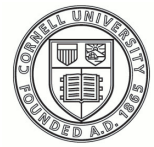

9-2010

The Impact of Class Size and Number of Students on Outcomes in Higher Education

James Monks
University of Richmond

Robert Schmidt
University of Richmond

The Impact of Class Size and Number of Students on Outcomes in Higher Education

Abstract

Numerous studies have investigated the impact of class size on student outcomes. This analysis contributes to this discussion by examining the impact of class size on student outcomes in higher education. Additionally, this paper investigates the importance of student load (total number of students taught across all courses) in educational outcomes. We find that both class size and student load negatively impact student assessments of courses and instructors. Large classes and heavy student loads appear to prompt faculty to alter their courses in ways deleterious to students.

Keywords

higher education, students, outcomes, performance, class size

The Impact of Class Size and Number of Students on Outcomes in Higher Education

James Monks

and

Robert Schmidt

Robins School of Business
University of Richmond
Richmond, VA 23173
jmonks@richmond.edu

September 2010

The Impact of Class Size and Number of Students on Outcomes in Higher Education

Abstract

Numerous studies have investigated the impact of class size on student outcomes. This analysis contributes to this discussion by examining the impact of class size on student outcomes in higher education. Additionally, this paper investigates the importance of student load (total number of students taught across all courses) in educational outcomes. We find that both class size and student load negatively impact student assessments of courses and instructors. Large classes and heavy student loads appear to prompt faculty to alter their courses in ways deleterious to students.

I. Introduction

Numerous studies have investigated the influence of class size on student attitudes, behaviors, and outcomes. The overwhelming majority of these studies have focused on elementary school and even pre-school effects of class size on student achievement. The conventional wisdom among parents, teachers, school administrators, and policy makers is that smaller class sizes translate to improvements in student learning and outcomes. This conventional wisdom, however, has not been universally supported by empirical evidence. While a number of studies have found support for the importance of class size on student achievement, others strongly refute this claim concluding that class size has little to no impact on objective student outcomes. The difficulties in assessing the causal influence of class size on student outcomes, such as achievement, are (1) class size itself is often not directly observed but rather proxied by pupil-teacher ratios at the state, district, or school level, (2) many data sets used to analyze this question are cross-sectional and thus do not allow one to control for fixed student, teacher, class, or school effects, and (3) class size itself may be endogenous in a student outcome equation. Nonetheless, the general consensus among researchers examining this issue is that if class size matters at all its influence is most pronounced at the lowest grade levels.

Only a handful of studies have focused on the role that class size may play in outcomes in tertiary education. Clearly, the educational environment is dramatically different from the classroom and learning environment of the elementary school setting. Even so the conventional wisdom of the benefits of small class size persists in postsecondary education, as well. This intractable perception is so prevalent that class size represents two of the fifteen inputs into the *U.S. News and World Report* college rankings formula, despite the lack of convincing evidence that class size has a significant impact on student outcomes. A university's rank is a function of the percentage of course sections that it offers with fewer than twenty students and the

percentage of course sections that it offers with fifty or more students. The former enters the rankings formula positively, and the latter negatively. While the *U.S. News and World Report*, and other college rankings, are often criticized for their focus on inputs rather than outputs in the educational process, rarely is the use of class size the primary focus of this criticism.

Not only is the educational setting dramatically different in tertiary education in comparison to elementary education, the primary focus of most of this literature, but class size itself may have a fundamentally different relationship to student outcomes in postsecondary education. In elementary school, where a teacher usually teaches the same class of students almost all school day long, class size and the number of students that a teacher is responsible for are equivalent. In higher education, a professor may teach one, two, three or more sections of a course each semester. In this case, class size -- the number of students in a class section -- and the total number of students that a professor is responsible for may or may not be the same; thus the existing studies of class size in higher education combine and confound class size effects and total student responsibility effects.

This study overcomes many of the shortcomings outlined above by taking advantage of a unique policy change within a business school at a private, selective university in the United States. A new dean of business at this university allowed professors who were teaching three sections of a course per semester to "super-size" these sections into two larger sections, of roughly equal total number of students. For example, a typical super-sized course went from three sections of thirty students to two sections of forty-five students. The impetus for this policy was an attempt to reduce the use of adjunct faculty and a low cost means of lowering the teaching loads of faculty. This practice persisted for approximately six years, until a new dean arrived at the school. With an eye on the new *BusinessWeek* undergraduate business school rankings, which is a function of average class size, the dean eliminated the use of super-sizing

and returned those faculty to three sections. This natural experiment allows us to compare student outcomes before, during, and after significant changes in class size as a result of the policy change and subsequent reversal. Additionally, because we are able to track the same professors over time, we are able to control for faculty and course fixed effects and thus estimate within instructor and course class size and number of student effects. As not all faculty were eligible for this policy change, because they taught fewer than three sections of the same course per semester, and because not all faculty opted to super-size even if eligible, we are able to compare the student outcomes of the super-size class sections to a control group of faculty over the same time period in the same courses. Furthermore, because the super-size policy did not significantly alter the total number of students that a faculty member was responsible for teaching, this policy allows us to identify the direct effect of class size on student outcomes separate from faculty student-load effects.

