the average grade awarded to students enrolled in courses offered during the 1973-74 academic year with the average grade awarded in a comparable course during 1993-94. Table 1 also records the variance in grades across the universities in our sample for each academic year and the number of universities that experienced grade inflation or deflation in each course.<sup>2</sup> The average GPA rose in 11 of the 12 courses. The only exception was Sociology and its average GPA is above the median in both years.

More importantly, this table shows that even among the courses that experienced grade inflation, the rate of inflation was not uniform. For example, the greatest increase occurred in English, with Biology and Chemistry following. The focus on

TABLE 1  
Mean Grades and Variances in Seven Universities, 1973-74 and 1993-94

Departments	1973-74		1993-94		Number of University Courses	Inflation*	Deflation*	No Change*
	GPA	Variance	GPA	Variance				
English	2.17	1.04	2.76	0.89	6	6	0	0
French	2.47	1.15	2.69	1.07	7	4	0	3
Music	2.89	0.95	3.02	1.16	4	2	1	1
Philosophy	2.38	1.15	2.54	1.07	7	4	1	2
Biology	2.18	1.19	2.52	1.19	7	5	1	1
Chemistry	1.88	1.57	2.18	1.51	7	4	1	2
Mathematics	2.14	2.00	2.19	1.86	7	2	1	4
Physics	2.17	1.65	2.38	1.63	7	4	1	2
Economics	2.07	1.53	2.18	1.44	7	3	1	3
Political Science	2.37	0.90	2.49	0.94	7	3	1	3
Psychology	2.31	1.16	2.40	1.17	7	3	2	2
Sociology	2.57	1.02	2.51	0.92	7	2	3	2
<b>Totals</b>					80	42	13	25

Note: \*10 percent level of significance. The means and variances for each of the 80 university courses were generated for the two reference years. Standard t-statistics for differences in means were calculated for each of the 80 pairings.

inflation instead of just GPA levels is important because the highest 1993-94 grades are found in Music, English, and French while the lowest grades are Economics, Mathematics, and Chemistry. Twenty years ago and with the marked exception of English, the same courses were highest and lowest grading.

The variances measure the average of the variances in each department across the different universities.<sup>3</sup> A decrease in the variances between 1973-74 and 1993-94 would demonstrate that the distribution of grades is shrinking around the mean. In all cases, except Music, Political Science, and Psychology the variances declined or stayed the same over the 20-year period.<sup>4</sup> The biggest decline in the variance is found in English.<sup>5</sup> Generally, and as expected, departments with the biggest decline in the variance of grades also had the biggest inflation. The notable exceptions are Mathematics and

Economics which, as discussed below, have lower variances because they have fewer failures in the early 1990s than in the 1970s.

Of the 80 individual course-university combinations on which we had data for both beginning and end period grades, 42 (53 percent) had grade inflation at the 10-percent level of significance, 25 (31percent) had no statistically significant change in their mean grades, and average grades fell in 13 (16 percent). These results indicate that while grade inflation is not uniform across all universities or even all courses, the pattern of rising grades is widespread. Our results are especially noticeable in the Arts where 16 of the 24 course-university combinations showed statistically significant grade inflation.

Only two courses (Psychology and Sociology) had grade deflation at more than one university, while further evidence suggests that the cases of

deflation may represent the idiosyncratic behaviour of specific professors. Nine of the 13 course-university combinations with grade deflation had exceptionally high grades in 1973-74 relative to the mean for that course at other universities in that year. On average, these nine courses were a half a grade (0.46) higher than the rest of the subjects in their field across the other five universities.

### DISTRIBUTION OF GRADES

For many purposes, the distribution of grades is more important than the average grade in a course. There have always been competing views concerning the purpose of a university, but students and their parents are increasingly concerned about getting jobs. As part of this view, employers would like prospective employees to be well trained.<sup>6</sup> Or, if the student is not well trained, employers could profitably hire him or her at a reduced wage if there were reliable information about a student. Letters of reference have well-known weaknesses. Simply graduating from a university is less distinctive than in earlier decades, which leaves grades as an important measure for employers, especially measuring grades relative to those of classmates.

Table 2 reports on the fraction of students who received As, Bs, Cs, Ds, and Fs. The horizontal sum of the first five columns, except for small rounding errors, totals 100 percent, as does the horizontal sum of the second five columns. The last two columns also show the 1993-94 percentages of As and Fs less the 1973-74 estimates. As expected, more students are now receiving high grades and fewer students are receiving low grades compared to 20 years ago. All of the introductory Arts courses have more than 50 percent of students receiving a B- or higher. Some people argue that most of the grade inflation has occurred at the B and C level, but that professors have preserved the value of an "A." These data suggest otherwise. Statistically significant increases in the fraction of a class receiving an A occurred in most of the courses that experienced grade inflation.

