IAPR TECHNICAL PAPER SERIES

DIFFERENTIAL GRADING STANDARDS AND STUDENT INCENTIVES

B. Curtis Eaton
Department of Economics
University of Calgary

and

Mukesh Eswaran
Department of Economics
University of British Columbia

Technical Paper No. TP-06011

Institute for Advanced Policy Research
University of Calgary
Calgary, Alberta
Canada

http://www.iapr.ca
iapr@ucalgary.ca

Correspondence: B. Curtis Eaton, Department of Economics, University of Calgary, eaton@ucalgary.ca

Differential Grading Standards and Student Incentives

B. Curtis Eaton, University of Calgary

Mukesh Eswaran, University of British Columbia

January 2001, Revised November 2006

ABSTRACT: We present data on grades in three Canadian universities that suggest that grading standards differ significantly across disciplines within universities. To explore the implications of differential grading standards, we develop a simple human capital model in which students are differentiated by their aptitude in each of two programs offered by some university. In each program the grading standard is a linear mapping from a student's human capital to her grade. In the post graduation labour market firms can observe grades but not human capital, so students who achieve the same grade are pooled. When grading standards differ, relative to the optimal allocation of students to programs, some students are induced to enrol in the wrong program, others who ought to choose the university option are induced to stay out, and yet others who ought not to be in university choose to enrol. Further, within their chosen programs, student incentives with respect to quantity of human capital are distorted, so that some students overachieve relative to the optimum and others underachieve. We suggest some non-intrusive ways to rectify the problem.

JEL Classification Nos: I21, J26

Key Words: Grading standards, human capital, pooling, screening, education

Acknowledgments: We would like to thank Chris Auld, Herb Emery, Lasheng Yuan and J. F. Wen for thoughtful comments on an earlier draft. Paul Anglin's unpublished work on grade inflation inspired our interest in the subject. We thank Sarah Eaton and Subrata Sarker for research assistance.

Correspondence: Address to Curtis Eaton, Department of Economics SS 454, The University of Calgary, 2500 University Drive N.W., Calgary, Alberta, CANADA T2N 1N4.

Email: eaton@ucalgary.ca and eswaran@econ.ubc.ca

1 Introduction

We examine some recent time series data on grades awarded by discipline in three Canadian universities. In these data, there are persistent and significant differences across disciplines in both the percentage of students who get high grades and in the average grade awarded. A student's grade in any course is just a mapping from the student's achievement in the course, as measured by her performance on assignments and examinations, to a grade for the course. If we suppose that instructors within disciplines use a common grading standard, we see that differences across disciplines in grades awarded can, in principle, be decomposed into a component that reflects differences in student achievement levels across disciplines, and a component that reflects differences across disciplines in the way student achievements are mapped to grades—that is, a component that is attributable to differences in grading standards. Regrettably, the data we have do not allow us to actually perform this decomposition. Nevertheless, it appears to us that some substantial portion of the differences in grades awarded across disciplines is due to differences in standards, and we certainly cannot reject this hypothesis.

This we think raises some fundamental issues for universities. It would seem that, on the instructional side of its mandate, the primary objective of a university is to promote student academic achievement. In support of this objective, considerable effort goes into the measurement and reporting of academic achievement, and students are rewarded in a variety of ways based on their reported achievement. At the interface of the measurement/reporting apparatus are the grading standards of individual instructors—the way in which individual instructors map observed achievements of the students in their courses to the grades that they report to the Registrar.

If grading standards are uniform, then (*i*) reported marks are comparable across courses and students, (*ii*) GPAs of different students can be meaningfully compared, (*iii*) a student's

GPA can be interpreted as an aggregate measure (depending on the common standard) of her achievement, and (*iv*) most importantly, GPAs can legitimately be used to reward performance and encourage achievement. Both inside universities, and to a lesser extent outside them, GPAs are used in precisely this way—to award scholarships, honours and degrees, and to ration access to courses, academic programs, and jobs.

In short, if grading standards are uniform, grades are a legitimate unit of account as regards student academic achievement. Conversely, to the extent that grading standards differ across instructors and disciplines, grades are not comparable across courses and students, and grades are not a legitimate unit of account.

In Section II we examine the data—they support a *prima facie* case that there are significant and persistent differences in grading standards across disciplines within these three universities. In Section III we use a human capital model to articulate the ways in which differential grading standards distort the decisions that students make. An essential feature of our model is that students with identical grades but different amounts of human capital are pooled in the labour market. In the concluding section we argue that it is possible to use the very extensive data that universities routinely collect to determine if the *prima facie* case we make here holds up on closer examination. In addition, we discuss ways in which the university could respond to what we think are the very serious issues that non-uniform grading standards raise.

It seems clear to us that the issues we raise in this paper are not peculiar to the three universities for which we have data. For one thing, we have seen fragmentary data for other Canadian universities that is entirely consistent with what we see in the data examined in this paper. [1]

[1] For another, there is a curious unwillingness to make this sort of data available. The protection of privacy is not the reason—it is impossible to infer anything about the academic performance of a particular student or the standards of a particular instructor from the data we use in this study— grades awarded by discipline,

We are, of course, not the first people to raise issues concerning grading standards. It is useful to distinguish two issues: grade inflation (the deterioration of grading standards with time) and differential standards. The classic study in the economics literature is Sabot and Wakeman-Linn (1991), which presents data on grades received by students in a number of disciplines for eight American colleges and universities for academic years 1962-63 and 1985-86. In the earlier year, within a university or college, there was a remarkable similarity in both the averages and the distributions of grades across different disciplines. In the following 23 years, in most disciplines, there was significant grade inflation, and the degree of inflation was far from constant, with the result that in 1985-86 the authors could readily identify, at a particular school, *high grading* and *low grading* departments. Anglin and Meng (2000) present similar results for universities in Ontario, Canada. They look at grades in introductory courses in 12 disciplines in seven universities for academic years 1973-74 and 1993-94, and they also find evidence of significant grade inflation and of significant divergence in grades across departments during this 20 year period. In 1993-94, in the three softest grading departments (English, French, and Music), more than 60% of all students received an A or a B, and less than 13% received a D or an F, while in the three hardest grading departments (Chemistry, Mathematics, and Economics), fewer than 45% of all students received an A or

by course level, by year. The following anecdotes suggests that it might just be the case that this reluctance stems for the recognition that there is a standards problem, and that it would be difficult to address. In the mid 90s in one Canadian university, extensive data on grades awarded in different disciplines was circulated to faculty members. People in some disciplines in which grades were relatively low were outraged by what they saw, because the data suggested to them that, by offering easy As, other disciplines were attracting students who otherwise would have enrolled in their own courses. As you might imagine, people in the "offending" disciplines had quite a different interpretation of the facts. Apparently, the discussion was heated, and the net result was that the university never again made such data available. We know of one university that routinely collects grades data of exactly the sort that is used in this study—but the data is circulated only to a very short list of administrators. Why? Perhaps because it is perceived that if the data were more widely circulated, there would be pressure to examine standards and perhaps rectify them.

a B, and more than 30% received a D or an F.

Various explanations for grade inflation, and especially for the differences in grade inflation across disciplines, have been suggested. Anglin and Meng (2000) offer the hypothesis that grade inflation is the consequence of competition among departments for students. Freeman (1999) offers a variation on this theme, namely that departments manage their enrolments by adopting grading standards commensurate with the market benefits of their courses, and presents supporting evidence. Siegfried and Fels (1979) and Nelson and Lynch (1984) examine the hypothesis that grade inflation results from the attempts of instructors to improve their teaching evaluations by offering students higher grades, and conclude that there is some evidence to support the hypothesis. Zangenehzadeh (1998) offers further evidence in support of this hypothesis and a suggested solution.

In addition, there has been some discussion of the distortions that may be associated with differential grading standards and grade inflation. Sabot and Wakeman-Linn (1991) examine student responses to differential grades and conclude that, in choosing their courses, students clearly do respond, and in the expected way. Both Sabot and Wakeman-Linn (1991) and Anglin and Meng (2000) discuss the distorted incentives that arise when grading standards differ across departments.

2 Grades Data for Three Universities

We have times series data on grades by department and course levels, for three Canadian universities: 10 years of data for University #1 (1996 through 2005), 8 years of data for University #2 (1997 through 2005, excluding 2002), and 7 years of data for University #3 (1999 through 2005). In our aggregation, in Universities #1 and #2 we separate the courses into junior level (1^{st} and 2^{nd} year) courses and senior level (3^{rd} and 4^{th} year) courses. In University #3, upper and lower levels refer to 400 and 200 level courses, respectively. To exclude very small departments, we used the data for any department at either level only if the department assigned at least 500 grades in at least one year.

We report on the *percentage of high grades* awarded and, to capture differences in the entire grade distribution, also on the *average grade* awarded by discipline. All together 53 disciplines are represented, and 18 of the disciplines are included at both the junior and senior levels at all three universities.[2] Fitting a simple time trend to the university average grades reveal that there has been statistically significant grade inflation in all three universities over the periods for which we have data. This inflation has been substantial in Universities #2 and #3, especially at the junior level, but somewhat less so at University #1.[3]

As a convenient way of summarizing the data for our purposes, we ran a number of OLS regressions of the following sort:

$$y_{it} = \beta_i D_i + \epsilon_{it}, \tag{1}$$

where i denotes the discipline, and t denotes time. The left hand side variable, y_{it}, is the deviation in period t from the corresponding university mean of a measure that captures the grades awarded by discipline i. We use two different measures: (a) the percentage of high grades (As) awarded by the discipline, and (b) the average grade awarded by the discipline. The average grade is measured on a 4-point scale for Universities #1 and #2, and on a 9-point scale in University #3. On the right hand side of (1), D_i is a dummy variable for discipline i (that is, $D_i = 1$ if y_{it} pertains to discipline i and $D_i = 0$ otherwise) , and ϵ_{it} is a normally distributed error term with mean zero. The coefficient β_i estimates the extent to which our measure of the grades in discipline i deviates from the university average. In an ideal world, this coefficient would reflect differences in student achievement across disciplines; but

[2] We refer to the units for which we have data as 'disciplines' with a slight abuse of terminology, because our data pertains to Departments. Sometimes the same discipline is taught in different Departments, and for some Departments it is not easy to define the dominant discipline.

