Class Size and Teacher Effects on Student Achievement and Dropout Rates in University-Level Calculus

Tyler J. Jarvis

Brigham Young University

Abstract

This paper describes two studies of class size effects on achievement and dropout rates in (traditional) calculus at the university level. The first involved one teacher in both a large and a small section of calculus. The second included four teachers, each in both large and small sections of calculus. After accounting for other significant factors, we found that class size itself was not a statistically significant factor. However, individual teachers did vary in their effectiveness in different class sizes. More importantly, the most effective teachers in large sections were more effective than almost all of the remaining teachers in small sections. Some implications for university administration are discussed.

Class Size and Teacher Effects on Student Achievement and Dropout Rates in University-Level Calculus

Introduction

It is widely believed by faculty, administrators, and students that university students are better served in small classes than in large ones. Indeed, most college ranking and evaluation systems reward schools with smaller classes by ranking them higher than their counterparts offering large classes (e.g., *U.S. News and World Report* Ranking of Colleges and Universities). Yet, schools are generally under substantial financial pressure to increase class size. And mathematics service courses (e.g. college algebra and calculus) are commonly quite large due to high student demand. It is, therefore, important to have reliable information about just how much, if any, students are helped by smaller classes in their study of elementary mathematics at the university.

Class size and its effect on students have been researched repeatedly in recent years, but most of these studies have focused on elementary and secondary schools rather than on university-level teaching. And very few studies have dealt with mathematics. Yet, Smith and Glass's meta-analysis (1980) gives evidence that class size effects vary with students' age. And others studies indicate that class size effects vary with subject matter—even within a discipline (McConnell & Sosin, 1984; Raimondo et al., 1990). This indicates a need for a specific study on the effects of class size in university-level mathematics courses. This paper begins to treat this need by studying class size effects on achievement and dropout rates in calculus classes at two large universities.

In a meta-analysis of class-size studies, Glass and Smith (1979; Smith & Glass, 1980) argue that for elementary school children the benefit of small classes is a logarithmic function of size, with the marginal benefit of reducing class size being most significant for classes of size 20 and fewer. Moreover, the marginal benefit is very small when classes are larger than 25 or 30 students; that is, there is little, if any, benefit to reducing class size if the small class has more than 25 or 30 students. Since most universities cannot afford to reduce class size in introductory mathematics courses to much below 30 students, Glass and Smith's results, if applicable to university-level instruction, would suggest that little or no benefit would be derived from reducing class sizes from relatively large to "small" classes of approximately 30.

Studies of university-level economics and accounting instruction have repeatedly shown little or no significant effect on student achievement from reduced class size (Bellante, 1972; Hill, 1998; Kennedy & Siegfried, 1997). But again, since class size effects vary with subject and discipline, it is important to study the effect of class size on student achievement in introductory-level mathematics.

One concern with many available class-size studies is the fact that, although one would expect the effect of class size to vary substantially with the teacher, few studies account for the teacher effect. Some studies include just one teacher (e.g., Thompson, 1991). These have the problem that they only show the significance of a size effect for the one teacher involved. But for many reasons, one teacher might be much less effective in a large class than another is. Other studies (e.g., Williams et al., 1985) include many teachers, some in large and some in small classes, but without accounting for teacher effects in the study. However, one would expect that the effect on student achievement due to variation among teachers is much larger than any effect due to class size; thus, any model that does not correct for the effect of different teachers will be unable to accurately identify a class-size effect. Indeed, the studies that fail to account for teacher effects have a very poor fit between their models and their data.

Of course, student achievement is not the only measure of teaching success. Students' attitudes may also be important, even if they are achieving more in one type of section (large or small). Studies of students' attitudes in large and small courses give conflicting results. For example, Wood et al. (1974) conclude that student ratings of instructors declined as enrollment increased to 240, but beyond that point, ratings began to improve. But Marsh et al. (1979) found little correlation between class size and students' attitudes about the course. And Sweeney et al. (1983) found that large economics courses were actually preferred over small ones.

Because student evaluations and attitude surveys are relatively unreliable measures and appear to vary widely from class to class, and even from day to day within a class, student dropout rates may be a more objective measure of student attitude. A secondary analysis of the effect of class size on dropout rates is also included here, but clearly more work remains to be done on the effects of class size on student attitudes.