This paper will examine the influence of significant changes in class size on student assessments of university business courses conditional on fixed instructor and course effects benchmarked against instructors and courses at the same university whose class sizes were not substantially altered over the same time period. The following section of the paper will briefly outline the extensive literature on the impact of class size on student achievement. This section is followed by a discussion of the data, and then direct empirical tests of the relationship between class size and faculty student-load on student outcomes. Finally, the conclusion of the paper discusses the implications of these results.

II. Literature Review

As mentioned above there is a vast literature on the role of class size on student achievement.[1] The bulk of this literature focuses on whether class size is instrumental in improving learning and academic achievement at the elementary school level, with the growing use of field experiments to overcome the selection problem inherent in many class size estimates. Perhaps the most famous and widely investigated experiment in class size is the Tennessee STAR program. This program ran from 1986 to 1989 and randomly assigned children entering kindergarten into classes of 15-17 students, 22-25 students, or 22-25 students with a teacher's aide. These students remained in these class sizes through the third grade, and then were returned to regular sized classes for the fourth grade. Numerous studies have examined the results from this experiment and found that the students in the smaller classes performed significantly better on standardized tests in mathematics and reading in kindergarten through the third grade (see Word et al. (1990); Finn et al. (1990); Nye et al. (2000 and 2001)). Additional research (Nye and Hedges (2001); Finn (1998); Mosteller (1995)) finds class size effects that persist at least until the seventh or eighth grade.

Krueger (1999) utilizes multiple regression analysis and the composite mathematics and reading score from the Stanford Achievement Test. Controlling for numerous family background and school characteristics, he finds that students in smaller classes performed approximately .2 to .3 standard deviations better on this standardized test than students in the larger classes over the first four years of schooling. The STAR program design has been criticized for a failure to pre-test the participating students to assure that the assignment was truly

[1] For a more complete discussion of class size on primary and secondary student achievement see "Exploring the Effect of Class Size on Student Achievement : What Have We Learned over the Past Two Decades." By Susan Averett and Michele McLennan in International Handbook on the Economics of Education, eds. Geraint Johnes and Jill Johnes, Edward Elgar Publishing, Cheltenham, UK. 2004.

random across class size. Additionally, some of the participating students were reassigned across class sizes due to behavior problems or at the request of the parents.

An additional field experiment was performed in North Carolina in 1991, partly in response to the criticisms of the STAR program design. Students were assigned to classes of either 15 or 25 students in first through third grade. Achilles et al. (1995) finds that students in the smaller classes achieved test scores that were .45 and .56 standard deviations higher than their peers in the larger classes, on the mathematics and reading tests, respectively.

Other studies have focused on non-experimental data to estimate the impact of class size on elementary student performance. For example, Angrist and Lavy (1997) and Hoxby (2000) rely on discontinuities in class size that arise from mandated caps on class size. For example, Angrist and Lavy examine class size information in Israel where Maimonides' Rule dictates that classes should be no larger than forty students. Similarly, Hoxby uses twenty four years of data from Connecticut for fourth and sixth graders. Some school districts cap enrollment at twenty five students per class. This creates variation in class sizes when the total number of students surpasses multiples of twenty five, and thus prompts an increase in the number of classes offered. Angrist and Lavy generally find a significant class size effect, while Hoxby generally does not find a significant class size effect.

Other studies have examined the influence of class sizes on older students using data from the National Educational Longitudinal Study (NELS). This study began in 1988 with a national (U.S.) sample of eighth grade students. A subset of the participants was re-interviewed in 1990, 1992, 1994, and again in 2000. Goldhaber and Brewer (1997) incorporate an exhaustive set of controls for teacher ability and behavior, as well as numerous student attributes including previous test scores to control for student ability. They report a positive and statistically significant effect of class size on standardized test scores in the tenth grade. Akerhielm (1995)

also utilizes the NELS data set and attempts to account for the possible endogeneity of class size in standardized test scores due to less able students being assigned to smaller classes. She uses two stage least squares with the average class size for a given subject in a school and eighth grade enrollment in the school as identifying instruments. She finds that once the endogeneity of class size is accounted for the influence of class size on test scores goes from positive to negative, and in some instances is statistically significant.

The studies most relevant to this paper examine the influence of class size on student outcomes in tertiary education. Because most students in higher education cannot be made or incentivized to take a single standardized test, the most commonly examined measures are student grades in a course and student self reported satisfaction with or assessment of the quality of a course or instructor. Studies that have examined student assessment of a course universally agree that class size has a negative impact on student course evaluations, with larger courses receiving statistically significant lower scores than smaller courses. For example, Bedard and Kuhn (2008) examine student evaluations of economics courses at the University of California at Santa Barbara, from 1997 to 2004. They find a large, highly significant, and nonlinear negative impact of class size on student evaluations of instructor effectiveness. Their result is robust to instructor and course fixed effects. Similarly, Walia (2008) utilizes 19 semesters of student evaluations of economics courses at Kansas State University. Once again class size is found to have a negative and statistically significant impact on student course evaluations. Illustrating that this result is not unique to economics students nor the United States, Westerlund (2008) reports that increases in the size of mathematics classes at Lund University in Sweden leads to significantly lower student course evaluations there, as well.