[3] The average grade point at the junior level has been increasing at 0.013 per year on a 4.0 scale in University #2, 0.053 per year on a 9.0 scale in University #3, and 0.0027 in University #1.

it may also reflect differences in grading standards. In general, we expect that the coefficient conflates both these effects.

The regression results for (a) are presented in the second and third columns of the following six Tables (Table 1.1 through Table 3.2), while those for (b) are presented in the fourth and fifth columns. The adjusted R squared statistics are reported in the second last line of each table, and they indicate that the discipline dummies explain a great deal of the variation in grades within a university, particularly for senior level courses. It is apparent from these tables that grades differ significantly, in some cases dramatically, across disciplines within each of these universities, and this is the phenomenon that we want to call attention to.

In discussing differences in grades across disciplines, we focus on the percentage of high grades awarded: for Universities #1 and #2, the percentage of grades that are either A^+, A, or A^-; for University #3, the percentage of A^+ grades awarded, because it is the only available index of high grades for this university. The patterns are similar for average grades.

Results for University #1 are reported in Table 1.1 for senior level courses and in Table 1.2 for junior level courses. The tables are organized to highlight differences in the percentage of As given in different disciplines. The beta coefficient in the second (fourth) column measures the divergence in the percent of As (average grade) given in that discipline from the university mean for all disciplines combined. The third and fifth columns give the corresponding T-statistics. In the Tables, the disciplines are ranked from lowest percentage of As to highest percentage of As, and disciplines are somewhat arbitrarily grouped into 'Low', 'Medium', and 'High' grade groups. In the Low Grade group, the percent of As is no larger than 75% of the average for all disciplines; in the High Grade group the percentage of As is no smaller than 125% of the average; the rest are in the Medium Grade group. The asterisks in the Table indicates discipline dummies that are significantly different from 0 at the 5% level of significance.

The university mean of the percentage of As awarded in senior level courses in University #1, call this average $\overline{PA}$, is 31.6%.[4] Denote the average percentage of As awarded in discipline i by PA_i. At one extreme, in French this percentage is 13.8 percentage points lower than the university average. At the other extreme, in Engineering Science this percentage is 21.5 percentage points higher. Notice that for 11 of the 27 disciplines, PA_i is less than $\overline{PA}$ at the 5% level of significance, and that for 9 of 27 disciplines PA_i is greater than $\overline{PA}$ at the 5% level of significance. Notice also that for 4 of the 27 disciplines $PA_i < 0.75\overline{PA}$, and that for 4 of 27 disciplines $PA_i > 1.25\overline{PA}$; by this criterion, nearly one third of all disciplines have grades that are either anomalously high or low. And this picture, dominated more by disparity than by similarity, is much the same within Faculties or disciplinary groupings.[5]

As regards high grades given in senior level courses at University #1, the picture that emerges from Table 1.1 is one of *disparity*, *dissimilarity*, and *divergence*. The picture is much the same when one focuses on the average grade given. The university average (over all the disciplines) is 3.01 on a 4.0 scale, and the range over disciplines extends from 2.54 to 3.39, and for three-quarters of the 27 disciplines the average grade is significantly different from 3.01 at the 5% level of significance. The picture is similar, though not quite so dramatic, for junior level courses at this university (Table 1.2).

The reader can peruse Tables 2.1 through 3.2 for the remaining two universities and persuade herself that disparity, dissimilarity, and divergence dominate these pictures as well. Furthermore, she will notice that even for a *given discipline* there is considerable variation

[4] The average percentage of As awarded in senior level (junior level) courses for Universities #1 and #2 over the period for which we have data are, respectively, 31.6% (20.3%), and 32.5% (23.5%). The average percentage of A^+s awarded in senior level (junior level) courses for Universities #3 is 9.6% (5.6%).

For Universities #1 and #2, the average grade awarded in senior level (junior level) courses over the period for which we have data are, respectively, 3.01 (2.68), and 3.03 (2.73) on a 4-point scale. For University #3, the average grade in the senior level (junior level) courses is 6.42 (5.14) on a 9-point scale.

[5] For these groupings, consult Table 4. below.

in the grades handed out in different universities. For example, in junior level philosophy courses the percentage of As given out was 15.7% in University #1, while it was 28.6% in University #2. In junior level kinesiology courses the percentage of As given out was 25.6% in University #1, while it was 37.8% in University #2.

In Table 4, we have somewhat arbitrarily grouped disciplines into six different disciplinary groups: Applied Science (APSC, 10 disciplines), Applied Social Science (APSS, 6 disciplines), Business (BU, 5 disciplines), Humanities & Arts (HA, 13 disciplines), Pure Science (PS, 11 disciplines), and Social Science (SS, 8 disciplines).

In Table 5 we focus on anomalously low and high grades by discipline and disciplinary groups. The data in the table are count data on the number of times a discipline appeared in the Low Grade and High Grade categories in Tables 1.1 through 3.2. The maximum count is 6 (lower and upper level courses at each of three universities). Only the first 18 disciplines listed in Table 5 are represented at both the junior and senior levels in all three universities, and in this sense they can be regarded as the *core disciplines*. Within the core disciplines anomalies are immediately apparent. History, for example, registers 6 counts in Low Grade and none in High Grade. Education registers only one count in Low Grade and 4 in High Grade.

In the core there are four Pure Science disciplines, five Social Science disciplines, five Humanities & Arts disciplines, one Applied Science discipline, one Applied Social Science discipline, and one Business discipline. Overall, in 29 of 102 cases (28%) the core disciplines were in the Low Grade category and in 15 of 102 cases (15%) they were in the High Grade category. In the non-core disciplines, there were relatively fewer cases in the Low Grade category, 8 of 84 (9%), and relatively more in the High Grade category, 29 of 84 (34%). It is apparent then that grades tend to be quite a lot higher in the non-core disciplines. However, it is also apparent that within the core disciplines there is considerable variation in the frequency of anomalously low and high grades.

Clearly, students' grades measure their achievement in their courses. It is widely presumed that one can meaningfully aggregate or average students' achievements and compare these averages. Otherwise we would not routinely report GPA on transcripts, and we would not use GPA internally in the ways that we do—to award financial aid, to identify students whom we want to honour for their achievements, to determine who will be admitted to various programs and who will be rejected, to determine who can continue their course of studies and who must withdraw. Given the very significant disparities that we see in the data in grades awarded by different disciplines in these three universities, we must ask whether the disparities reflect differences in achievement or differences in grading standards. This would seem to be an important question. Because, to the extent the variation across disciplines in grades is due not to differences in achievement but to differences in standards, we are awarding scarce financial aid to the some of the wrong students, honouring some of the wrong students, admitting some of the wrong students to our programs, and encouraging some of the wrong students to continue their studies. More subtly, we are inducing students to enroll in the wrong programs, to take the wrong courses, and to allocate their time in socially inefficient ways. We elaborate on these important points in the following section.

Given the data we have it is impossible to discern the extent to which standards vary across disciplines. But, drawing on impressions formed over many years in a number of universities, we do not believe that most of the variation across disciplines in grades awarded is attributable to differences in achievement. We cite just a few examples. It seems implausible that the levels of achievement of students in junior level contemporary arts courses at University #1, who receive 14 percentage points more As than the university mean (20.3%), are significantly higher than the levels of achievement of students in junior level business administration courses at this university, who get 8 percentage points fewer As than the university mean;

it seems implausible that the levels of achievement of students in senior level kinesiology courses at University #2, who receive 18 percentage points more As than the university mean (32.5%), are significantly higher than the levels of achievement of students in senior level chemistry courses at this university, who receive 10 percentage points fewer As than the university mean; it seems implausible that the levels of achievement of students in senior level psychology courses at University #3, who receive 7 percentage points more A^+s than the university average (9.6%), are significantly higher than the levels of achievement of students in senior level economics courses at this university, who receive 1 percentage point fewer A^+s than the university mean.

For University #1 we do have some recent data on grades by Faculty that strongly suggest that standards differ substantially across Faculties at this university. In Table 6, grades data pertaining to the five Faculties listed in the third column of the table are presented. The Faculties are Applied Science (APSC), Arts & Social Science (A&SS), Business (BUS), Education (EDUC), and Science (SCI). The first column reports the percentage of As in all courses offered in the five Faculties in 2005, and the second column reports the percentage of As received by students who were registered in the five Faculties in 2005. The ordinal rankings along the columns of these percentages are given in parentheses. Notice that students registered in the Education Faculty (EDUC) received the highest proportion of As (58%), and that the highest proportion of As are given in courses offered in this Faculty (49%). Both proportions are far higher than in the other Faculties. Similarly, students registered in the Applied Science Faculty (APSC) get more As (30.0%) than do students in any Faculty other than Education, and more As are given in courses offered in this Faculty (29.7%) than in any Faculty other than Education.

