Preliminary Draft – Comments Welcome
Please do not circulate or quote

Mass Instruction or Higher Learning? The Impact of Class Size in Higher Education

Eric P. Bettinger, Case Western Reserve University & NBER
Bridget Terry Long, Harvard University & NBER

February 2005

ABSTRACT

This paper attempts to measure the effects of collegiate class size on student outcomes such as persistence, subsequent course selection, and major choice. While class size is a perennial issue in literature focusing on primary and secondary schooling, it has been largely ignored in research on higher education. Nonetheless, the effect of class size in college is potentially very important as it impacts everything from competitive rankings (20 percent of the *U.S. News and World Report* formula focuses on class size) to university budgets. Moreover, the range of possible class sizes is substantial in higher education (as many as 700 per lecture in our data) suggesting large differences in students' classroom experiences and possibly large effects on students' outcomes as well. Few researchers have focused on collegiate class size because data are rarely available. In addition, there are inherent selection problems as higher ability students appear to systematically avoid large classes. To address this hole in the literature, we use a unique data set that includes detailed class size, faculty, and student characteristics and track two cohorts of nearly 41,000 students in four-year, public colleges in Ohio. Using an instrumental variables approach to control for selection bias, the preliminary results suggest that an increase in class size leads to a higher likelihood of dropping out and a reduction in subsequent student interest in a subject.

Send comments to bettinger@cwru.edu and longbr@gse.harvard.edu. The authors thank the Ohio Board of Regents for their support during this research project. Rod Chu, Darrell Glenn, Robert Sheehan, and Andy Lechler provided invaluable help with the data. Amanda Starc and James Carlson provided excellent research assistance. All opinions and mistakes are our own.

I. Introduction

Class size is a perennial topic in the economics of education. Policymakers focused on primary and secondary schooling continually debate the use of class size as a lever, and many states have passed legislation aimed at reducing class size (e.g. Morgan-Hart Class Size Reduction Act in California in 1989). In addition, researchers have examined the impact of class size on student achievement in the primary and secondary context (Hoxby 2000, Angrist and Lavy 1999, Krueger and Whitmore 2000). Yet the role and effects of class size in postsecondary settings are largely ignored by both researchers and policymakers.

The lack of research or policies on class size in higher education is not due to a perceived lack of importance. Indeed, measures of class size comprise 20 percent of the indicators used by the *U.S. News and World Report* in its popular college rankings, and class size is central to the costs and production of higher education. Class size may be even more important in the postsecondary context because there is much more variation in the size of college classes. While the debate in K-12 revolves around shifts in five to fifteen students per class, college courses can have as many as 1000 students. Over 25 percent of classes in our data are larger than 120, and the largest class has 704 students.

The lack of research on the effects of collegiate class size is largely because of a lack of data. None of the national datasets (e.g. National Education Longitudinal Study of 1988) contain information about students' class sizes. Additionally, an individual course may oftentimes have different class sizes for each of its meetings. For instance, a lecture on Monday may join multiple sections while a lecture on Wednesday may be a different size, and a lab on Friday may serve only a subset of student. In order to measure class size accurately, researchers must be able to identify the accurate size of each meeting and control for any variation within a given section of a course.

Selection issues are also pervasive in the study of the effects of collegiate class size thereby making research on the subject even more difficult. Unlike in primary and secondary school where class size is often assigned or otherwise exogenously determined, class size in college is more ubiquitous. Students choose classes, class sizes, and professors, and this may be done in nonrandom ways. Students who take a humanities-laden schedule will generally have smaller class sizes than students who choose physics or chemistry. Meanwhile, they may also have different underlying probabilities of success thus confounding any direct comparison of outcomes by class size. Moreover, students, regardless of the subject, often will try to take the "best" professor, and good teachers generally attract more students and thus have full classes. Large classes could therefore be associated with the "best" teachers making it difficult to compare class sizes.

To resolve these empirical difficulties, we measure the effects of collegiate class size using a comprehensive dataset of nearly 41,000 students collected and maintained by the Ohio Board of Regents. The sample includes transcripts for students who first enrolled in any public, four-year college in Ohio during Fall 1998 and Fall 1999. Detailed information is available on the class size for every meeting of every section of every course. Therefore, we distinguish between lecture and labs and track differences in class size throughout the week for each course. Moreover, the data identify the instructors of each class so that we can control for selection issues concerning specific faculty members in the analysis.

The analysis also uses exogenous variation in class size to address selection issues. Using additional information provided by the Ohio Board of Regents, the paper tracks changes in the pool of instructors for each department in each school over six years. These deviations result from retirements, leaves of absence, promotion decisions, and hiring decisions and influence the size of the courses students face. Using this variation, we further address the possible selection into specific classes.