Fewer studies still have examined the impact of class size on student performance in higher education. Bandiera et al. (2009) examine administrative records from a leading UK

university and find a significant negative, but highly non-linear effect of class size on student tests results. They conclude that changes in class size have a significant impact on student performance but only at the very top and bottom of the class size distribution. Furthermore, they find that students at the top of the grade distribution are most negatively affected by class size, particularly in large class sections. They rule out class size effects being due to non-random assignment of faculty across class size, student self-selection into class size, omitted inputs, or changes in grading policies based on class size. Similarly, Kokkelenberg et al. (2008) find that average grades decline significantly with class size at a public northeastern U.S. university. They find that grades drop dramatically with class size up to twenty students, and less steeply but nonetheless monotonically thereafter.

Correa (1993) posits a theoretical explanation for the importance of class size in the education production function that focuses on the role of individual faculty-student interaction. His model describes teachers that must weigh efforts directed to the whole class versus individual student attention. The larger the class the greater the instructor effort devoted to class-wide activities at the expense of individual attention. In this way, individual student learning and outcomes decline as class size increases. His model illustrates the importance of separating class size effects from total student responsibilities effects. Lazear (2001) outlines a theoretical model where class size itself is important due to the role that class size plays in setting the class room environment. Large classes (more students) may allow students to be more disruptive, allow them to "hide" from participation, engagement, or even attendance, while small classes may more easily lend themselves to pedagogical activities that improve learning, such as hands on activities and student-faculty classroom interaction.

While the existing literature on primary and secondary class size effects is mixed, the evidence of class size on student assessment and student grades in higher education is more

consistent. Unfortunately, the higher education literature suffers from a lack of random, substantial changes in class size within instructors and courses, and confounds class size effects with student oversight effects. This paper contributes to this literature by exploiting a natural experiment in class sizes that allows for an approximate fifty percent increases in class size, within instructor and course, without a change in the total number of students taught by that instructor.

III. Data

The data for this study come from administrative records and student course evaluations at a private, highly selective university on the east coast of the United States. The faculty and courses utilized in this analysis are restricted to the undergraduate business school within this university, as the student course evaluation instrument and the practice of super-sizing outlined above was limited to undergraduate business school courses at this institution. The sample period covers the academic years 1996 through 2008. This allows for three years of data preceding the implementation of the super-size policy, six years of super-sizing, and three years following the suspension of this practice. The sample includes 48 individual faculty members, 88 separate courses, and 1,928 course sections. In total, 8 faculty, 14 courses, and 84 sections were super-sized over this period.

The outcomes available for and examined in this paper are the average course ratings from the student course evaluations for: (1) overall instructor rating (1-5, with 5 being the best); (2) amount learned (1-5, with 5 being the most); (3) overall course rating (1-5, with 5 being the best), and (4) the average of two questions asking students their lowest and highest expected grade in the course (calculated as expected course GPA).

The independent variables fall into several categories. First, course specific variables such as the average grade point average of the students in the class, the gender mix of the class (proportion male and the proportion of the class the same sex as the instructor), the grade level mix of the class (proportion seniors, juniors, sophomores, and freshmen), the meeting time of the class (early or late), class length (50 versus 75 minutes) type of class (pre-business and required versus elective), and the prior level of student interest in the course. Additionally, two controls are included for institutional factors: (a) a trend variable to allow for observed evaluation inflation and (b) an "online" binary indicating the evaluation was administered online (discussed later). The second category of independent variables captures the number of students. Class size (CS) is taken from administrative records that list the actual number of students enrolled in the class after the drop/add period in the second week of a fifteen week semester. Additionally, we are also able to control for student load, the total number of students that each instructor had enrolled in all of his or her class sections that semester (SL). This will allow us to determine if total number of students taught per semester plays a role in student assessments independent from any class size effect. An increase in either could affect the amount of individual attention students receive. The final set of independent variables is instructor and course fixed effects. Additionally, we include a trend variable to control for an upward drift in average course ratings that occurred at this institution over this time period. This allows us to control for course, faculty, and trend effects that otherwise would confound the estimation of impacts on average course ratings for those super-sizing their course sections, and thus dramatically increasing their class sizes, to those who did not see significant changes in their class sizes over the same period.