Given the data in the first two columns of Table 6, we might be tempted to conclude that students in the Education and Applied Science Faculties are somewhat better than students in the other Faculties of this university. The fourth through eighth columns of the Table point to quite a different conclusion. These columns report data on the percentage of As awarded to students registered in each of the Faculties listed in the third column in courses they have taken in different Faculties. So 22.8% of the students registered in the Applied Science Faculty (APSC), for example, received As in courses offered by the Faculty of Arts & Social Science (A&SS), while they received As in 33.4% of the courses offered by their own faculty. In terms of ordinal ranking, students registered in the Faculty of Education are near the bottom of the heap in the courses they take in Science (SCI), where their rank is 4^{th} out of 5, and Applied Science (APSC), where their rank is 5^{th} out of 5, and although they do relatively well in the courses they take in Faculty of Arts and Social Science (A&SS), the percentage of As they get in these courses (28.6%) is far below the percentage they get in courses they take in their own Faculty (65.8%). It looks like Education students get a lot of As mostly because they take a lot courses in the Faculty of Education where lots of As are awarded. Similarly it looks like the high percentage of As received by students in Applied Science is largely attributable to the fact that they take a lot of courses in Applied Science where many As are awarded.

The above perceptions are reinforced by the observation that the cross-faculty grades data is driven mostly by cross over in junior level courses, where grades are systematically lower than they are in senior level courses. This means that the relatively high percentage of As that Education (or Applied Science) students get in Education (or Applied Science) courses is perhaps driven by the fact that they take many senior level courses in their own Faculties where grades are much higher than they are in the junior level courses where most of the cross over occurs. The data in Table 6 do not support the proposition that students in Education and Applied Science are high achievers, relative to students in other Faculties. Looking only at the second column, one would conclude that students in the Faculty of Business are modest achievers, in the middle of the pack. But when we look at the cross-Faculty comparisons, it is apparent that these students as a group are very high achievers, and the fact that they finish in the middle of the pack in column 2 seems to be attributable to the high standards in the Business Faculty.

One might argue that students registered in the Faculty of Education in this university have aptitudes specific to that Faculty, and so it is not a surprise that their performance in courses outside that Faculty is worse. The argument, however, is unlikely to be valid. Most Faculties contain many departments, each offering many courses. It is therefore unlikely that there is a dearth of courses outside of Education that require the skills that are specific to that Faculty.

To us the conclusion seems inescapable: there are very significant differences in grading standards across disciplines and Faculties in these three universities and, given the ways in which GPA is used in universities, possibly serious distortions throughout these universities. We have seen comparable grades data for just one year for three other Canadian universities, and there is nothing is these data to suggest that they also do not suffer from the same problems as the three universities examined here.

3 A Simple Human Capital Model

This section offers a simple model that brings out some of the distortions that arise when different disciplines within a university have different grading standards. Our model is in the tradition of economic models of grading standards [Becker (1975), Betts (1998), Becker and Rosen (1992), Costrell (1994), Dickson (1984), McKenzie (1975)], but our concerns are different. We focus on *differential* grading standards and their effects.

Students' preferences are identical and defined over the present value of lifetime income, y, and the effort expended, e, to develop their human capital. They are captured in the following utility function:

$$U(y,e) = y - \frac{1}{2}e^2. \tag{2}$$

The present value of the earnings generated by a unit of human capital is equal to β. The cost incurred in training a student is γ, which is independent of the effort exerted by the student, the student's potential, and the program chosen by the student. This cost is borne by the student.

There is a university that offers two programs, Program 1 and Program 2. Each student is described by an innate aptitude, or ability-pair (α_1, α_2), where α_1 is her aptitude for Program 1 and α_2 is that for Program 2. We shall also refer to α_1 and α_2 as Aptitude 1 and Aptitude 2, respectively. The density function describing the distribution of aptitudes among potential students is

$$f(\alpha_1, \alpha_2) > 0 \quad \alpha_1 \geq 0 \quad \alpha_2 \geq 0.$$

In the university, students expend effort to develop their human capital, and human capital, $h_i(e_i)$, as a function of student effort in Program i is given by

$$h_i(e_i) = \alpha_i e_i. \tag{3}$$

A given effort generates more human capital for students with higher aptitudes.

There is an outside option (which we dub Program 0) that is available to all students and offers a lifetime utility equal to ω.

3.1 Optimality

The *cost-benefit optimal allocation* maximizes the sum of realized utility over all students. Because there is no interaction among individuals, we can characterize the optimal allocation by maximizing the utility of each student. This, of course, implies that *cost-benefit optimal allocation* is *Pareto-optimal.* If student α is assigned to university Program i, her utility is just

$$U_i(e_i, \alpha_i) = (\beta\alpha_i e_i - \gamma) - \frac{1}{2}e_i^2. \tag{4}$$

The first term on the right, $(\beta\alpha_i e_i - \gamma)$, is lifetime income net of the cost of her education, and the second term, $\frac{1}{2}e_i^2$, is the utility cost of her effort. The effort level that maximizes total utility is

$$e_i^*(\alpha_i) = \beta\alpha_i, \tag{5}$$

the optimal quantity of human capital is

$$h_i^*(\alpha_i) = \beta\alpha_i^2, \tag{6}$$

and maximized utility is

$$U_i^*(\alpha_i) = \frac{1}{2}\beta^2\alpha_i^2 - \gamma. \tag{7}$$

Optimal effort, human capital, and maximized utility from either Program are increasing functions of the student's aptitude for that program. This is because the returns to effort are higher for students with greater aptitude.

The outside option (Program 0) will dominate university Program i if $\omega > U_i^*(\alpha_i)$, or if

$$\alpha_i < \underline{\alpha}\,, where \quad \underline{\alpha} \equiv \sqrt{2(\gamma+\omega)}/\beta. \tag{8}$$

The optimal assignment of students to programs is illustrated in Figure 1, which displays the aptitude pairs (α_1, α_2) of students. The ray OA is the 45^0 line. Students below the line have better abilities for Program 1; students above OA for Program 2; and students along OA have equal abilities for both programs. Students with low abilities, however, cannot generate sufficient human capital to justify foregoing their outside option. Those with α_1 less than OB ($= \underline{\alpha}$) would be better off with their outside option compared to enrolling in Program 1. Likewise, students with α_2 less than OD ($= \underline{\alpha}$) would be better off with their outside option compared to enrolling in Program 2. Thus it is optimal for students with ability pairs in the square OBCD to not enroll in university. It will be optimal for students with ability pairs to the right of BC and below CA to enroll in Program 1, and for those with ability pairs above DC and CA to enroll in Program 2. Among students enrolled in either of the university programs, the achieved level of human capital will be increasing in their aptitude for that program.

3.2 Equilibrium

We treat the university as an intermediary in the market for human capital. It provides the opportunity for students to develop their human capital, observes their realized human capital, and assigns grades that reflect their achievements. Assigned grades are then used by employers to assess students' human capital. In this framework, the grading standards in Programs 1 and 2 are, in essence, the price system that guides students' choices.

The informational assumptions in our model are the following. Prior to enrolment, the university knows nothing about students, and it accepts all students who choose to enrol. The university observes the realized human capital of students when they complete their programs, but not their true aptitudes, and it assigns grades that reflect these achievements. Students know their own aptitudes, the utility of the outside option (ω), the present values of the student's worth per unit of human capital (β), and the cost of a university education (γ). Employers cannot observe human capital, and instead rely on assigned grades to assess students' human capital. That employers never observe the human capital achievements of their employees is, obviously, a strong assumption, but it will be apparent that none of the qualitative features of the distortions that we highlight are driven by it as long as employers cannot immediately assess their new hires.

In addition, we assume that student's pay the cost of their university education, and that the labor market is competitive, so that students in aggregate capture as earnings all that they produce.

The role of the university is then to educate students, observe their human capital achievements, and assign grades. The questions we ask concern the university's grading standards. What are the properties of optimal standards? What distortions arise when non-optimal standards are used?

3.2.1 Grading Standards

A *grading standard* for program i is a mapping from the realized human capital, h_i, of a student who has completed the program to a grade or mark, $m_i(h_i)$. A student's grade in Program i is a linear function of her human capital:

$$m_i(h_i) = \theta_i \, h_i \qquad \theta_i \geq 1. \tag{9}$$

The parameter θ_i is an inverse measure of the grading standard in Program i. We assume that Program 2 has the lower standard and, since it is *differences* in standards that distort student incentives (as we shall see), we fix the standard in Program 1. Accordingly, we assume that $\theta_1 = 1$. Notice that, in Program 1, a student's grade is identical to her human capital. Relative to Program 1, standards in Program 2 are possibly diluted ($\theta_2 \geq 1$). In particular, except in the case where $\theta_2 = 1$, a student's grade in Program 2 is less than her human capital.

In this environment, differential grading standards distort student incentives in that students are induced to manage their grades instead of their human capital.

3.3 Equilibrium Choices

Since firms cannot observe human capital, they rely on the grades provided by the university to infer the human capital of students whom they hire. Hence, in the labour market students are pooled by their grades. And, if programs use different grading standards, students in the same grade pool will have different amounts of human capital. Pooling across different disciplines, especially for students with only a Bachelors degree, is a very reasonable scenario. An entry level job in the banking sector, for example, requires generic skills that could be satisfied by students with degrees in Economics, Psychology, Mathematics, English, Physics or Business, to name just a few disciplines.[6]

Some circumstantial evidence can be brought to bear on our claim that students are pooled in the labour market. Sabot and Wakeman-Linn (1991) have examined the effect of a high grade in an introductory course on the probability of subsequent enrollment by the student in that discipline's courses. They find that a high grade in a particular course induces students to choose subsequent courses in that discipline with higher probability even after they account for the comparative advantage information contained in the grade (by evaluating their relative position in that course in comparison to the relative ranking of their overall GPA). The inducement offered by the absolute grade suggests that students expect that potential employers cannot glean full information about their achievement through their transcript—presumably because of pooling.