There are two levels of analysis. In the first, we measure the effects of class size on students' engagement in a specific department. This is done by observing the class size in a student's initial introduction to a subject and tracing whether they take additional courses in the same subject as well as the number of subsequent credits taken and the likelihood the student majored in the department. Unbiased estimates are provided by using faculty fixed effects and exogenous variation in faculty composition as an instrumental variable. In the second level of analysis, we examine the effects of class size across all of the courses in a student's first semester on the likelihood that they dropout after one year. Throughout the analyses, we explore various extensions such as allowing specifications where class size has nonlinear effects on students.

II. Research on Class Size

Economists have long been interested in the effects of class size on student outcomes. However, most of the previous work focuses on primary and secondary schooling, and much of the research suffers from selection bias by comparing the outcomes of students in small and large classes without taking into account confounding factors that may influence both the class size the student faces and their outcomes. Several recent studies identify the effects of class size at the primary and secondary level using exogenous variation to address these concerns.

Mosteller (1995) examines the effects of the Tennessee STAR Program, an experiment which randomly assigned students in kindergarten through third grade to class sizes ranging from

22 to 15. While he found positive effects on early achievement, Krueger and Whitmore (2000) provide a more recent examination of the program's long term effects. Even years after the experiment, students who had been placed in a smaller class were more likely to attend college later and performed better on their SAT scores.

One of the criticisms of research based on the STAR Program is that teachers knew that they were part of an experiment – an experiment that could have led to a more desirable outcome for the teachers regardless of the impact on students. Hoxby (2000) raises this possibility and therefore utilizes an alternative strategy to deal with possible biases. She estimates the impact of class size using an instrumental variables strategy that exploits year to year changes in the student population. Due to randomness in the number of children ready for school each year, there is natural variation in class size over time. She uses this variation and finds class size does not appear to affect student outcomes.

Finally, Angrist and Lavy (1999) estimate the impact of class size on student outcomes in Israel. Following Maimonides' teachings centuries ago, Israel largely follows a formula in determining class size. As the average class sizes reach 40 in a specific grade, the school must add an additional class. In small schools, this creates dramatic variation. A school with 40 students should have two classes with an average of 20 students per class while a school with 39 students will only have one class. Angrist and Lavy (1999) use this exogenous variation to estimate the effect of class size on student outcomes and find that smaller classes lead to better student test scores.

There are a number of reasons why the results vary across these studies. First, the respective authors look at distinct populations. Class size may matter in some settings but not others. Second, the source of variation in each of these studies is from different parts of the class size distribution. The effects of class size could be very nonlinear with positive effects over some parts of the distribution and no effects at other parts. These nonlinearities are likely to be important in any analysis of class size in higher education because of the range of possible sizes. For example, in Ohio's public, four-year colleges, the 25th percentile of class size is 25 students while the 75th percentile is 123 students.

As noted above, there are few studies that evaluate the effects of class size on student outcomes in college. A number of papers which investigate the effects of class size on teaching pedagogy (e.g. Kennedy and Siegfried 1995; Becker and Powers 2001). Most of these studies attempt to help educators understand what changes in teaching style or presentation must be made in large classes. However, one study looks at the effects of class size in college by focusing on a single institution. Dillon, Kokkelenberg, and Christy (2002) find that class size affects the grade distribution of a course. Their paper compares small and large classes but does not fully address potential selection bias in how students sort into classes. Finally, Kennedy and Siegfried (1997) use a national economic education database (TUCE III) and treat each class as an observation. They find class size does not affect student achievement. To our knowledge, no study uses a large sample of individual-level data to examine the impact of class size on student outcomes.

III. Data and Empirical Strategy

Data

Our data come from institutional records collected and maintained by the Ohio Board of Regents (OBR). Through a collaborative agreement with the OBR, we have access to anonymous student-level data from all of Ohio’s public institutions. The data include information on student demographics, ACT and SAT test results and surveys, students' collegiate transcripts, enrollment patterns, and degree completion if applicable. Most students in Ohio take the ACT exam, and the ACT records include the highest test score as well as the most recent responses to the ACT survey which includes student-reported information on high school performance and their intended college major.

This study focuses on students whose first collegiate enrollment was in Fall 1999 or Fall 1999 at one of Ohio's public, four-year campuses. We restrict the sample to full-time students who were age 18, 19, or 20 in the year that they first enrolled in college and for whom we have ACT records. The ACT restriction affects some of the out-of-state students at Ohio colleges while the age restriction limits our analysis to traditional college students. Nontraditional, older students are excluded because they often have different motivations and background experiences.