The econometric specification we estimate is:

$$(1) \qquad Y_{ijt} = \beta_0 + \beta_1(X_{ijt}) + \beta_2(SL_{jt}) + \beta_3(CS_{ijt}) + \delta_i + \alpha_j + \nu_t + \varepsilon_{ijt}$$

Where Y represents the average course outcomes outlined above for instructor i, in course j, at time (semester) t. The vector X indicates course specific variables such as average student grade point average which may influence student evaluations, SL is the student-load of the faculty member that semester over all courses, and CS is the class size. The error components δ, α, ν represent instructor, course, and time effects, respectively. Equation (1) is estimated using weighted least squares to account for the heteroskedasticity inherent when using dependent variables derived from averages calculated across groups of varying size.

IV. Empirical Results

We begin by examining the influence of class size and student load on the self-reported outcomes of: (1) how much the students reported learning in the course; (2) overall instructor rating for the course; (3) overall course rating; and (4) the average expected grade for the course. These four measures are used to gauge the impact of class size and student load on course outcomes. Table 2 presents the results of the weighted least squares regressions controlling for dummy variables for whether it is a required class, a pre-business class (microeconomics, macroeconomics, two introductory accounting classes, and a business statistics class), early morning class (classes that begin before 9AM), late afternoon class (classes that begin after 2PM), met three times a week for 50 minutes (versus twice a week for 75 minutes), and whether the evaluations were administered online versus in paper (faculty had the option of using online evaluations the last two semesters of the sample period). Additionally, controls for the proportion of the class that was male, the proportion of the instructor's sex, the proportion in each grade (sophomore, junior, senior), the average self-reported GPA of the class, and the

reported level of interest in the subject matter prior to the course. All regression results include a trend term and faculty fixed effects, so that all estimated coefficients represent within faculty effects of changes in course enrollment and student load on outcomes.

Before turning to the enrollment effects it is interesting to note that a few of the course characteristics have significant affects on course outcomes. For example, required courses have significantly (at the 99 percent level) lower levels of amount learned, instructor and course rating, and expected grades. Similarly, pre-business courses have significantly higher instructor and course ratings (but not amount learned or expected grade). Faculty interested in maximizing their course evaluation ratings should avoid early morning classes as these usually earn lower ratings in the four outcomes, and should choose courses that meet three times a week for fifty minutes as these generally receive higher average ratings for amount learned and instructor and course ratings. It is also interesting to note that the higher the proportion of the class that is male the lower the average rating for amount learned and instructor and course rating, while the proportion of the class that is the same sex as the instructor raises these scores. This implies that for male instructors these two effects largely cancel each other out, but for female instructors the higher the proportion of the course that is male has an amplifying, and negative, impact on their ratings. Upper classmen give slightly higher marks for the instructor and course than freshmen, and expect higher grades, as well.

The higher the average GPA of the class the lower the reported amount learned, instructor rating, and course rating. As expected, the higher the average GPA the higher the expected grade in the class. It is not surprising, but nonetheless interesting, that the level of reported prior interest in the subject matter has a positive and statistically significant (99 percent level) impact on all four course outcomes. Students' level of a priori enthusiasm for the material has a large impact on how they rate the outcomes of a course.

Finally, turning to the influence of course size and student load on course outcomes we find that the larger the section size (number enrolled) the lower the self-reported amount learned, the instructor rating, the course rating, and the expected grade. Clearly, students feel that they learn less and get less out of large class sections, even conditional on the number of total students an instructor is responsible for. Similarly, faculty who are responsible for many students in a semester (student load), even conditional on class size, receive lower average scores for amount learned (at the 95 percent level), and lower scores for instructor and course ratings, although these results are not statistically significant (p-values of .378 and .295, respectively). Additionally, the more students a faculty member is responsible for in a semester the higher the expected grades of his or her students (significantly different from zero at the 95 percent level).

As mentioned above, this institution experimented in the middle years of this sample with combining three sections of a course into two larger sections. This introduced substantial increases in class size, both within faculty members and within courses. This consolidation usually involved increasing class sizes from approximately 30 students to 45 students per section. This policy allows for the estimation of substantial changes in class size on student outcomes, while holding student load constant, within courses and faculty. Table 3 presents the results incorporating course fixed effects into the regressors. Note that two variables, required course and pre-business, are not included in Table 3 because course type is subsumed within the course fixed effects.

The results are largely consistent with those from Table 2. The major exception to this consistency is that the higher the proportion of upper-classmen in the course the higher the instructor rating, course rating, and expected grade. While this was true excluding course fixed effects, the results now are even larger and statistically significant at a higher level of confidence. In terms of class size, courses with larger enrollments continue to receive

significantly lower ratings on amount learned, instructor rating, and course rating. The impact of class size on expected grade is no longer statistically significant, once course fixed effects are incorporated. Even conditional on class size, student load has a negative and statistically significant impact on amount learned, and a positive and statistically significant impact on expected grade, and is near statistical significance, at the 90 percent level, for the instructor and course ratings (p-values of .103 and .133, respectively). Qualitatively similar results were found when using a dummy variable for super-sized versus non-super sized sections, rather than the continuous actual class enrollment. Similarly, a quadratic in class size was also attempted and found not to be statistically significant. For these reasons, class size is entered into the regressions linearly.