Let $\overline{h}(m)$ denote the average human capital acquired by the students who receive grade m. Then, since firms pool students with identical grades, lifetime earnings of a student who receives grade m will be $\beta\overline{h}(m)$. To achieve grade m in Program i, a student must acquire m/θ_i units of human capital, which requires that she expend $m/(\alpha_i\theta_i)$ units of effort. Hence, the typical student's maximization problem is to choose a program i $(i = 1, 2)$, and a grade m, to maximize her utility:

$$\max_{i,m} \quad \beta\overline{h}(m) - \gamma - \frac{1}{2}(m/(\alpha_i\theta_i))^2. \tag{10}$$

A student can get grade m in either program and, since earnings net of the cost of an

[6] In jobs that only require high school graduates, the pooling is even coarser. Bishop (1988) argues that employers only look for information on years of schooling, area of specialization, and certification; they do not seek to identify competence or level of achievement (by asking for transcripts or references).

education, $(\beta\overline{h}(m)-\gamma)$, are independent of the program chosen, a student will be indifferent between getting grade m in Program 1 or in Program 2 if the effort required to get grade m is identical in the two programs, that is, if

$$\frac{m}{\alpha_1} = \frac{m}{\alpha_2\theta_2}.$$

Hence, the locus of indifference is

$$\alpha_2 = \alpha_1/\theta_2. \tag{11}$$

This locus is represented in Figure 2 as dashed ray OE, which has a slope less than 1 if $\theta_2 > 1$. Let us isolate on this ray a student, represented by point F, with the aptitude pair $\alpha \equiv (\alpha_1, \alpha_2)$ such that the grade m solves her maximization problem. She is indifferent between getting grade m in Program 1 and getting grade m in Program 2 . Students $\widehat{\alpha} = (\widehat{\alpha}_1, \widehat{\alpha}_2)$ on the horizontal line segment HF (with $\widehat{\alpha}_2 = \alpha_2, \widehat{\alpha}_1 < \alpha_1$) are just as able as student F in Program 2 but less able in Program 1. Consequently, all these students will choose to get grade m in Program 2. In other words, if opting for grade m in Program 2 solves the maximization problem for student α, it also solves the maximization problem for every student $\widehat{\alpha} = (\widehat{\alpha}_1, \widehat{\alpha}_2)$ with $\widehat{\alpha}_2 = \alpha_2$ and $\widehat{\alpha}_1 < \alpha_1$. Analogously, since grade m also solves the maximization problem for student α in Program 1 (because, by assumption, she is indifferent between the two programs), this Program will also solve the maximization problem of every student $\widehat{\alpha} = (\widehat{\alpha}_1, \widehat{\alpha}_2)$ with $\widehat{\alpha}_1 = \alpha_1$ and $\widehat{\alpha}_2 < \alpha_2$. These students have aptitude pairs lying on the vertical line GF.

The grade pool m will comprise all the students with ability pairs on the lines GF and HF. Students on GF all have aptitude α_1 for the program they've enrolled in (Program 1), while on HF all students have an aptitude α_1/θ_2 for the program they've enrolled in (Program 2). By integrating the density function $f(\widehat{\alpha}_1, \widehat{\alpha}_2)$ over the line segments FG and HF in Figure 2, we can compute the proportions of students from Programs 1 and 2, $P_1(\alpha_1, \theta_2)$ and $P_2(\alpha_1, \theta_2)$

respectively, in grade pool m.[7]

We can now determine the average level of human capital in any grade pool. To earn grade m, a student in Program 1 must acquire m units of human capital, and a student in Program 2 must acquire m/θ_2 units. So the average human capital of students in grade pool m is

$$\overline{h}(m) = mP_1(\alpha_1, \theta_2) + (m/\theta_2)P_2(\alpha_1, \theta_2),$$

or

$$\overline{h}(m) = mD(\alpha_1, \theta_2), \tag{12}$$

where

$$D(\alpha_1, \theta_2) \equiv P_1(\alpha_1, \theta_2) + (1/\theta_2)P_2(\alpha_1, \theta_2). \tag{13}$$

Note that, since $\theta_2 \geq 1$, $D(\alpha_1, \theta_2) \leq 1$. Furthermore, $D(\alpha_1, \theta_2) = 1$ if $\theta_2 = 1$ and is declining in θ_2. Consequently, the average amount of human capital in grade pool m is less than or equal to m. If there is dilution of grading standards in Program 2 ($\theta_2 > 1$), this average is strictly less than m because the human capital that students from Program 2 have is less than m.

There is a continuum of students, so no individual student contributes a measurable amount to the human capital in any grade pool. Consequently, in choosing their grade pool m, each student will take the proportions $P_1(\alpha_1, \theta_2)$ and $P_2(\alpha_1, \theta_2)$, and hence $D(\alpha_1, \theta_2)$, as given. Therefore, from the perspective of an individual student, the rate of change of lifetime

[7] These proportions are given by

$$P_1(\alpha_1, \theta_2) = \frac{\int_0^{\alpha_1/\theta_2} f(\alpha_1, \widehat{\alpha}_2) d\widehat{\alpha}_2}{\int_0^{\alpha_1/\theta_2} f(\alpha_1, \widehat{\alpha}_2) d\widehat{\alpha}_2 + \int_0^{\alpha_1} f(\widehat{\alpha}_1, \alpha_1/\theta_2) d\widehat{\alpha}_1}$$

and

$$P_2(\alpha_1, \theta_2) = \frac{\int_0^{\alpha_1} f(\widehat{\alpha}_1, \alpha_1/\theta_2) d\widehat{\alpha}_1}{\int_0^{\alpha_1/\theta_2} f(\alpha_1, \widehat{\alpha}_2) d\widehat{\alpha}_2 + \int_0^{\alpha_1} f(\widehat{\alpha}_1, \alpha_1/\theta_2) d\widehat{\alpha}_1}.$$

earnings with respect to grade m is

$$\frac{\partial \beta \overline{h}(m)}{\partial m} = \beta D(\alpha_1, \theta_2).$$

Using this, the first order condition with respect to m for a student who enrols in Program 1 is just

$$\beta D(\alpha_1, \theta_2) - \frac{m}{\alpha_1^2} = 0,$$

which can written as

$$m = \alpha_1^2 \beta D(\alpha_1, \theta_2). \tag{14}$$

This equation identifies the grade m that solves the maximization problem of students who enrol in Program 1 in terms of her aptitude for that program. By substituting $\alpha_1 = \alpha_2 \theta_2$ in the above equation, we obtain the grade m that solves the maximization problem of the students enrolling in Program 2 in terms of her aptitude for that program as

$$m = (\alpha_2 \theta_2)^2 \beta D(\alpha_2 \theta_2, \theta_2). \tag{15}$$

The locus OE in Figure 3 partitions the (α_1, α_2) space into students who prefer Program 1 to Program 2 and vice versa. Equations (14) and (15) pin down the grade m achieved by students in terms of their aptitude for the program they enrol in. This grade, naturally, is increasing in the aptitude for the program in which the student is enrolled.

We can use Figure 3 to identify the equilibrium allocation of students to programs and compare this outcome with the optimum. In the figure as drawn, we are assuming that standards in Program 2 are diluted; that is, that $\theta_2 > 1$. We have seen that in the optimum, students with ability less than OB would opt for their outside option rather than enroll in Program 1. In equilibrium, students in Program 1 earn lower returns on their human capital since they are pooled with students from Program 2, which has lower standards for the grades of its students. Consequently, the ability cut-off for students to forego their outside option in order to enroll in Program 1 is higher in equilibrium than in the optimum and is denoted by OI in Figure 3. Students in Program 2 are pooled with Program 1 students, who have higher human capital. Thus students in Program 2 receive a higher return from enrolling in Program 2 than is warranted by their human capital. Consequently, the ability cut-off in equilibrium (denoted by OK) for students to forego their outside option in order to enroll in Program 2 is lower than that in the optimum (OD). In equilibrium, students with aptitude pairs falling in the rectangle OIJK will opt for Program 0 (their outside option). Students with aptitude pairs falling in the region to the right of IJ and below JE will enroll in Program 1; those with aptitude pairs lying above KJ and JE will enroll in Program 2.

3.4 Equilibrium versus Optimum

Many authors, including Sabot and Wakeman-Linn (1991) and Anglin and Meng (2000), have argued that identical grading standards are, in some sense, optimal. In our model, identical grading standards ($\theta_2 = 1$) will lead to the optimal allocation. We see from (11) that ray OE will coincide with ray OA in Figure 3 when $\theta_2 = 1$. Also, we see from (13) that $D(\alpha_1, 1) = 1$ and so (12) says that the average human capital embodied in a grade pool coincides with the grade. In Figure 3, rectangle OIJK coincides with rectangle OBCD. Even though students from different programs are pooled, their grade is a perfect indicator of their productivity. Even if we allowed for a dilution of the grading standard in Program 1 ($\theta_1 > 1$), we see from the logic of the model that, as long as the degree of dilution is identical for the two programs ($\theta_1 = \theta_2$), grades convey perfectly all the information that is needed to achieve the optimal allocation within the university.

Various authors have commented on the distortions with respect to program choice one should expect to see when grading standards differ [see especially the discussion in Anglin and Meng (2000)]. Three sorts of distortion are apparent from Figure 3:

(i) Optimality dictates that students in area KLCD stay out of university (Program 0), but in equilibrium they choose Program 2. By choosing Program 2 instead of their outside option, these students are pooled with others from Program 1 who acquire more human capital that they do. Hence there is an implicit cross-subsidy that is sufficient to induce these students to choose Program 2 instead of Program 0.

