Technical Report: Instructor and Department Effects on UNC Grades
EPC Grading Subcomittee
March 16, 2007

One of the reasons for implementing the Achievement Index at UNC is the possibility that the grades students receive in different departments may vary due to differing grading practices in different departments and by different instructors. If that is the case, then a student's GPA is affected by two important factors other than her performance in class: the mix of instructors who teach her classes and the mix of departments in which she takes classes.

In order to evaluate the effects of instructor and departmental effects on student grades and GPA, we carried out two regression analyses. Both are based on ten years' worth of data of the population of students who took at least 10 classes at UNC.

Analysis 1: How much of students' grades is due to instructors and departments?

To evaluate this question, we took a 10% sample because of computational limitations. We used a linear mixed-effects model to estimate the relative contribution of student, instructor, and department effects on grades issued. The analysis shows that approximately 6.7% of variation in grades is attributable to instructor grading patterns; 9.6% to departmental grading patterns; and 29.7% to student-level, cross-section consistency (i.e., the idea of the "strong student"). The remaining 53% is, appropriately enough, attributable to variation at the level of the grade itself, which means it's variation *within* the student based on the particular course in which s/he receives each grade. Thus, **approximately 15% of the variation in student grades at UNC appears to be attributable to the grading habits of the instructor and the prevailing practices in the department of the class.**

For this analysis, N=167,654 grades for 5,471 students issued by 7,198 different instructors across 97 departments.

Analysis 2: Does AI do a better job of factoring out instructor and department effects than GPA does?

We test this question by modeling each student's GPA and AI as functions of the mix of instructors and departments from which the student took classes. We began by calculating the mean grade issued by each instructor and in each department in the full data. We then took each student's mix of instructors and departments, assigning each student a departmental-mix mean and an instructor-mix mean. We then model students' GPA and AI as functions of departmental and instructor means.

Unsurprisingly, given an N of 54,711 students, both are highly significant predictors of both GPA and AI. What's more interesting is the variation explained (R-squared) in the two models. For GPA, R-Squared: 0.09028, while for AI, R-Squared: 0.03101. Thus about 9% of variation in GPA is explained by the effects of instructor and department mix, while only about 3% of the variation in AI is explained thus. In other words, **using AI cuts the artificial effects of instructor and departmental grading practices on student rankings by 66%.**

Given that outcome, why go to the trouble of calculating the AI, instead of simply using a model like the one above to correct for instructor and department effects? There are two reasons why the AI remains a better solution: one technical, the other conceptual.

The technical reason is that correcting GPA using a linear model as the ones above could penalize students for the instructors and departments from which they took classes. The AI, by contrast, does not

penalize students who make good grades, regardless of the grading environments in which they study.

The conceptual reason is that the AI corrects for a wide range of extraneous effects on student grades, not just those attributable to instructor and department effects. We therefore do not need to decide ahead of time that we are concerned only with departmental or instructor-based inequality; we can correct for extraneous inequality in general.