Preliminary Study: University A

Models

The object of the preliminary study, which was performed at a large land-grant institution in the southern United States (*University A*), was to study the effect of class size (S) on the performance of students in a traditional (as opposed to "reformed") Calculus I course, while adjusting for other important effects such as students' general mathematical ability as measured by their ACT score (A), students' sex (G), students' ethnic group (E), teacher (T), and semester (M). Performance was measured principally by a standardized final exam, but we were also interested to see if students' success (as measured by their grade) in Calculus II might be differently affected by class size. The following statistical models were used for analysis of the data.

Model to measure size effects on final exam scores

Let $y_{l(i,j,k)}$ be the score on a standardized final exam of the student l in the class k taught by the teacher j in a class of type i (large if i=1, small if i=0).

Several factors were included that might affect $y_{l(i,j,k)}$. These are
Size of the class: S_1 := effect of large class on student performance. S_0 := effect of small class on student performance.
Teacher. T_j := effect of teacher j.
Sex. $G_{l(i,j,k)}$:= effect of sex.
Ethnic Group. $E_{l(i,j,k)}$:= effect of ethnic group.
Semester. M_K := effect of semester in which the course was taught.
There are also potential effects due to interaction between different effects, for example:
$SG_{il(j,k)}$:= interaction effect between class size and sex.
$SE_{il(j,k)}$:= interaction effect between class size and ethnic group.
$GT_{lj(i,k)}$:= interaction effect between sex and teacher.
$ET_{lj(i,k)}$:= interaction effect between ethnic group and teacher.
Also, there is an adjustment for initial aptitude and preparation, as measured by the ACT score A_l and scaled by a linear factor β.

Letting μ be the overall mean, the basic model is

$$\begin{aligned} y_{l(i,j,k)} = \mu + S_i + T_j + G_{l(i,j,k)} + E_{l(i,j,k)} + M_k + ST_{ij} + SG_{il(j,k)} \\ + SE_{il(j,k)} + \ldots + SGET_{ijkl} + \beta\ A_{l(i,j,k)} + \delta_{k(i,j)} + \varepsilon_{l(i,j,k)} \end{aligned} \quad (1)$$

Here $\delta_{k(i,j)}$ is the random error term associated to each class, and $\varepsilon_{l(i,j,k)}$ the random error associated to each student.

The initial factors of sex and ethnic group were decided on because other studies have found them to be significant (see, for example, McConnell and Sosin, 1984). It also seemed very likely that the teacher would have a significant effect on achievement, and so teacher was also included as an initial factor.

This statistical model is similar to a standard split-plot model, but differs in that we have an adjustment for the covariate (ACT score). The use of the covariate is important to compensate for the fact that various factors beyond our control may affect the types of students enrolled in the different-sized classes. For example, the small classes fill up more quickly, leaving those who register late in the large classes (this may also partially account for the common perception that large classes are worse). The covariate helps to account for these differences.

Model to Measure Calculus I Size Effects on Performance in Calculus II

The preceding model describes the effects of class size on student mastery of the material in Calculus I as measured by the departmental final exam. But it also seemed possible that class size in Calculus I might have a different effect on students' performance in the follow-up course, Calculus II, than it did on performance on the Calculus I final exam. The model we used to test this hypothesis was similar to the preceding model, but it used the restricted data set of only those students that continued on to Calculus II. Success in Calculus II was measured using students' final grade in Calculus II, after adjusting for the additional effect of different teachers for Calculus II and adjusting for the higher order interactions that might be associated with this teacher effect.

Although it would be interesting and important to know the effects of large class size on subsequent enrollment in mathematics, it is probably impossible to draw conclusions about such effects from available data. The difficulty arises because student enrollment in subsequent mathematics courses, especially in Calculus II, is primarily dependent upon program requirements rather than student preference.

Preliminary Study Data

In this study, one instructor taught both a large (ca. 90 students) and a small (ca. 35 students) section of introductory calculus in the course of one academic year (1995-6). Eight other instructors taught small sections, and their data are used to standardize the final exam, in particular to account for differences between semesters. Their data also helped identify some changes that needed to be made for the main study at University B. This set contains data for 293 students.