(ii) Optimality dictates that students in area BIJL be allocated to Program 1, but in equilibrium they choose Program 0. The story here is the reverse of the one told above. If these individuals were to choose Program 1, they would be pooled with others from Program 2 who would acquire less human capital than they would, and the burden of the cross-subsidization that they would have to bear is so large that they are better off choosing their outside option instead of Program 1.

(iii) Optimality dictates that students in area ACLJE be allocated to Program 1, but in equilibrium they choose Program 2 instead. Once again the implicit cross-subsidy from pooling is driving the misallocation.

So, in this model, differential grading standards clearly distort student incentives with respect to program choice. In addition, whenever students from different programs are pooled their incentives to acquire human capital within their chosen program are distorted. Since at any grade students are pooled from Programs 1 and 2, the returns to effort are reduced for students in Program 1 and raised for those in Program 2. Thus students in Program 1 will underachieve in equilibrium relative to their optimal achievement, whereas those in Program 2 will overachieve. The cross-subsidization makes students of Program 1 worse off in equilibrium and students of Program 2 better off. All these distortions are due to the wedge that is created by differential grading standards between true human capital and signaled human capital. Clearly, the only way in which these misallocations in equilibrium can be eliminated is by having identical grading standards, that is, by setting $\theta_2 = 1$.

We record the essential points of the above discussion.

1. If the grading standard in Program 2 is diluted ($\theta_2 > 1$):

    - Some students who should optimally stay out of university (Program 0) and some who are optimally allocated to Program 1 choose Program 2, and some students who are optimally allocated to Program 1 choose to stay out of university.
    - When students from Programs One and Two are pooled, the students from Program 1 acquire less human capital than is optimal and the students from Program 2 acquire more human capital than is optimal.
    - All students who are optimally allocated to Program 1 are worse off in equilibrium, and those who are optimally allocated to Programs 0 and 2 but are pooled with students from Program 1 in equilibrium are better off.

2. The equilibrium allocation is optimal if and only if grading standards are identical.

By eliminating the cross-subsidy in the labour market associated with pooling workers with different amounts of human capital, identical grading standards eliminate the misallocation of students across programs in our model. In reality, however, students from different universities compete in the labour market. So even if the grading standards of all the disciplines within a given university are common, these common standards may differ across universities. If employers possess imperfect knowledge of the precise extent of this difference, there would be some pooling in the labour market of workers with different amounts of human capital. There would, then, obtain a misallocation of students across universities (even within a given discipline). In other words, while a common standard within a university sorts out the misallocation problem between its disciplines in our model, there would still obtain a misallocation across universities if the labour market cannot perfectly discern human capital differences.

In the formal analysis above, we have taken the grading standards as given. It would be possible to construct a model in which these standards are endogenized. Disciplines might be interested in increasing the number of students enrolled in their programs, since funding depends on the student-to-faculty ratio and Departments see research advantages to having more faculty members. As a first cut one might reasonably attribute to each discipline an objective function that is the difference between this enrollment-dependent funding and the costs of managing the program. Student enrollment depends on the grading schemes used in the two programs, as analyzed above. Such a model would predict that, if the standards in the two programs were chosen under Nash conjectures, the equilibrium level would entail a dilution of grading standards in both. This would rationalize an informal observation made by Shea (1994). He reports that, after the study of Sabot and Wakemen-Linn (1991) documenting differential grading practices in Williams College was published, one did not subsequently see stricter grading in the easy grading disciplines of this College. Rather, what transpired was an easing up of grading in mathematics and science (the traditionally hard grading disciplines). Grade inflation in these disciplines after 1985-86 (the last year for which Sabot and Wakeman-Linn presented data) was 9%, while that in arts and languages (the traditionally easy grading disciplines) was 5%.

If different Departments in a university make decentralized decisions about their grade distributions, there is thus reason to believe that dilution of grading standards will obtain. This is because such a grading policy is individually rational for each discipline. By luring away potential students, each Department inflicts a negative externality on others that it does not take into account. An adverse consequence of this is that, at each grade level, a student embodies less human capital.[8] Therefore, there are grounds for the university administration

[8] This would be true even if the standards end up being identical in the Nash equilibrium, as long as there is pooling in the labour market of students from different universities with differing standards (as discussed above). Using data from high schools in Florida, Figlio and Lucas (2004) have shown empirically that higher

to step in with measures that would internalize these externalities and prevent a race to the bottom. These measures need not be as draconian as imposing strict, common guidelines for grade distributions in all disciplines; reporting students' percentile ranking in their courses would go some distance in resolving this problem.

4 Conclusions

We analyzed here time series data on grades awarded in three Canadian universities. The data covers periods that begin in the mid 90s and range in length from 5 to 10 years. In all three universities we see very substantial and persistent differences across disciplines in both the average grade awarded and in the number of high grades awarded. In principle, differences in grades across disciplines can be decomposed into a component driven by differences in achievement levels, attributable one imagines to differences in the academic potential of students that different disciplines attract, and a component driven by differential grading standards, that is, by differences across disciplines in the way in which observed achievement in a course is mapped to a grade. Although the data that is available to us do not permit this decomposition, it appears to us that much, perhaps most, of the differences in grades across disciplines within each of these universities is attributable to differences in grading standards. We have seen fragmentary data for a number of other Canadian universities, and in each case the data supports the same conclusion. In addition, differential grading standards are apparent in various studies of grade inflation [see especially Anglin and Meng (2000) and Sabot and Wakeman-Linn (1991)]. It appears to us that substantial differences in grading standards across disciplines is the norm in our universities.

Assuming for the moment that this assessment is correct, we think there is a serious

grading standards lead to better performance in state-wide test scores. So inferior standards do lead to worse outcomes in terms of human capital.

problem: to the extent that grading standards differ across disciplines, grades are not a legitimate unit of account as regards the academic achievement of students, and they cannot be legitimately used to award scholarships, honours and degrees, nor to ration access to courses, academic programs, and jobs in excess demand.

Where do we go from here? It might be argued that what is done in different disciplines is so different that comparisons of achievement across them are simply impossible. If this argument were accepted, then it would seem that we ought to abandon the pretense of comparability; that is, that we ought to abandon the practice of requiring different disciplines to report student's achievements to the Registrar using the same set of grades. This, we think, would cause more problems than it solves. If they are to compare the transcripts of different students, users of transcripts necessarily require some index of aggregate achievement, and if the university does not supply the index the users themselves will. At least implicitly, users will devise their own aggregate indices of achievement. This would seem to be unfortunate, because universities have far more information at their disposal on matters pertaining to the measurement and comparison of student achievement than do typical users of transcripts.

It seems to us that the first thing that needs to be done is to actually decompose differences in grades awarded across different disciplines into a component attributable to differences in standards and a component driven by differences in achievement. Only then will we actually grasp the extent of the problem. Only then will we have the information necessary to reform our grading practices. Importantly, universities have in their data banks the information that is necessary to tackle these issues.

Given the way in which universities operate, it would be difficult to impose uniform grading standards. But, it is not necessary that grading standards actually be uniform. If it is possible to estimate grading standards by discipline, and we think it is, then it is possible to construct an aggregate index of achievement that is purged of the biases associated with grading standards that differ across disciplines. This aggregate is what we ought to be calculating and reporting on transcripts, not GPA. Importantly, this one measure would appear eliminate at a stroke the incentives that Departments have to dilute their standards.

In the meantime, we would argue that universities ought now to calculate and report for each student the student's average relative performance, and that this ought to be given the same prominence that is now given to GPA on transcripts.[9] This could be as simple as an average of the student's percentile ranking in each of her courses. Users can then at least spot students who have a high GPA, not because their achievement is at the high end of the distribution of achievement in the classes that they take but because they are enrolled in courses where everyone gets a high mark. They can likewise spot students who have a low GPA, not because they are near the bottom of the distribution of achievement in the courses they take but because they are enrolled in courses where everyone gets a low mark.

In this paper, we have analyzed the consequences of differential grading standards. We have not examined the reason for differences in standards. Various hypotheses have been suggested in the literature for the latter [eg. Dickson (1984), Freeman (1999), McKenzie (1975), Nelson and Lynch (1984), Zangenehzadeh (1998)], and many of these have to do with the incentives of individual instructors. Even if one argues that disciplines like to maintain high ratios of students to teachers, a discipline-wide easy grading practise would need to be coordinated—an outcome that would succeed only to the extent that individual instructors in the discipline feel vulnerable. Differential grading standards may well be partly a consequence of the allocation procedures in place in most universities, whereby funding is tethered firmly to student enrollments. In this environment, as long as instructors' efficacy in teaching is assessed through student evaluations and as long as students' perceptions of instructors can be influenced by the grades they receive as the term progresses, dilution of grading standards is

[9] In espousing this view we agree with Dickson (1984), who argues that GPAs should be standardized by discipline.

a temptation. As with individual disciplines, so with individual instructors: it is individually rational not to hold out against a rising tide of grade inflation. Any university that is serious about improving the quality of its teaching has to stem this tide if it is to succeed.

Finally, we note that in the model of this paper we assumed that students know their innate abilities. In reality, few university students (especially undergraduates) have this degree of self-knowledge. They go to university to take a variety of courses and learn what they are good at. Then they choose a major. They implicitly expect their professors to inform them of their strengths through the grades they assign. Differential grading standards, by providing misinformation, betray that expectation. The resulting misallocation within universities we have discussed at length here. The greater tragedy of mismatched careers and missed callings perpetrated by this misinformation can only be imagined.