While it is clear from these results that students generally rate courses and faculty in large sections less favorably than smaller sections, it is not obvious as to why that is the case. In Table 4 we estimate the impact of class size and student load on specific course attributes and faculty practices. In each case, all of the above control variables, including faculty and course fixed effects, are also included in the regressions (but the results are not shown). We present only the results of the influence of class size and student load on the course attributes listed down the left hand side of Table 4.[2] We find that class size has a negative and statistically significant (at the 95 percent level or greater) impact on the amount of critical and analytical thinking required in the course, the clarity of presentations, the effectiveness of teaching methods, the daily preparedness of the instructor for class, the instructor's effectiveness in stimulating student interest, the instructor's enthusiasm for the class, the instructor's availability outside of class, the

[2] All regressions in Table 4 are based on 1,928 sections/observations with the exception of the amount of workload in the course. This question changed in the 2008 course evaluations, and so only the 1,794 sections prior to 2008 are used in the workload regression, and in the results presented in Table 5.

instructor's respect for his/her students, the applicability of graded material to the course content, the adequacy of comments on student work, the timeliness of feedback, and even the usefulness of the text to learning. On the other hand, class size was found to have a positive and significant impact on the reported pace of the course. The only course attributes that were not statistically significantly influenced by class size were the level of course workload and the level of difficulty of the course.

Similarly, student load has a negative and statistically significant impact on the level of workload, the level of difficulty, the amount of critical and analytical thinking, the clarity of presentations, the effectiveness of presentations, the instructor's daily preparedness for class, the instructor's ability to stimulate student interest, the instructor's enthusiasm for the class, the applicability of graded material to the course content, the adequacy of comments on student work, and the timeliness of feedback. Student load was not found to significantly influence the pace of the course, the availability of the instructor outside of class, his/her respect for students, and the usefulness of the text to learning.

Clearly, class size and student load are perceived by students as significantly altering many aspects of their courses. In the following analyses, we investigate the impact of these course attributes on the four course outcomes discussed above to determine if class size and student load, in and of themselves, have a direct influence on students' course ratings or if they work through these intermediate course attributes to influence overall course outcomes, or both. Table 5 incorporates all of the course attributes from Table 4 into the course outcomes regressions, including all of the previous regressors. The results reveal that many of the course attributes negatively affected by class size and student load positively influence the four course outcomes. For example, the level of critical and analytical thinking, the instructor's daily preparedness for class, the instructor's effectiveness in stimulating student interest, and the

applicability of graded material to the course content all positively and significantly influence all four course outcomes (and are themselves negatively influenced by class size and student load). Similarly, the clarity of presentations and effectiveness of teaching methods positively impact the course outcomes of amount learned, instructor rating, and course rating. Many of the other intermediate course attributes significantly influence one or more of the four course outcomes, and are themselves a function of class size and student load. Nonetheless, class size continues to have a negative and statistically significant impact on the amount learned, instructor rating, and course rating conditional on the course attributes, although the magnitudes of these effects are dramatically smaller once the course attributes are incorporated. Student load is no longer significant in explaining variation in the amount learned and expected grade, and in fact is now positive and statistically significant (at the 90 percent level) in influencing the instructor and course rating. These results suggest that student load's negative influence on course outcomes works through altering the attributes of the instructors' courses. Faculty who must handle a large number of students in a semester appear to change their courses in ways that negatively affect course outcomes. Class size has a similar indirect impact in that larger classes seem also to prompt instructors to alter their courses in ways that negatively affect course outcomes, but that class size itself has a reinforcing direct negative impact on course outcomes, as well.

V. Conclusion

The evidence found in this analysis unequivocally leads to the conclusion that both class size and the total number of students that a faculty member is responsible for teaching have a negative impact on the self-reported outcomes of amount learned, instructor rating, course rating, and expected course grade. These negative relationships between class size and student load with student outcomes are found conditional on faculty and course fixed effects, and thus

represent changes in student outcomes within instructors and courses, and are not attributable to endogenous variation in class size across instructors. Additionally, the analysis above reveals that class size and student load primarily influence student outcomes by altering certain aspects of courses that students find beneficial and helpful in learning. For example, large class sizes and higher student loads are correlated with less critical and analytical thinking, less clarity in class presentations, and lower ratings on the instructor's ability to stimulate student interest. In turn, these course and instructor attributes are positively related to students' course assessments.

Reducing class sizes will help to improve student outcomes, but ignores the impact that student load plays in how faculty structure their courses. Reducing class sizes and the total number of students that a faculty member is responsible for teaching in a semester will lead to significant improvements in student outcomes. Administrative policies of hiring adjunct faculty to teach numerous sections of a course in order to minimize class sizes ignores the important role that total student responsibility plays in how faculty actually teach those courses. Course attributes important to student learning and how much students get out of a course suffer when class size and student loads increase. Policies designed to reduce class sizes in order to fare better in institutional rankings should be weighed against the impact these policies may have on student load and the equally important impact it plays in student outcomes.