Conclusions of the Preliminary Study

It was impossible to measure the interaction of teacher and class size because only one teacher taught both large and small sections. However, for that one teacher, class size was not statistically significant. The interaction effects of class size with ethnic group and with sex were also not significant. See Tables 1 and 2 for more details.

By far the most significant influences on student performance were initial preparation and aptitude, as measured by the ACT math score, and teacher (although only one teacher taught both large and small sections, eight others taught small sections). The fact that the teacher would have an effect is not surprising. But the magnitude of this effect was large compared to all others (except the ACT scores). This seems to indicate that class size effects cannot be effectively measured without carefully adjusting for teacher. Moreover, the widely varying nature of results of other studies that did not adjust for teacher effect (or which only include one teacher) indicates the potential for a large interaction between size and teacher.

Another large effect was associated with the semester in which the course was taught. This is probably due to the fact that students who are well-prepared for Calculus I by their high school program are likely to take Calculus I in the Fall semester of their first year, whereas the remaining students are more likely to take Calculus I after first taking a semester of prerequisites.

The time of day the courses were offered varied through the regular school day (8 am to 4 pm), but despite some expectations to the contrary, time was not significant. It appears that the covariate accounted for essentially all variation associated to differences in time of day.

The model measuring the performance of students in the subsequent class, Calculus II, showed no additional information given by using students' grade in Calculus II to measure performance in Calculus I. In fact, after systematic removal of insignificant factors, the best model for student performance in Calculus II appeared to be one that depends only on teacher of the Calculus II section and the students' score on the Calculus I final. Because tracking students to Calculus II gave no information about student mastery of Calculus I beyond that given by the final, that aspect was dropped in the main study.

Finally, although the R^2 value of .49 for this model was stronger than many that have been published on class size (ranging from R^2=.39 in Glass and Smith, 1979, down to R^2 =.01 in Williams et al., 1985) it still seemed relatively weak, indicating a need for a better covariate and for inclusion of other significant factors.

Implications of the Preliminary Study for the Main Study

The results of the preliminary study indicated several things that were important to change for the main study. First, since teacher effect was so large relative to other effects, and since the teacher-size interaction term was suspected to be large as well, the main study needed several different teachers in both large and small sections. Second, since the semester effect was significant, whereas the time of day was not, the teachers should, when possible, teach both small and large sections in the same semester. Third, additional factors needed to be considered and a better covariate used in order to improve the model.

Main Study: University B

The goal of the main study was to decide if size had a significant effect on student achievement in calculus, and to see if the teacher-size interaction was significant. This study was conducted at a large, private university in the western United States (*University B*).

Main Study Model

The model used for the main study was similar to the preliminary study, with some additional factors considered. These included a pretest as an additional covariate, instead of using just the ACT. We also included age, course load, major college, and total hours earned as factors that might affect student achievement. The pretest was included as an additional covariate because a significant number of students had actually studied some calculus before enrolling in the sections included in this study. Although students who had previously taken calculus for university credit could be identified and excluded from the study, those who had taken calculus in high school could not. The ACT math score seemed to be a good measure of general mathematical aptitude, but our pretest was designed to test preparation specifically relevant to calculus.

We originally expected the ACT and the pretest to be linearly dependent (at least after accounting for other factors such as students' age). However, a test of this hypothesis showed no significant correlation between the ACT math score and the pretest. This is probably because, as explained above, they actually test different things. Consequently, we included both in the model.

Main Study Data

The primary data consist of pretest and final exam scores for 1,984 students in first-semester calculus and 134 students in second-semester calculus, collected over two years at University B. The data also include the various other potential factors described in the previous section that might influence student achievement.

The final exam was written by a departmental committee to represent the core topics and skills that were considered most important for students to know. This was considered a good measure of learning since it represented the consensus of a large number of mathematics instructors about what constitutes successful (traditional) calculus learning.

For the purposes of this study, small classes are classes with 20-35 students, while large classes contain 150-240 students. Both kinds of classes included review sessions (20-35 students) twice a week in addition to the main lectures.

Students who had taken the course previously were not included. Students who dropped the class were also not included, since they did not take the final exam. There was some concern that weak students might be more likely to drop from a large section, but a separate logistic regression showed that for a given teacher, and after adjusting for pretest scores, students drop essentially randomly.