Bibliography

Anglin, P. M. and R. Meng (2000), "Evidence on Grades and Grade Inflation in Ontario's Universities", *Canadian Public Policy*, Vol. XXVI, No. 3, pp. 361-368.

Becker, W.E. (1975), "The University Professor as a Utility Maximizer and Producer of Learning, Research, and Income", *The Journal of Human Resources*, Winter, pp. 107-115.

Becker, W.E. and S. Rosen (1992), "The Learning Effect of Assessment and Evaluation in High School", *Economics of Education Review*, 11, pp. 107-118.

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

Bishop (1988), "Employment Testing and Incentives to Learn", *Journal of Vocational Behavior*, 33, pp. 404-423.

Costrell, R.M. (1994), "A Simple Model of Educational Standards", *American Economic Review*,84, pp. 956-971.

Dickson, V.A. (1984), "An Economic Model of Faculty Grading Practices", *Journal of Economic Education*, Summer, pp. 197-203.

Figlio, D. N. and M. E. Lucas (2004), "Do Grading Standards Affect Student Performance?", *Journal of Public Economics*, 88, pp. 1815-1834.

Freeman, D. G. (1999), "Grade Divergence as a Market Outcome", *Journal of Economic Education*, Fall 1999, pp. 344-351.

McKenzie, R.B. (1975), "The Economic Effects of Grade Inflation on Instructor Evaluations: A Theoretical Approach", *Journal of Economic Education*, 6, pp. 99-106.

Nelson, J. P. and K. A. Lynch (1984), "Grade Inflation, Real Income, Simultaneity, and Teaching Evaluations", *Journal of Economic Education*, Winter 1984, pp. 21-37.

Sabot, R. and J. Wakeman-Linn (1991), “Grade Inflation and Course Choice”, *The Journal of Economic Perspectives*, Winter 1991, pp. 159-170.

Shea, C. (1994), “Grade Inflation’s Consequences”, *Chronicles of Higher Education*, 40 (18), A 45-46.

Siegfried, J. J. and R. Fels (1979), “Research on Teaching College Economics: A Survey”, *Journal of Economic Literature*, September 1979, pp. 923-969.

Spence, M. (1973), “Job Market Signaling”, *Quarterly Journal of Economics*, August 1973, pp. 355-379.

Stiglitz, J. (1975), “The Theory of Screening, Education, and the Distribution of Income”, *The American Economic Review*, June 1975, pp. 283-300.

Zangenehzadeh, H. (1998), “Grade Inflation: A Way Out”, *Journal of Economic Education*, Summer 1988, pp. 217-231.

Table 1.1

Senior Level: University # 1

	Deviation of Percentage A from Univ. Mean (31.6%)		Deviation of Average Grade from Univ. Mean (3.01 / 4.0)	
	Beta	**T-Stat.**	**Beta**	**T-Stat.**
Low Marks				
French	-13.81*	-11.60	-0.11*	-3.98
Math & Computer Science	-9.99*	-8.39	-0.47*	-16.69
Economics	-9.74*	-8.18	-0.27*	-9.47
History	-9.59*	-8.05	-0.13*	-4.58
Medium Marks				
Business Administration	-7.63*	-6.41	-0.08*	-2.96
Statistics	-7.51*	-6.31	-0.33*	-11.76
Sociology	-6.19*	-5.20	-0.03	-1.06
Geography	-6.03*	-5.06	-0.05	-1.69
Mathematics	-4.62*	-3.88	-0.29*	-10.14
Computer Science	-3.80*	-3.19	-0.07*	-2.39
Psychology	-2.93*	-2.46	-0.14*	-4.82
Philosophy	-2.21	-1.86	-0.15*	-5.39
Linguistics	-1.67	-1.40	-0.10*	-3.66
Biological Sciences	-0.62	-0.52	-0.05	-1.90
English	-0.49	-0.41	0.07*	2.43
Chemistry	0.36	0.30	-0.07*	-2.61
Criminology	0.71	0.60	0.11*	3.77
Political Science	0.74	0.62	-0.01	-0.42
Mol. Biology & Biochem.	4.60*	2.73	0.06	1.49
Communication	4.65*	3.91	0.17*	5.95
Kinesiology	4.78*	4.01	0.09*	3.17
Archaeology	6.16*	5.17	0.12*	4.15
Physics	7.11*	5.97	0.02	0.70
High Marks				
Humanities	8.62*	7.24	0.20*	6.90
Contemporary Arts	18.05*	15.16	0.35*	12.46
Education	20.91*	17.56	0.38*	13.34
Engineering Science	21.48*	18.04	0.38*	13.41
Adjusted R-Square	0.85		0.84	
Number of Observations	265		265	

* denotes significance at 5% level

Table 1.2

Junior Level: University # 1

	Deviation of Percentage A from Univ. Mean (20.3%)		Deviation of Average Grade from Univ. Mean (2.68 / 4.0)	
	Beta	**T-Stat.**	**Beta**	**T-Stat.**
Low Marks				
Business Administration	-8.03*	-8.23	-0.14*	-5.51
Economics	-5.47*	-5.61	-0.25*	-9.51
Criminology	-5.24*	-5.37	0.05	1.81
History	-5.07*	-5.2	-0.02	-0.85
Medium Marks				
Philosophy	-4.57*	-4.68	-0.23*	-8.78
Sociology	-4.03*	-4.13	0.13*	4.89
Psychology	-3.73*	-3.82	-0.06*	-2.46
Communication	-3.10*	-3.18	0.14*	5.2
English	-2.03*	-2.08	0.18*	6.78
Geography	-1.83	-1.88	0.05	1.77
Biological Sciences	-1.75	-1.79	-0.18*	-6.97
Mathematics	-1.58	-1.62	-0.30*	-11.48
French	-1.2	-1.23	0.16*	6.28
Statistics	-1.18	-1.21	-0.22*	-8.55
Chemistry	-0.66	-0.68	-0.05*	-2.04
Math & Computer Science	-0.39	-0.4	-0.22*	-8.55
Linguistics	-0.06	-0.06	0.10*	4.01
Political Science	0.12	0.12	0.02	0.69
Mol. Biology & Biochem.	0.26	0.19	-0.20*	-5.39
Physics	2.55*	2.61	-0.03	-1.27
Archaeology	2.82*	2.89	0.10*	3.89
Computer Science	4.83*	4.95	-0.04	-1.39
High Marks				
Kinesiology	5.31*	5.44	0.08*	3.24
Education	10.39*	10.65	0.28*	10.86
Humanities	13.35*	13.68	0.39*	14.98
Contemporary Arts	13.71*	14.05	0.48*	18.68
Engineering Science	28.48*	29.19	0.60*	23.03
Adjusted R-Square	0.86		0.88	
Number of Observations	265		265	

* denotes significance at 5% level

Table 2.1
Senior Level: University # 2

	Deviation of Percentage A from Univ. Mean (32.5%)		Deviation of Average Grade from Univ. Mean (3.03 / 4.0)	
	Beta	**T-Stat.**	**Beta**	**T-Stat.**
Low Marks				
Law	-20.758*	-13.99	-0.12*	-3.77
Policy & Environment	-15.75*	-7.51	-0.07	-1.60
Engineering	-12.63*	-8.51	-0.46*	-13.95
French	-10.75*	-7.25	-0.15*	-4.53
Accounting	-10.63*	-7.16	-0.38*	-11.31
Chemistry	-10.00*	-6.74	-0.28*	-8.30
History	-8.38*	-5.65	-0.04	-1.13
Biology	-8.13*	-5.48	-0.24*	-7.17
Medium Marks				
Zoology	-7.63*	-5.14	-0.14*	-4.15
Economics	-7.38*	-4.97	-0.26*	-7.92
Strategy and General Management	-7.00*	-3.34	0.00	0.00
Linguistics	-6.75*	-4.55	-0.12*	-3.77
Mathematics	-6.75*	-4.55	-0.43*	-12.82
Finance	-6.75*	-4.55	-0.12*	-3.77
Geology	-6.50*	-4.38	-0.05	-1.51
Computer Science	-5.75*	-3.88	-0.15*	-4.53
Chemical & Petroleum Engineering	-5.63*	-3.79	-0.20*	-6.03
Physics	-5.00*	-3.37	-0.20*	-6.03
Biochemistry	-4.88*	-3.29	-0.26*	-7.92
Sociology	-4.75*	-3.20	-0.09*	-2.64
Political Science	-3.63*	-2.44	-0.03	-0.75
Mechanical Engineering	-3.63*	-2.44	-0.11*	-3.39
English	-3.13*	-2.11	0.00	0.00
Civil Engineering	-2.88	-1.94	-0.03	-0.75
Anthropology	-1.63	-1.10	-0.01	-0.38
Electrical & Computer Eng.	-1.38	-0.93	-0.10*	-3.02
Marketing	-0.63	-0.42	0.15*	4.53
Geography	-0.50	-0.34	0.01	0.38
Archaeology	1.63	1.10	-0.01	-0.38
Art History	2.00	1.07	0.00	0.00
Cell., Mol. & Microb. Biology	2.25	1.52	-0.06	-1.89
Religious Studies	2.50	1.69	0.13*	3.77
General Business	2.88	1.94	0.16*	4.90
Psychology	6.63*	4.47	0.08*	2.26