References

Achilles, C.M., P. Harman and P. Egelson (1995). "Using Research Results on Class Size to Improve Pupil Achievement Outcomes." *Research in Schools*. vol 2(2), pp. 23-30.

Akerhielm, Karen (1995). "Does Class Size Matter?" *Economics of Education Review*. vol 14(3), pp. 229-41.

Angrist, Joshua, and Lavy, Victor (1999). "Using Maimondides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics*, vol??, pp. 533-75.

Bandiera, Oriana; Larcinese, Valentino; Rasul, Imran (2009). "Heterogeneous Class Size Effects: New Evidence from a Panel of University Students." C.E.P.R. Discussion Papers, CEPR Discussion Papers: # 7512, 2009.

Bedard, Kelly; Kuhn, Peter (2008). "Where Class Size Really Matters: Class Size and Student Ratings of Instructor Effectiveness." *Economics of Education Review*, vol. 27, no. 3, June 2008, pp. 253-65.

Correa, Hector (1993). "An Economic Analysis of Class Size and Achievement in Education." *Education Economics*, vol 1(2), pp. 129-35.

Finn, Jeremy (1998). "Class Size and Students at Risk: What is Known? What is Next?" Department of Education, Washington, DC.

Finn, Jeremy, and Achilles, Charles (1990). "Answers and Questions about Class Size: A State-wide Experiment." *American Educational Research Journal*, vol 27(3), pp. 557-77.

Goldhaber, Daniel, and Brewer, Dominick (1997). "Why Don't Schools and Teachers Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity." *Journal of Human Resources*, vol 32, pp. 505-23.

Hoxby, Caroline (2000). "The Effects of Class Size and Composition on Student Achievement: New Evidence from Natural Population Variation." *Quarterly Journal of Economics*, pp. 1239-85.

Johnes, Geraint; Johnes, Jill, eds. International Handbook on the Economics of Education. Cheltenham, U.K. and Northampton, Mass.: Elgar, 2004, pp. 860.

Kokkelenberg, Edward C; Dillon, Michael; Christy, Sean M (2008). "The Effects of Class Size on Student Grades at a Public University." *Economics of Education Review*, vol. 27, no. 2, April 2008, pp. 221-33.

Krueger, Alan (1999). "Experimental Estimates of Education Production Functions." *Quarterly Journal of Economics*, vol 114(2), pp. 497-532.

Lazear, Edward (2001). "Educational Production." *Quarterly Journal of Economics*, vol 116(3), pp. 777-803.

Mosteller, Frederick (1995). "The Tennessee Study of Class Size in the Early School Grades." *The Future of Children*, vol 5(2), pp. 113-27.

Nye, Barbara, and Larry Hedges (2001). "Are Effects of Small Classes Cumulative? Evidence from a Tennessee Experiment." *The Journal of Educational Research*, vol 94(6), pp. 336.

Nye, Barbara, Hedges, Larry, and Konstantopoulos, Spyros (2000). "The Effect of Small Classes on Academic Achievement: The Results of the Tennessee Class Size Experiment." *American Educational Research Journal*, vol 37(1), pp. 123-52.

_____ (2001). "The Long Term Effects of Small Classes in Early Grades: Lasting Benefits in Mathematics Achievement in Grade 9." *The Journal of Experimental Education*, vol 69(3), pp. 245.

Walia, Bhavneet (2008). Three Essays in Health and Labor Economics Ph.D. Dissertation. Kansas State University, 2008.

Westerlund, Joakim (2008). "Class Size and Student Evaluations in Sweden." *Education Economics*, vol. 16, no. 1, March 2008, pp. 19-28.

Word, Elizabeth, Johnston, John, Pate-Bain, Helen, Fulton, DeWayne, Boyd-Zaharias, Jayne, Achilles, Charles, Lintz, Martha, Folger, John, and Breda, Carolyn (1990). "The State of Tennessee's Student/Teacher Achievement Ratio Project Final Summary Report 1985-1990." Tennessee State Department of Education.