Calculus I Data

The Calculus I data cover four semesters (Fall and Winter of 1997 and 1998), and 27 teachers. One teacher (*Teacher Q*) taught both small and large sections in Winter 1997, a different teacher (*Teacher L*) taught both small and large sections in Fall 1997, a third teacher (*Teacher T*) taught both small and large sections in Winter 1998, and a fourth teacher (*Teacher AA*) taught both small and large in Fall 1998. Some of these four taught large and small sections in other semesters, but not simultaneously. One other teacher (*Teacher M*) taught only large sections, and the remaining teachers taught only small sections. These other teachers are included for purposes of standardizing the pretest and final, and for estimating the relative magnitude of the teacher-size effect.

This set contains data for 1,984 students. They are divided into six ethnic groups and 11 major colleges (and also the option of an undeclared major).

Calculus II Data

One teacher taught both small and large sections of Calculus II in Fall 1997. The total number of students in the data set (after removing students who dropped or who had taken the course before) is 134. The same demographic data were included here as those included in the Calculus I data set.

Main Study Analysis

Calculus I Analysis

Initially the model considered the potential effects of many factors, as described above. Class size alone was not significant, nor were most of its higher-order interactions (see Table 3). After a standard, systematic elimination of insignificant variables, the model had as its main factors ACT, pretest, teacher, semester, major college, and the teacher-size interaction. Unlike in the case of University A, at University B the interactions between teacher and ethnicity and between teacher and sex were not significant (see Table 4).

Major college is probably significant because those who have aptitudes in mathematics are most likely to major in mathematically challenging fields like engineering. But since interest and aptitude are not easily changed by the university, they are of relatively little interest to teachers and administrators.

The factor that is the most interesting is the (weakly significant) teacher-size interaction term. The fact that size itself was not significant and that this interaction term was weakly significant in the final model shows that the size effect, if there is any size effect at all, depends primarily upon characteristics of each individual teacher.

This final model had an R^2 value of .61—a substantial improvement over the preliminary (University A) study.

Calculus II Analysis

Again, a variety of different potential factors were included and then insignificant terms were systematically eliminated (see Table 5). And again, size was not significant. Teacher-size interaction was not measurable here, since only one teacher is involved.

Unlike first-semester calculus, student's age seemed to be a significant factor, as well as a student's current course load. The appearance of age as a factor may be because many students delay taking their second semester of calculus, sometimes for several years, whereas most students take their first semester of calculus in their first year. On the other hand, many students with high school background in calculus will take second-semester calculus immediately in their first year. This gap between some but not all students' first and second semesters of calculus seems the most likely explanation for the role that students' age plays. The interaction between ACT math score and age may be significant because students who take second-semester calculus immediately in their freshman year have also just recently taken the ACT, whereas others who delay taking second-semester calculus may have taken the ACT several years earlier, thus its predictive value will likely vary somewhat with the student's age.

The effect of student course load is harder to explain, but a possible explanation would be that students in their first semester at the university will often simply take the recommended general education courses (including Calculus I) and the recommended total hours, whereas more experienced students will vary from the norm, perhaps because they think they know better what they want and how to accomplish it. Consequently, course load in the second-semester calculus courses may reflect a student's personal choices and attitudes toward school work, rather than reflecting the advice of a the university or a counselor.

Major college plays no role in Calculus II, but it is significant in Calculus I. We conjecture that this is because many majors require only first-semester calculus, which also fulfills some university general education requirements, whereas most majors that require second-semester calculus are either in the College of Engineering or the College of Physical and Mathematical Sciences, which have relatively similar coefficients in the first-semester model (see Table 6). Moreover, few students take the course as an elective.

Conclusions of the Main Study

Class size was not significant, and even the teacher-size interaction effect was only weakly significant (i.e., significant at p=.05, but not at p=.01). No other interaction terms involving size were significant. This suggests that if there is any effect on students' achievement due to class size, it is a function of the individual teacher and her or his ability and attitude, rather than a function of the size alone.

The important question to ask about class size is whether it is in the students' and the university's best interest to increase or decrease class sizes. The insignificance of size as factor in achievement is, taken alone, not enough to answer that question. In particular, we must ask whether some teachers in large classes are more effective than others in small classes. Also, it is important to know if more students drop out of large classes, since their data would not be included in the study without final exam scores (failing students who did not drop were part of the main study). These two questions are the subject of the additional analyses described in the next two sections.