High Marks				
Communication Studies	8.38*	5.65	0.29*	8.67
Philosophy	9.13*	6.15	0.23*	6.79
Social Work	15.38*	10.37	0.36*	10.94
Art	17.63*	11.88	0.38*	11.31
Kinesiology	18.38*	12.39	0.31*	9.43
Nursing	21.38*	14.41	0.39*	11.69
Medicine	24.50*	8.26	0.35*	5.28
Drama	25.25*	17.02	0.39*	11.69
Education	26.13*	17.61	0.46*	13.95
Music	30.13*	20.31	0.46*	13.95
Adjusted R-Square	0.88		0.86	
Number of Observations	335		335	

* denotes significance at 5% level

Table 2.2
Junior Level: University # 2

	Deviation of Percentage A from Univ. Mean (23.5%)		Deviation of Average Grade from Univ. Mean (2.73 / 4.0)	
	Beta	**T-Stat.**	**Beta**	**T-Stat.**
Low Marks				
History	-9.63*	-3.55	0.00	0.00
Chemistry	-7.50*	-2.77	-0.19*	-2.92
Linguistics	-7.25*	-2.68	-0.14*	-2.14
English	-6.75*	-2.49	0.08	1.17
General Business	-6.75*	-2.49	0.16*	2.53
French	-6.13*	-2.26	0.02	0.39
Medium Marks				
Computer Science	-4.25	-1.57	-0.14*	-2.14
Mathematics	-4.13	-1.52	-0.35*	-5.46
Geology	-3.88	-1.43	-0.06	-0.97
Political Science	-3.25	-1.20	-0.04	-0.58
Sociology	-3.13	-1.15	-0.04	-0.58
Astronomy	-2.88	-1.06	-0.03	-0.39
Archaeology	-2.25	-0.83	-0.03	-0.39
Economics	-2.13	-0.78	0.00	0.00
Biology	-1.63	-0.60	-0.05	-0.78
Statistics	-1.13	-0.42	-0.19*	-2.92
Geography	0.00	0.00	0.05	0.78
Engineering	0.63	0.23	-0.15*	-2.34
Anthropology	1.00	0.37	-0.06	-0.97
Social Work	1.00	0.37	0.30*	4.68
Physics	1.88	0.69	0.08	1.17
Psychology	2.63	0.97	0.09	1.36
Religious Studies	2.63	0.97	0.21*	3.31
Classical Studies	4.00	1.48	0.16*	2.53
Philosophy	5.13	1.89	0.14*	2.14
Nursing	5.63*	2.08	0.39*	6.04
High Marks				
Communication Studies	6.13*	2.26	0.26*	4.09
Medicine	6.14*	2.12	-0.03	-0.42
Art History	7.00*	2.04	0.18*	2.22
Kinesiology	14.25*	5.26	0.43*	6.63
Art	17.63*	6.51	0.54*	8.38
Drama	18.00*	6.65	0.50*	7.80
Education	24.88*	9.18	0.40*	6.24
Music	31.88*	11.77	0.64*	9.94
Adjusted R-Square	0.59		0.63	
Number of Observations	268		268	

* denotes significance at 5% level

Table 3.1
400 Level Courses: University # 3

	Deviation of Percentage A+ from Univ. Mean (9.6%)		Deviation of Average Grade from Univ. Mean (6.42 / 9.0)	
	Beta	**T-Stat.**	**Beta**	**T-Stat.**
Low Marks				
History	-8.21*	-5.96	-0.86*	-8.86
English	-6.70*	-4.86	-0.64*	-6.60
Business	-6.30*	-4.57	-0.26*	-2.69
Philosophy	-5.80*	-4.21	-0.81*	-8.29
Creative Writing	-5.26*	-3.81	0.06	0.59
Political Science	-4.26*	-3.09	-0.41*	-4.25
French	-2.97*	-2.15	-0.47*	-4.80
Medium Marks				
Economics	-1.27	-0.92	-1.16*	-11.93
Sociology	-1.00	-0.73	-0.21*	-2.20
Child and Youth Care	-0.54	-0.39	0.32*	3.25
Education	-0.16	-0.11	0.46*	4.68
Chemistry	0.01	0.01	-0.39*	-3.98
Biochemistry and Microbiology	0.04	0.03	-0.45*	-4.63
Public Administration	0.26	0.19	0.27*	2.77
Social Work	0.34	0.25	0.43*	4.45
Mechanical Engineering	0.61	0.45	-0.36*	-3.73
Geography	0.91	0.66	0.37*	3.76
Anthropology	1.21	0.88	-0.45*	-4.66
Biology	1.23	0.89	-0.06	-0.59
History in Arts	1.23	0.89	0.15	1.58
High Marks				
Nursing	3.74*	2.71	0.49*	5.04
Physics	3.77*	2.74	-0.73*	-7.50
General Engineering	4.56*	3.30	-0.47*	-4.77
Linguistics	5.03*	3.65	-0.01	-0.13
Theatre	5.34*	3.87	0.81*	8.30
Psychology	7.31*	5.30	0.34*	3.51
Eletrical and Computer Engineering	8.43*	6.11	-0.05	-0.56
Computer Science	10.17*	7.38	-0.09	-0.95
Music	13.71*	9.95	0.68*	6.94
Mathematics	22.23*	16.12	0.38*	3.86
Adjusted R-Square	0.75		0.78	
Number of Observations	210		210	

* denotes significance at 5% level

Table 3.2
200 Level Courses: University # 3

	Deviation of Percentage A^+ from Univ. Mean (5.6%)		Deviation of Average Grade from Univ. Mean (5.14 / 9.0)	
Low Marks	**Beta**	**T-Stat.**	**Beta**	**T-Stat.**
History	-4.64*	-4.89	-0.35*	-3.04
English	-4.01*	-4.23	0.15	1.28
Political Science	-3.34*	-3.52	-0.18	-1.53
Education	-3.03*	-3.19	0.70*	6.10
Creative Writing	-2.37*	-2.50	0.99*	8.67
Business	-2.29*	-2.41	-0.05	-0.40
Biochemistry	-2.14*	-2.26	-0.79*	-6.88
Biology	-1.61	-1.70	-0.15	-1.31
Medium Marks				
French	-1.16	-1.22	0.06	0.50
History in Arts	-1.16	-1.22	0.45*	3.91
Anthropology	-1.13	-1.19	0.31*	2.69
Social Work	-1.13	-1.19	0.80*	6.96
Sociology	-0.70	-0.74	0.34*	2.98
Geography	-0.34	-0.36	0.52*	4.55
Economics	-0.14	-0.15	-0.55*	-4.78
Chemistry	-0.11	-0.12	-0.76*	-6.59
Psychology	0.39	0.41	-0.08	-0.72
Theatre	0.46	0.48	1.08*	9.40
General Engineering	0.74	0.78	0.25*	2.17
Philosophy	1.01	1.07	-0.07	-0.61
High Marks				
Computer Science	1.73	1.82	-0.12	-1.03
Physics	1.83	1.93	-0.33*	-2.91
Mathematics	2.09*	2.20	-1.10*	-9.59
Linguistics	3.26*	3.43	0.54*	4.72
Mechanical Engineering	3.99*	4.20	-0.16	-1.43
Child and Youth Care	4.40*	4.64	0.99*	8.63
Music	5.00*	5.27	0.96*	8.37
Electrical and Computer Engineering	5.87*	6.19	0.27*	2.39
Adjusted R-Square	0.50		0.78	
Number of Observations	196		196	

* denotes significance at 5% level

TABLE 4

Key to Disciplinary Codes

Applied Science (APSC)	
	Chemical & Petroleum Engineering
	Civil Engineering
	Computer Science
	Electrical & Computer Engineering
	General Engineering
	Kinesiology
	Mathematics & Computer Science
	Mechanical Engineering
	Medicine
	Nursing
Applied Social Science (APSS)	
	Child & Youth Care
	Education
	Law
	Policy & Environment
	Public Administration
	Social Work
Business (BUS)	
	Accounting
	Finance
	General Business
	Marketing
	Strategy and Management
Humanities and Arts (HA)	
	Art
	Art History
	Humanities
	Communication
	Creative Writing
	Drama
	English
	French
	History
	Linguistics
	Music
	Philosophy
	Religious Studies

Pure Science (PS)

Biochemistry
Biology
Cellular Biology
Chemistry
Geology
Mathematics
Molecular Biology
Physics
Statistics
Zoology

Social Science (SS)

Anthropology
Archaeology
Criminology
Economics
Geography
Political Science
Psychology
Sociology

Table 5

Anomalies by Discipline

Core Disciplines	**Low**	**High**
Biology	2	0
Chemistry	2	0
Mathematics	0	2
Physics	0	2
Pure Science Total	**4/24=0.17**	**4/24=0.17**
Economics	2	0
Geography	0	0
Political Science	2	0
Psychology	0	1
Sociology	0	0
Social Science Total	**4/30=0.13**	**1/30=0.03**
English	3	0
French	4	0
History	6	0
Linguistics	1	2
Philosophy	1	1
Humanities & Arts Total	**15/30=0.50**	**3/30=0.10**
General Engineering	1	3
Education	1	4
General Business	4	0
Core Disciplines Total	**29/102=0.28**	**15/102=0.15**
Other APSS	2	2
Other APSC	1	12
Other BUS	1	0
Other PS	1	0
Other SS	1	0
Other HA	2	15
Non-Core Disciplines	**8/84=0.09**	**29/84=0.34**