These data indicate that the time of the "Gentleman's C" is gone. It would be more accurate to refer to the time of the "Gentleperson's B" (or even A in some cases). One could respond that these courses attract better students but we have rather limited evidence on this hypothesis. One could argue that students should take as many introductory courses as possible in order to be well-rounded citizens. Even if true, not all of these students should be encouraged to continue. For this reason, the widespread decreases in the fraction of a class receiving Ds or Fs are worrisome.

The results presented in the second last column,  $\Delta$  (%A), indicate that 11 of the 12 subjects assigned more As in 1993-94 than in 1973-74. The only exception is Sociology. In the case of English, for instance, the fraction of the class with As during these two periods more than doubled from 8.5 to 18.4 percent. Averaged across courses, the increase in the fraction receiving an A over a 20-year period was 4.7 percentage points.

Except for Music, which already has an exceptionally low failure rate, fewer Fs are being assigned. Some disciplines like English, French, Music, and Biology have significant grade inflation at the upper tail of the distribution. Interestingly, a number of other subjects, especially in the sciences, have significantly lower failure rates with small increases in the percentage of As. These subjects apparently have grade inflation at the lower end of the grade distribution.

It is clear from Table 2 that widely different grading standards apply within a university. There are now the hardest grading departments (Chemistry, Mathematics, and Economics) and the softest grading departments (English, French, and Music). A few disciplines, like Political Science, Physics, and Psychology are medium hard while Philosophy, Biology, and Sociology would be classified as medium-soft.<sup>7</sup> While there has been a general upward drift in grades, the differences in discipline-specific grading practices is quite revealing. Most depart-

TABLE 2  
Distribution of Grades: 1973-74 and 1993-94

Departments	1973-74					1993-94					$\Delta(\%A)$	$\Delta(\%F)$
	%A	%B	%C	%D	%F	%A	%B	%C	%D	%F		
English	8.5	31.6	36.8	15.0	8.1	18.4	45.6	26.7	5.8	3.6	9.9	-4.5
French	17.7	36.2	27.7	13.0	5.4	24.5	36.6	26.5	8.2	4.2	6.8	-1.2
Music	29.2	43.4	17.4	7.3	2.8	42.7	29.4	17.7	7.1	3.1	13.5	0.3
Philosophy	14.5	35.3	31.6	11.3	7.2	18.1	38.2	28.3	10.4	4.9	3.6	-2.3
Biology	12.8	29.1	30.3	18.8	9.1	22.6	31.9	25.7	14.0	5.7	9.8	-3.4
Chemistry	15.1	20.2	24.6	22.0	18.1	18.0	24.9	26.2	20.2	10.7	2.9	-7.4
Mathematics	22.9	21.9	19.7	17.1	18.4	23.6	20.9	21.5	19.2	14.7	0.7	-3.7
Physics	18.7	26.6	21.7	18.9	14.0	21.1	25.0	25.4	18.5	9.9	2.4	-4.1
Economics	14.0	25.8	26.6	20.3	13.4	17.1	24.5	27.8	20.7	10.0	3.1	-3.4
Political Science	9.6	38.7	37.1	9.2	5.4	12.7	42.9	30.2	9.0	5.1	3.1	-0.3
Psychology	15.9	30.8	30.0	16.4	6.8	17.2	31.2	31.0	15.6	5.0	1.4	-1.8
Sociology	16.9	43.5	25.6	8.2	5.8	15.9	36.9	33.1	10.8	3.4	-1.0	-2.4
Average	16.3	31.9	27.4	14.8	9.5	21.0	32.3	26.7	13.3	6.7	4.7	-2.8

ments, historically and presently gave either high (e.g., Music) or low (e.g., Economics) grades. However, English and Biology would have been considered among the hardest grading departments in the early 1970s, but are soft grading departments in the 1990s.

## DISCUSSION

Our survey of selected Ontario universities finds significant grade inflation. Are these results typical of other universities in Canada or elsewhere? We know of no comparable Canadian study while Sabot and Wakeman-Linn's (1991) results imply that grade inflation is common throughout the United States.

A vast number of hypotheses have been put forward to explain grade inflation. To generalize, grades could increase because: (a) professors are

more generous<sup>8</sup> for a given group of students, (b) the quality (e.g., intelligence, study effort) of a group of students has increased over time, (c) for a given group of students, teaching methods improved, and (d) random effects that apply to a particular observation (i.e., specific course-university effects). Most people believe that the first hypothesis is the most prevalent and represents the greatest concern to policymakers while acknowledging that the second and third make careful analysis difficult. We acknowledge the final hypothesis by noting which changes are statistically significant. The essential problem in studying or verifying any of these hypotheses is that they rely on data that are not readily available. Hopefully, our results will encourage researchers who have access to the required data to conduct appropriate studies.