Additional analysis—Net Effect of Teacher and Teacher-Size Effects

In order to decide whether a good teacher in a large section was more effective than other teachers in small sections, we solved for the (biased) coefficients in the previous Calculus I model.

The results are listed in Table 6. We found that the best teachers in large sections (three of four who taught large sections) were better for student achievement than all but four of the remaining 24 teachers who taught in small sections.

In particular, teacher Q had an effect of 10.3 and a teacher-size interaction effect of -5.2 for large classes, making a total effect of 5.1 to a student's final exam score in teacher Q's large section. Teacher M only taught large sections, and had an effect of 3.7, and teacher AA had a total effect of 2.62 in large classes. However, only four small-section teachers (*B*=14.1, *H*=8.8, *L*=5.4, and *O*=7.3) had a better effect in their small sections than these three teachers (M, Q, and AA) of large sections. The remaining 20 teachers taught only small sections, and they had an effect that ranged from -5.8 up to 2.57. The vast majority of the students were enrolled in small sections (about 1,300) but only about 120 students were enrolled in the courses of the most effective small teachers (B, H, L, and O). The remaining 1,180 small-section students would have been better off in the large sections of M, Q, and AA.

Also, note that the teacher-size coefficient for teacher AA is positive—indicating that teacher AA was actually *more effective* in the large class than the small one. The remaining three teachers of both large and small sections (Q, L, and T) appear to have been slightly more effective in small rather than large sections. In all cases, the teacher-size interaction, if it should be included at all (being only weakly significant), is dwarfed by the effect due to teacher alone: the variation due to class size (i.e., the variation among the teacher-size interaction terms) was only 6.08 compared to a variation of 19.97 due to teachers. This helps explain the widely differing conclusions and the poor fit between model and data in existing class size studies that do not account for variation due to teacher—any size effects are completely masked by teacher effects.

Additional Analysis—Dropping out

Some critics of large sections have claimed that students dislike large sections. This seems difficult to test carefully, so we ask instead if more students drop out of large sections than small sections.

Drop Model and Analysis

Using a standard logistic regression, we analyze the influence of class size on dropping in both first- and second-semester calculus. For each of the teachers in the University B Calculus I and Calculus II data sets who taught both large and small sections simultaneously, we let $D_{i(j)}$ denote the odds ratio (that is, p/(1-p)) that student *i* in class *j* of type *k* (large or small) will drop the class. Let P_i denote student *i*'s pretest score, which will be scaled by a linear factor α, and let S_k denote the effect due to being in a class of type *k* on the odds ratio of dropping.

For each teacher we compare the two models

$$\log(D_{i(j,k)}) = S_k + \alpha P_i + \delta_{jk} + \varepsilon_{i(j,k)}, \qquad 2)$$

and

$$\log(D_{i(j,k)}) = \alpha P_i + \delta_{jk} + \varepsilon_{i(j,k)}, \qquad 3)$$

where δ_{jk} is the random error term associated to each class, and $\varepsilon_{i(j,k)}$ is the random error associated to each student.

For two of the four Calculus I teachers (teachers Q and AA) the total number of students who dropped was so small, (3 of 183 and 8 of 238, respectively), that no conclusions about dropout rates could reasonably be drawn from their classes. For both of the remaining two teachers of Calculus I and the teacher of Calculus II, the Wald χ^2 indicated that size was not significant in the first model (2), and the value of *c* (a goodness-of-fit test comparable to R^2) did not change significantly when size was deleted (model (3)). These results are summarized in Table 7.

As in the case of achievement, the influence of class size on students' dropping appears to be small or nonexistent. If it is a factor, it probably varies with the teacher, but it appears to be insignificant for the three teachers involved in this study. However, the goodness-of-fit statistic *c* in these models is fairly close to .5, which seems to indicate a relatively poor fit, so more studies and a better model are needed to get a clear answer about the effect of class size on students' dropping out.

Caveats

The results of all of these studies apply only to the difference between classes of about 30 and classes of about 90 to 180. It is very possible (and even likely, based on evidence from primary and secondary schools—see Glass and Smith, 1979) that a significant difference in achievement will exist between very small classes (ten or fewer students) and those which are small for a university mathematics class (20-30 students).