Table 6							
Grade Comparisons Within and Across Faculties							
Within Faculties			**Across Faculties**				
% As for courses within faculty	% As for students registered in faculty	Faculty	APSC	A&SS	BUS	EDUC	SCI
29.7 (2)	30.0 (2)	APSC	33.4 (2)	22.8 (4)	19.8 (3)	33.3 (5)	23.4 (3)
23.6 (3)	22.5 (5)	A&SS	17.2 (4)	23.5 (3)	14.8 (4)	37.2 (3)	14.7 (5)
20.9 (5)	26.5 (3)	BUS	35.2 (1)	29.1 (1)	23.8 (1)	45.0 (2)	32.0 (1)
48.8 (1)	58.3 (1)	EDUC	16.7 (5)	28.6 (2)		65.8 (1)	18.2 (4)
22.7 (4)	24.0 (4)	SCI	25.7 (3)	20.8 (5)	20.6 (2)	33.7 (4)	24.8 (2)

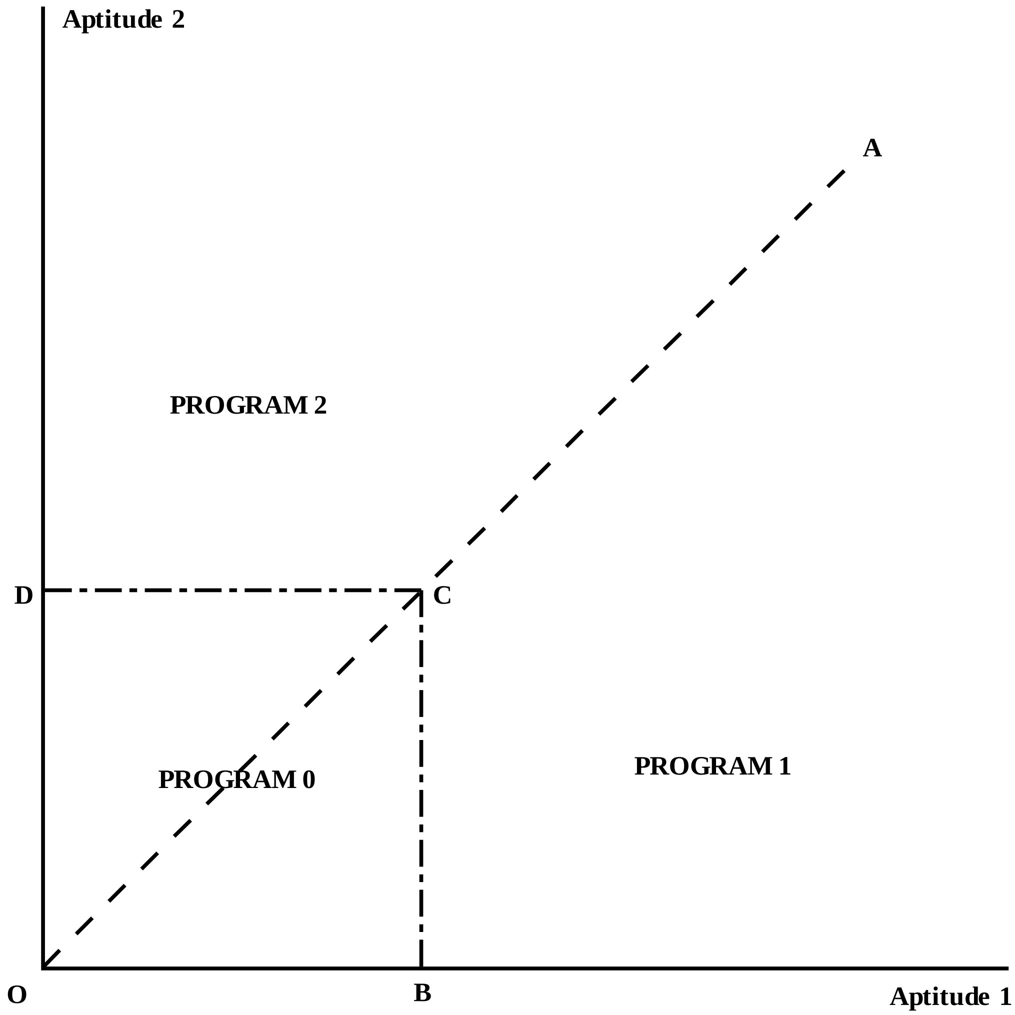

Figure 1: The optimal assignment of students to Programs

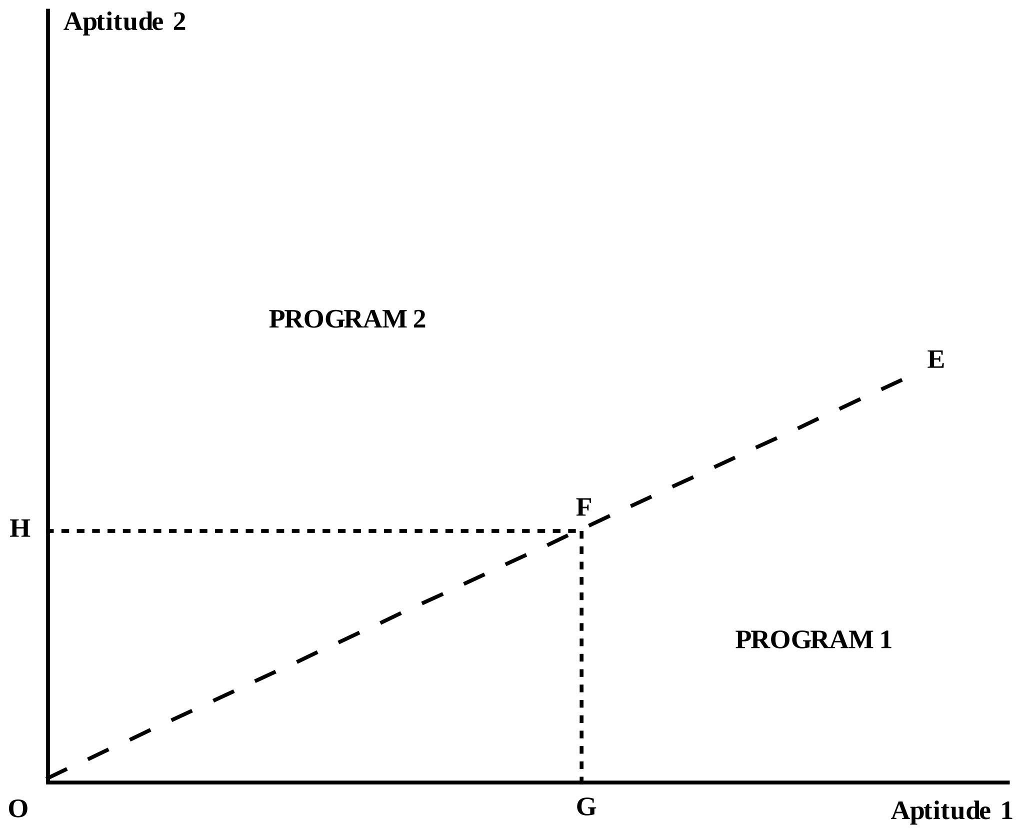

Figure 2: The pooling of students with a given grade

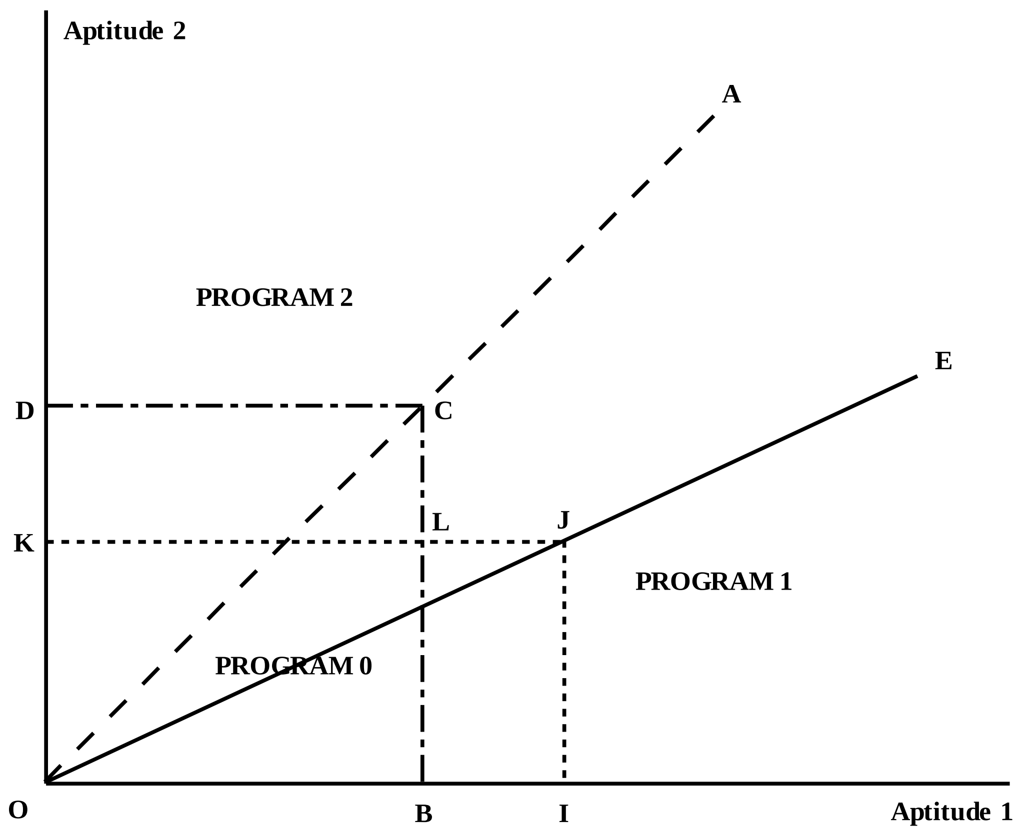

Figure 3: Comparison of the optimal and equilibrium allocations of students to Programs