Canadian universities are under pressure, mostly due to declining federal and provincial grants for

postsecondary education, to attract more students and to retain them. The competition between universities to attract students, and the resources that go with them, has intensified over the last decade. There is also greater rivalry between departments within each university because increased student enrolment for any given subject will mean that, at a minimum, faculty size will not fall as fast as in other less fortunate departments. This simple-minded intuition requires two points of clarification.

First, professors sometimes feel compelled to award higher grades, ask less of students, or anticipate lower enrolment (Dickson 1984). To our knowledge, only one study attempts to estimate the response of student choice with respect to grades. Sabot and Wakeman-Lin (1991) studied a representative sample of students at Williams College to see if course choice was a function of grades. They developed a panel data set that tracked their students through their undergraduate studies. Starting with first year and using a probit model they measured the influence of the grade received by a student on the possibility that the same student would take a second course in the same department. After controlling for such things as relative ranking of the student (the marginal grade versus the average grade in the class), gender, and several socio-economic background characteristics, their results indicate that subsequent enrolment in low-grading departments is adversely affected by the absolute grade the student receives in their introductory courses. In the case of Economics, also a low-grading department at Williams (and other US schools), changing a student's grade from a B to an A increased the probability of taking a second course in the department by 15 percent.

Second, are university administrations aware of the differential grade inflation and its possible effects on resource allocation? In universities with their specialization on advanced knowledge, such awareness is not clear, and for this reason, there is a potential unresolved governance question of how administrations can deal with grade inflation while

professors profess their specializations. If universities offer a differentiated product, in the sense of specializing in certain disciplines, grade inflation is a potential way for less favoured disciplines at a university to attract attention and students; conversely, a favoured discipline with excess students may find that tougher grading is a way to cut costs. Thus, when this "monopolistically competitive" market reaches an equilibrium, the net effects of differential grade inflation may be hard to sort out. How to make this informal explanation precise is an open research question.

Given these issues and as long as resources are tied to student enrolment, it is possible that grades will continue to rise and employers will have an increasingly difficult time distinguishing a well-trained student from someone who just "gets by." If the informational value of a degree's grades deteriorates, employers will demand an even greater signal to ensure quality. Therefore, a debate on grading leads to a debate on "credentialism." Credentialism refers to the idea that a university degree is valuable because employers use it to reduce the pool of applicants rather than because it demonstrates particular skills. Some (e.g., Orr 1997) claim that all courses "look the same on a transcript." If a credential no longer provides added value to the economy then a university degree represents an enormous expense for a government without producing a public benefit.

## NOTES

We would like to thank Pamela Kirch, Ruisheng Pan and Linda Qiu for their research assistance as well as Charles Beach, Felice Martinello and three anonymous referees for their valuable comments. We are also indebted to our colleagues at various Ontario universities and the Registrar's Office at the University of Windsor for supplying us with the data used in this study. Comments can be sent to Department of Economics, University of Windsor, Windsor, Ontario, N9B 3P4, or by email to panglin@uwindsor.ca or to rmeng@uwindsor.ca

<sup>1</sup>Of the seven universities in our survey, the beginning period for two universities was 1972 and for another it was 1975. The authors contacted 15 universities in the province in order to obtain data on grades. Only about half gave us data for the two periods.

<sup>2</sup>The potential number of comparisons for 12 subjects at the seven universities is 84. However, three universities did not have Music grades for at least one period and one university could not provide their 1973-74 English grades.

<sup>3</sup>Throughout this analysis we have used unweighted averages and variances. If we weight the summary statistics by student enrolment, the numbers from one or two large universities swamp the smaller institutions. By weighting our estimates, we would also be studying the wrong question. Our goal is to look at the behaviour of a "typical" program at a representative university.

<sup>4</sup>The results for Music should be viewed with caution since they use data from only four universities.

<sup>5</sup>As far as the authors are aware none of the courses were compulsory for a wide range of students in 1973-74. Being a required course could change the situation if professors view their classes a "captive audience" and would not have to compete with other departments for students.

<sup>6</sup>Reading the *London Times*, a leading British newspaper, reveals an interesting fact: job advertisements ask for minimum standing, e.g., "a 2:1 honors degree" at graduation. Our thanks to Marion Steele for noting this fact.

<sup>7</sup>Our lines of demarcation between departments are somewhat arbitrary, but Table 1 suggests there are noticeable clusters of disciplines. The hardest grading departments have less than a 2.2 GPA in 1993-94, while the softest grading departments have a GPA greater than 2.6 and others are, of course, in between the two averages.