It is also important to remember that, although both the preliminary and the main studies found no significant effect due to size of the lecture section, all classes in the main study were supplemented by small review sessions (held twice weekly). These review sections were all small—about 30 students each. In the preliminary study, neither the large nor small sections had associated review sections. While the preliminary study showed no significant effect due to class size even without these small review sections, the review sections did appear to be helpful to students in both large and small sections alike, so without further research, we cannot recommend replacing small sections by large ones that have no smaller review sections. It is possible that the size of the review session does have an impact on student achievement even when the size of the main lecture does not. Indeed, Kennedy and Siegfried (1997) cite work of Attiyeh and Lumsden from 1972 which shows some evidence that this is the case in introductory economics classes. Further research in this direction is warranted.

Summary

Our main results are that class size effects on achievement and dropout rates in university calculus classes appear to be a function of the teacher only. In particular, averaging over all teachers in the study, class size had no significant effect on students' achievement on the final exam in calculus. However, some teachers are substantially less effective in a large class than other teachers are, and some teachers are more effective in a large class than in a small one. More significant is the conclusion of the second analysis, comparing the relative effectiveness of different teachers; namely, three teachers of large sections were more effective than all but four of 24 teachers of small sections. Consequently, value added (in terms of student achievement on the final exam) was generally greater in the large sections with these more-effective teachers than it was in the small ones with the less-effective teachers.

Final Conclusions

More careful studies on the effects of class size versus teacher effect are warranted. The work described in this paper is the only study to date on the effects of class size in mathematics classes at the university level, despite the fact that research shows class size effects vary depending on discipline and students' age. This study is also one of the few carefully designed studies on effects of class size at the university level in any subject. Despite the enormous, well-demonstrated effect of teacher on achievement, few studies even attempt to correct for teacher effects, let alone attempt to design an experiment explicitly controlling for teacher, as done here. Similarly, despite the many reasons that students of a given background and ability might tend to enroll in a larger or a smaller section, thus skewing results, very few studies of class size attempt to control for or account for this phenomenon, as we have done here with the pretest and ACT scores. This lack is especially pronounced in studies at the university level, and even more so among those studies that also control for teacher effects.

This study shows that teacher effects probably dwarf all other effects, but a study with more teachers and more sections would be very valuable. Unfortunately, logistic and political barriers make it difficult to conduct a larger study. Since both large and small sections' being offered in the same semester requires large numbers of students, and because the number of faculty both available and willing to teach both small and large sections of calculus in the same semester is relatively limited, even at a large university like University B, no larger study than this one is likely to be conducted at a single university. Any study involving more teachers will probably require cooperation from faculty and administrators at several universities to ensure, among other things, a common syllabus, pretest and final exam, and appropriate scheduling of instructors in both large and small sections. Thus, university faculty and administrators must understand the importance of settling this question of class size effects and cooperate in such studies.

If, as this and other smaller studies seem to indicate, the individual teacher has a major effect, and class size is an insignificant factor, both in terms of student achievement and in terms of students' likelihood to complete the course, then even if a university can afford to offer calculus in small sections with first-rate teachers, it still may not be the university's best strategy. One of those teachers may well be more effective in a large class than most or all of the others in a small class; therefore, students might be better served in the large section than in most or all of the small ones. At most universities, however, small classes are achieved only by hiring adjuncts, graduate students, and less-qualified faculty. In such a case, reducing class size may be doing students a disservice while simultaneously increasing instruction costs.

References

Bellante, D. M. (1972). A summary report on student performance in mass lecture classes of economics. *Journal of Economic Education*, 4(1):53-54.

Glass, G. V. and Smith, M. L. (1979). Meta-analysis of research on class size and achievement. *Educational Evaluation and Policy Analysis*, 1:2-16.

Hill, M. C. (1998). Class size and student performance in introductory accounting courses: Further evidence. *Issues in Accounting Education*, 13(1):47-64.

Kennedy, P. E. and Siegfried, J. J. (1997). Class size and achievement in introductory economics: evidence from the TUCE III data. *Economics of Education Review*, 16(4):385-94.

Marsh, H. W., Overall, J. U., and Kesler, S. (1979). Class size, students' evaluations, and instructional effectiveness. *American Educational Research Journal*, 16:57-70.