Table 1
Summary Measures

	Minimum	Maximum	Mean	Std. Dev.
Class Size	2	45	23.39	6.862
Student load	10	150	63.46	18.751
Pre-business Course	0	1	.24	.426
Required Course	0	1	.46	.499
Early	0	1	.10	.297
Late	0	1	.28	.451
Fifty Minute Classes	0	1	.31	.463
Male	0	1	.60	.151
Same Gender as Instructor	0	1	.58	.162
Percent Sophomores	0	1	.23	.315
Percent Junior	0	1	.34	.325
Percent Senior	0	1	.37	.376
GPA	2.75	3.69	3.18	.142
Expected Grade	2.04	3.97	3.05	.292
Interest in this subject prior to the course	1.75	5.00	3.47	.545
Course Workload	2.27	4.92	3.39	.463
Course Level of Difficulty	2.45	5.00	3.66	.501
Level of Critical & Analytical Thinking	2.20	5.00	4.00	.447
Pace of Course	2.20	4.45	3.15	.207
Number of Evals. Completed by Student	1.00	4.67	2.66	.824
Clear and Understandable Presentation	2.00	5.00	4.09	.485
Effectiveness of Teaching Methods	2.04	5.00	4.05	.500
Instructor's Daily Preparation for Class	2.54	5.00	4.50	.361
Effectiveness in Stimulating Interest	2.03	5.00	3.94	.514
Enthusiasm for Teaching Course	2.87	5.00	4.48	.366
Availability Outside the Classroom	2.36	5.00	4.41	.370
Respect for Students in the Class	2.50	5.00	4.34	.425
Graded Material vs. Course Content	1.44	5.00	3.98	.499
Adequacy of Comments on Student Work	1.78	5.00	3.76	.483
Timeliness of Feedback on Student Work	1.26	5.00	4.26	.496
Usefulness of text, etc. to Learning	1.66	5.00	3.84	.512
Amount Learned as a Result of Course	2.45	5.00	4.12	.449
Instructor's Overall Teaching Ability	2.08	5.00	4.18	.479
Overall Quality of Course	2.30	5.00	4.00	.475
Number of Faculty	48			
Number of Courses	88			
Number of Sections	1,928			

Table 2
Regression Results
Impact of Enrollment on Course Outcomes

	How much learned		Instructor Rating		Course Rating		Expected Grade	
Intercept	3.797		4.243		3.453		.759	
trend	.014	***	.004	***	.008	***	-.001	***
	(0.001)		(0.001)		(0.001)		(0.001)	
Class Size	-.009	***	-.009	***	-.009	***	-.004	***
	(0.001)		(0.002)		(0.002)		(0.001)	
Student load	-.001	**	-.001		-.001		.001	**
	(0.001)		(0.001)		(0.001)		(0.0003)	
Required	-.129	***	-.109	***	-.132	***	-.064	***
	(0.026)		(0.030)		(0.028)		(0.016)	
Pre-Business	.063		.193	***	.152		.020	
	(0.050)		(0.057)		(0.054)		(0.030)	
Early	-.095	***	-.107	***	-.110	***	-.067	***
	(0.026)		(0.030)		(0.029)		(0.016)	
Late	-.002		.008		.018		.019	*
	(0.018)		(0.021)		(0.020)		(0.011)	
50 Min. Class	.035	*	.052	**	.055	***	.003	
	(0.019)		(0.021)		(0.020)		(0.011)	
Online	-.096		-.042		-.063		-.076	
	(0.081)		(0.092)		(0.087)		(0.048)	
Male	-.126	*	-.149	*	-.168	**	-.005	
	(0.069)		(0.078)		(0.074)		(0.041)	
Same Gender	.118	*	.201	***	.184	***	.068	*
	(0.066)		(0.076)		(0.072)		(0.040)	
Percent: Sophomore	-.011		.053		.078		.081	**
	(0.057)		(0.065)		(0.061)		(0.034)	
Junior	.020		.140	*	.145	*	.113	***
	(0.072)		(0.082)		(0.078)		(0.043)	
Senior	-.083		.185	**	.142	*	.183	***
	(0.073)		(0.084)		(0.079)		(0.044)	
GPA	-.170	***	-.189	**	-.193	***	.584	***
	(0.066)		(0.075)		(0.071)		(0.039)	
Prior Interest	.320	***	.229	***	.356	***	.091	***
	(0.023)		(0.026)		(0.025)		(0.014)	
Faculty Effects	Yes		Yes		Yes		Yes	
R-squared	0.596		0.541		0.575		0.645	

Table 3
Regression Results
Impact of Enrollment on Course Outcomes with Course Fixed Effects

	How much learned		Instructor Rating		Course Rating		Expected Grade	
Intercept	3.317		3.807		2.998		.554	
Trend	.014	***	.005	***	.009	***	.000	
	(0.001)		(0.001)		(0.001)		(0.001)	
Class Size	-.009	***	-.008	***	-.009	***	-.001	
	(0.002)		(0.002)		(0.002)		(0.001)	
Student load	-.001	**	-.001		-.001		.001	*
	(0.001)		(0.001)		(0.001)		(0.0003)	
Early	-.085	***	-.095	***	-.102	***	-.065	***
	(0.026)		(0.030)		(0.028)		(0.014)	
Late	.005		.007		.017		.007	
	(0.018)		(0.021)		(0.020)		(0.010)	
50 Min. Class	.021		.026		.029		.008	
	(0.020)		(0.023)		(0.021)		(0.011)	
Online	-.116		-.065		-.077		-.070	
	(0.078)		(0.090)		(0.084)		(0.043)	
Male	-.114		-.172	**	-.207	***	.042	
	(0.073)		(0.084)		(0.079)		(0.040)	
Same Gender	.109		.210	***	.213	***	.043	
	(0.068)		(0.079)		(0.074)		(0.038)	
Percent:								
Sophomore	.082		.185	***	.209	***	.123	***
	(0.058)		(0.067)		(0.063)		(0.032)	
Junior	.189	**	.335	***	.353	***	.147	***
	(0.075)		(0.087)		(0.081)		(0.042)	
Senior	.109		.354	***	.328	***	.147	***
	(0.080)		(0.093)		(0.087)		(0.045)	
GPA	-.123	*	-.205	***	-.179	**	.577	***
	(0.068)		(0.078)		(0.073)		(0.038)	
Prior Interest	.371	***	.338	***	.433	***	.119	***
	(0.030)		(0.034)		(0.032)		(0.016)	
Faculty Effect	Yes		Yes		Yes		Yes	
Course Effect	Yes		Yes		Yes		Yes	
R-squared	0.656		0.598		0.634		0.736	
No. of obs.	1928							