<sup>8</sup>"Generosity" should be interpreted broadly. Without a structural model of behaviour, we do not distinguish between changes in professors' tastes and changes in their incentives, which could be referred to as hypotheses (a)(i) and (a)(ii). At the very least, identifying incentives that are compatible with an "equilibrium" requires careful study of the competition for students within a university and between universities.

## REFERENCES

- Bruneau, W. 1994. "The Perils of Performance Indicators," Canadian Association of University Teachers, unpublished paper.
- Canadian Association of University Teachers (CAUT). 1996. "Governments, Too, Are Accountable," *CAUT Bulletin*, Supplement, May.
- Carmichael, L. 1999. "Restructuring the University System: What Level of Public Support?" *Canadian Public Policy/Analyse de Politiques* 25(2):133-40.
- Dickson, V. 1984. "An Economic Model of Faculty Grading Practices," *Journal of Economic Education* 15:197-203.
- Dooley, M. 1986. "The Overeducated Canadian? Changes in the Relationship among Earnings, Education and Age for Canadian Men: 1971-81," *Canadian Journal of Economics* 19:142-59.
- Durm, M. 1993. "An A is not an A is not an A: A History of Grading," *The Education Forum* 57:294-97.
- Freeman, D.G. 1999. "Grade Divergence as a Market Outcome," *Journal of Economic Education* 30(3):344-51.
- Juarez, T. 1996. "Why Any Grades at all, Father?" *Phi Delta Kappan*, pp. 374-77.
- Lavoie, M. and R. Finnie. 1999. "Is It Worth Doing a Science or Technology Degree in Canada? Empirical Evidence and Policy Implications." *Canadian Public Policy/Analyse de Politiques* 25(2):101-21.
- Martin, B. 1998. "Great Grade Inflation Scam," *National Jurist* (March/April):26-32.
- McKenzie, R.B. 1979. *The Political Economy of the Education Process*. Boston: Martinus Nijhoff Publishers.
- MacMillan, M. and D. Datta. 1998. "A DEA Analysis of University Productivity," *Canadian Public Policy/Analyse de Politiques* 24(4):485-511.
- Millar, M. 1997. "Marking at Queen's: A Comparative Study," Queen's University Alma Mater Society, Kingston, Ontario.
- Orr, R. 1997. "The Executive Dean's Corner: More on Getting the Grades You Want...", *The Lance* 70(13):23.
- Sabot, R. and J. Wakeman-Linn. 1991. "Grade Inflation and Course Choice," *Journal of Economic Perspectives* 5:159-70.
- Vaillancourt, F. 1995. "Private and Total Returns to Education in Canada, 1985," *Canadian Journal of Economics* 28:532-54.
- West, E. 1988. *Higher Education in Canada*. Vancouver, BC: Fraser Institute.

## **APPENDIX CONSTRUCTING DATA**

We collected data on grades at different universities at different times. Universities have a surprising number of ways of reporting grades. Plus, these ways differ from the numerical 0-100 scale often used on tests that are returned to students.

We compared grades in 12 courses because they are core courses and are offered at nearly every university. We excluded professional courses such as engineering or business. We focused on first-year courses because they have no prerequisites that might limit students and that might change over time. Grades in higher level courses (see Millar 1997, for some data) reflect some self-selection by students and are less pure measures of grade inflation.

Our aggregation of grades may raise some questions. All universities report an A to F scale but not all use the same method of converting it to a GPA. We assume that an A counts as a 4, B is 3, C is 2, D is 1, F is 0. We ignore the difference between a B+ and a B-, counting both as 3 points. The difference between a B+ and an A- is important because an A- counts as 4 points. Some universities, especially in the earlier years, did not give +/- . A more serious problem might occur if a mass of students now receive a B- who would have received a C+ under the older grading policy because the measured change would overstate the true change in grades. The sign of this effect is ambiguous, though, since a change from B- to B+ has no effect on our measure of GPA.

In the normal course of events, some students withdraw from a course while others take a course for a second time in hopes of getting a higher grade. Since these people have some information on their abilities, they are unlikely to be a typical first-year student. Students who withdraw from a class may find the course too demanding or because they want to avoid a low expected grade or because of personal problems unrelated to their studies. Students who take a course for a second time tend to have received a low grade the first time. For these reasons, selection bias could affect our estimates of the “average” GPA in a class at any point in time. This potential is especially relevant when university rules change over time and differ between universities. While acknowledging this problem, we feel that we cannot solve it without much more data and data of a different type. We chose to count every student who finished a class equally: that is, students who drop a course are ignored; students retaking a course are treated as though taking it for the first time.

Some readers may be concerned that, even if there is no grade inflation, there may be disguised grade inflation as professors adjust the content of their courses. As with several other hypotheses that were noted in our discussion, we cannot confirm or refute this hypothesis without more data of a different kind.