McConnell, C. R. and Sosin, K. (1984). Some determinants of student attitudes toward large classes. *Journal of Economic Education*, 15(3):181-90.

Raimondo, H. J., Esposito, L., and Gershenberg, I. (1990). Introductory class size and student performance in intermediate theory courses. *Journal of Economic Education*, 21(4):369-381.

Smith, M. L. and Glass, G. V. (1980). Meta-analysis of research on class size and its relationship to attitudes and instruction. *American Educational Research Journal*, 17(4):419-33.

Sweeney, M. J. B., Siegfried, J. J., Raymond, J. E., and Wilkinson, J. T. (1983). The structure of the introductory economics course in united states colleges. *Journal of Economic Education*, 14(4):68-75.

Thompson, P. A. (1991). The effect of section size on student performance in a statistics course. *College student journal*, 25:388-395.

Williams, D. D., Cook, P. F., Quinn, B., and Jensen, R. P. (1985). University class size: Is smaller better. *Research in Higher Education*, 23(3):307-18.

Wood, K., Linsky, A. S., and Straus, M. A. (1974). Class size and student evaluations of faculty. *Journal of Higher Education*, 45:524-34.

Author Note

This research supported in part by NSF grant DMS-9501617. I am grateful to Govinda Weerakody and David Whiting for help with the statistical models, and to Ralph Brown, Bruce Collings, Tamara Cooper, Missie Elkins, Pedro Geoffrey, Donald Jarvis, Matt Johnson, and Krishnaswamy Venkata for helpful discussions.

Correspondence concerning this article should be addressed to Tyler J. Jarvis, Department of Mathematics, Brigham Young University, Provo, UT 84602. Electronic mail may be sent to jarvis@math.byu.edu

Table 1. University A: Initial model. The initial model included many potential factors, but most were not significant. Neither class size nor any of its higher-order interactions are significant.

Model	
R^2	Pr > F
0.501564	0.0001
Variables	
Source	Pr > F
A = ACTMATH	0.0001
T = TEACHER	0.0002
M = SEMESTER	0.0287
E = ETHN	0.3037
G = SEX	0.6897
S = SIZE	0.5010
ET = ETHN*TEACHER	0.0908
EM = ETHN*SEMESTER	0.3148
AE = ACTMATH*ETHN	0.9240
GT = SEX *TEACHER	0.0289
GM = SEX *SEMESTER	0.3812
AG = ACTMATH*SEX	0.9755
SE = ETHN*SIZE	0.2297
SG = SEX *SIZE	0.6945

Table 2: University A: Final Model This model is the result of systematic elimination of the insignificant variables. Note that ethnic group and sex are not significant themselves, but only in their interaction terms with teacher. That coincides with the intuition that sex and ethnicity themselves do not play a role in achievement, but the way the teacher responds to them does.

Model	
R^2	Pr > F
0.490446	0.0001
Variables	
Source	Pr > F
TEACHER	0.0001
ACTMATH	0.0001
SEMESTER	0.0276
ETHN	0.2957
SEX	0.6877
ETHN*TEACHER	0.0796
SEX *TEACHER	0.0306

Table 3: University B: Calculus I. Initial model. The initial model included a large number of potential factors, including total hours of credit the student had earned (THOURS), student's age (AGE), ethnic group (ETHN), sex (SEX), current course load (LOAD), and major college (MAJCOLL). Most of these were not significant and were systematically removed.

Model	
R^2	Pr > F
0.641257	0.0001

Variables	
Source	Pr > F
SEMESTER	0.0001
ACTMATH	0.0001
PRETEST*SEMESTER	0.0001
TEACHER	0.0001
LOAD	0.0572
AGE	0.0161
AGE*AGE	0.0074
MAJCOLL	0.0292
SEX	0.9126
ETHN	0.5431
SIZE	0.1204
THOURS	0.7504
THOURS*THOURS	0.3043
TEACHER*SEX	0.9633
ACTMATH*AGE	0.0589
LOAD*MAJCOLL	0.4501
LOAD*LOAD	0.2868
SIZE*SEX	0.4114
SIZE*ETHN	0.4553
TEACHER*SIZE	0.1491
TEACHER*ETHN	0.4676
SEMESTER*TEACHER	0.0091
AGE*THOURS	0.4642

Table 4: University B: Calculus I. The final model, developed by systematic elimination of insignificant variables (including class size), shows that class size itself is not a significant factor, but the teacher-size interaction term might be significant.