Table 4
Regression Results
Impact of Class Size and Student load on Course Attributes
Dependent variables are in the rows

	Class Size		Student load		R-squared
Workload	-.001		-.002	***	0.765
	(0.001)		(0.001)		
Difficulty	.002		-.002	***	0.829
	(0.001)		(0.001)		
Critical Thinking	-.003	***	-.002	***	0.794
	(0.001)		(0.001)		
Pace	.002	**	-.0003		
	(0.001)		(0.003)		
Clarity	-.007	***	-.001	**	0.617
	(0.002)		(0.001)		
Effectiveness	-.007	***	-.001	*	0.61
	(0.002)		(0.001)		
Preparation	-.005	***	-.001	**	0.648
	(0.001)		(0.001)		
Stimulates Interest	-.006	***	-.002	***	0.648
	(0.002)		(0.001)		
Enthusiasm	-.003	***	-.002	***	0.687
	(0.001)		(0.001)		
Availability	-.005	***	-.001		0.522
	(0.001)		(0.001)		
Respect	-.004	***	-.001		0.631
	(0.001)		(0.001)		
Graded Material	-.005	***	-.001	**	0.611
	(0.002)		(0.001)		
Quality of Feedback	-.005	***	-.001	**	0.625
	(0.002)		(0.001)		
Timeliness of Feedback	-.005	***	-.001	**	0.662
	(0.002)		(0.001)		
Usefulness of Text	-.006	***	-.0001		0.609
	(0.002)		(0.001)		

Table 5
Regression Results
Impact of Course Attributes on Course Outcomes

	How much learned		Instructor Rating		Course Rating		Expected Grade	
Intercept	-.147		-.626		-.403		1.203	
Trend	.003	***	-.008	***	-.006	***	-.005	***
	(0.001)		(0.001)		(0.001)		(0.001)	
Class Size	-.004	***	-.001	*	-.003	***	.0001	
	(0.001)		(0.001)		(0.001)		(0.001)	
Student load	.0001		.001	*	.001	*	.0001	
	(0.001)		(0.001)		(0.0005)		(0.0005)	
Workload	.101	***	-.019		.002		.048	***
	(0.020)		(0.018)		(0.018)		(0.017)	
Difficulty	.017		.019		.002		-.318	***
	(0.025)		(0.023)		(0.023)		(0.021)	
Critical Thinking	.168	***	.070	***	.113	***	.050	***
	(0.022)		(0.020)		(0.020)		(0.019)	
Pace	-.060	*	-.037		-.110	***	-.050	***
	(0.033)		(0.030)		(0.030)		(0.019)	
Clarity	.068	**	.380	***	.177	***	-.036	
	(0.028)		(0.025)		(0.025)		(0.024)	
Effectiveness	.169	***	.173	***	.183	***	.024	
	(0.028)		(0.026)		(0.026)		(0.024)	
Preparation	.044	*	.146	***	.073	***	-.049	**
	(0.026)		(0.023)		(0.023)		(0.022)	
Stimulates Interest	.276	***	.164	***	.263	***	.058	***
	(0.022)		(0.020)		(0.020)		(0.019)	
Enthusiasm	.014		.132	***	.044	*	-.019	
	(0.025)		(0.023)		(0.023)		(0.022)	
Availability	-.008		-.008		-.022		-.011	
	(0.017)		(0.016)		(0.016)		(0.015)	
Respect	-.035	*	.101	***	.041	**	.018	
	(0.019)		(0.017)		(0.017)		(0.016)	
Graded Material	.128	***	.063	***	.146	***	.133	***
	(0.018)		(0.017)		(0.017)		(0.016)	
Quality of Feedback	-0.005		-0.016		-0.011		-0.028	*
	(0.019)		(0.017)		(0.017)		(0.016)	
Timeliness of Feedback	0.020		0.000		0.003		-0.013	
	(0.014)		(0.013)		(0.013)		(0.012)	
Usefulness of Text	0.140	***	-0.019		0.079	***	-0.016	
	(0.013)		(0.012)		(0.012)		(0.011)	
Faculty Effects	Yes		Yes		Yes		Yes	
Course Effects	Yes		Yes		Yes		Yes	
R-squared	0.908		0.933		0.931		0.836	

Note: All regressions include variables for time of class, length of class, gender mix, grade mix, average GPA, and level of student interest.