Model	
R^2	Pr > F
0.612772	0.0001
Variables	
Source	Pr > F
SEMESTER	0.0001
ACTMATH	0.0001
TEACHER	0.0001
PRETEST	0.0001
MAJCOLL	0.0265
TEACHER*SIZE	0.0475

Table 5: University B: Calculus II. This model, developed by systematic elimination of insignificant variables, shows that class size is not significant for this particular teacher.

Model	
R^2	Pr > F
0.452498	0.0001
Variables	
Source	Pr > F
ACTMATH	0.0131
PRETEST	0.0001
SIZE	0.6948
LOAD	0.0014
AGE	0.0238
ACTMATH*AGE	0.0255

Table 6. University B: Calculus I. Estimated coefficients (biased) for the University B Calculus I model.

Parameter	Estimate	Parameter	Estimate
INTERCEPT	10.34158770	TEACHER	
ACTMATH	1.81378293	Teacher A	2.97328398
PRETEST	0.28145443	Teacher B	14.14475384
		Teacher C	-2.84780267
SEMESTER		Teacher D	-5.83704943
1997 Winter	-22.81697752	Teacher E	-1.72930695
1997 Fall	-30.78852062	Teacher F	1.42747917
1998 Winter	-31.56136904	Teacher G	2.36437273
1998 Fall	0.00000000	Teacher H	8.75866094
		Teacher I	1.99462535
MAJOR COLLEGE		Teacher J	-1.34914128
Art	1.06288729	Teacher K	1.36048855
Biology	3.09426616	(large and small) Teacher L	5.36603618
Business	1.70441351	(large only) Teacher M	3.67414925
Education	-1.64708684	Teacher N	-3.54302678
Engineering	0.95833944	Teacher O	7.33831105
Family science	-8.03393284	Teacher P	-4.96070232
Health/PE	0.72086673	(large and small) Teacher Q	10.34994654
Humanities	2.41684870	Teacher R	0.41012342
Nursing	1.55783270	Teacher S	2.57185348
Phys. & math. science	2.44752224	(large and small) Teacher T	1.01064676
Social science	-0.36290885	Teacher U	-0.05621085
Undeclared	0.00000000	Teacher V	-0.53756900
		Teacher W	-0.65055273
TEACHER*SIZE		Teacher X	-4.58577073
Teacher L large	-6.08437615	Teacher Y	-0.33776313
Teacher L small	0.00000000	Teacher Z	1.50086908
Teacher Q large	-5.25948571	(large and small) Teacher AA	0.00000000
Teacher Q small	0.00000000		
Teacher T large	-3.17380207		
Teacher T small	0.00000000		
Teacher AA large	2.61853700		
Teacher AA small	0.00000000		

Table 7. Dropout rates: Logistic regression to evaluate the effect of size on a student's likelihood of dropping the class shows that there is no significant effect due to class size for these teachers.

Calculus I Teacher L	Winter 1998	Calculus I Teacher T	Fall 1997	Calculus II	
218 students	21 dropped	257 students	82 dropped	170 students	36 dropped
With size (Equation (2))		With size (Equation (2))		With size (Equation (2))	
Variable	Pr > Wald χ^2	Variable	Pr > Wald χ^2	Variable	Pr > Wald χ^2
INTERCPT	0.0061	INTERCPT	0.0999	INTERCPT	0.4633
SIZE	0.5187	SIZE	0.2835	SIZE	0.5509
PRETEST	0.0001	PRETEST	0.4654	PRETEST	0.0036
	c= 0.664		c = 0.587		c= 0.654
Without size (Equation (3))		Without size (Equation (3))		Without size (Equation (3))	
Variable	Pr > Wald χ^2	Variable	Pr > Wald χ^2	Variable	Pr > Wald χ^2
INTERCPT	0.0004	INTERCPT	0.0976	INTERCPT	0.2871
PRETEST	0.0001	PRETEST	0.4179	PRETEST	0.0038
	c = 0.658		c = 0.541		c = 0.652