One of the strengths of the data is that we observe students transcripts even if they transfer out of one of the public, four-year colleges in our sample. Most studies in higher education cannot track these students and consider them dropouts; however, many students transfer to other schools during their collegiate careers and capturing this mobility is important in trying to measure outcomes. We observe all classes that these students enrolled in so long as they continued to attend a public college or university within the state. While our capability of tracking students exceeds that of other studies, we cannot track students who transfer from Ohio public institutions to private colleges or institutions located in other states. These students are indistinguishable in the data from students who truly withdraw from higher education altogether. This potential measurement error, however, should be very small since the percentage of students counted as dropout who instead likely transferred makes up a small fraction of the total number of observed dropouts.[1]

Empirical Strategy

The paper analyzes two types of outcomes: one at the student level and another at the student-by-subject level. For the former outcome, we relate the overall class size faced by a student during their first semester to persistence. In comparison, the student-by-subject outcomes focus on a student's interaction with a particular department. We identify the first time a student takes a class in a specific subject and note the class size of that course. Then, using college transcripts, we measure subsequent student interest: whether the student ever took additional courses in the same department, the number of additional credits taken, and whether the student elected to major in that subject. Because specific variables from the student-by-subject level analysis shed light on our empirical strategy for estimating effects at the student level, we explain this empirical strategy first.

Our basic student-by-subject empirical strategy compares the outcomes measuring student interest as a function of student characteristics and class size:

(1) $$y_{ik} = X_i\beta + Z_{ik}\gamma + f(class\_size_{ik}, \lambda) + FE + \varepsilon_{ik}$$

Equation (1) models these outcomes for student *i* in subject *k*. Student outcomes are a function of: (i) student-specific characteristics (X_i) including race, gender, and ACT scores and (ii) students-subject-specific characteristics (Z_{ik}) such as whether the student intended to major in

[1] The Integrated Postsecondary Education Data System (IPEDS) tracks the number of transfers at each institution but does not record the state of residence of transfer students although it does track the states of residence for incoming freshman. Assuming that transfer students are geographically representative of the incoming freshman class, then one would expect around 650 Ohio students to transfer to the non-Ohio schools with substantial Ohio enrollments. If we further assume that *all* 650 transfer students just finished their first year of school, then about 4.3 percent of observed dropouts are actually transfer students.

this specific department as designated on their ACT survey and an interaction between this "pre-college" major and students' ACT scores.

We track three different class sizes. For every section of every course, our primary metric is the lecture class size. If the lecture size varies within a section over a week, we use the average lecture size. We attempt multiple functional forms for average class size in an attempt to identify nonlinearities in class size. If a lab exists, we control for its size separately from the lecture size. Finally, we also control for class sizes in other types of sections. These special meetings may include required study groups, discussion sessions, or active learning exercises. As with the class size in the labs, we control for whether a specific section includes one of these other types of meetings and for the class size in this meeting. For now, we do not specify a functional form for class size.

The final term in Equation (1) focuses on sets of fixed effects. In our baseline specifications, we include fixed effects for the campus of attendance, course subject, and student cohort (Fall 1998 versus Fall 1999). When estimating Equation (1), we report standard errors that control for correlation across observations of the same student.

Equation (1) will be biased if we omit variables related to student interest in a subject that are correlated with class size. For example, faculty members may play an important role in influencing subsequent student interest in a subject, and these same faculty members are likely to attract students at different rates than other instructors. Failure to control for these faculty interactions would bias our results, and so, in some of our specifications, we include faculty fixed effects. To control for the fact that faculty members may draw different class sizes across different courses, we can also include faculty-by-course effects. When using faculty fixed effects, the campus and subject fixed effects are dropped due to collinearity.

While we can control for faculty fixed effects, there may be other unobservable variables (e.g. underlying student interest) that affect the class sizes that students choose and the likelihood of subsequent engagement in a subject. The first way we address this is by controlling for the student's intended college major as given when they took the ACT. In addition, we use an instrumental variables specification to overcome this selection bias. A good instrument should be correlated with the endogenous regressor (class size) but uncorrelated with the residual outcome (subsequent engagement). Movements in the composition of a department's faculty may be such an instrument. Using additional data from OBR, we measure the six-year steady-state proportion of classes taught by assistant, associate, and full professors and non-ranked instructors (i.e. adjuncts) for each term. Departments deviate from this steady state each term due to retirements, promotion decisions, sabbaticals, and unexpected outcomes in the hiring process. Such deviations are related to the size of the courses taught in that subject that term. For example, if a faculty member is absent one term, it puts pressure on his or her department to consolidate classes during that period. In the presence of the faculty fixed effects, the variation from which we identify is from the interaction over time between department-wide shifts in the distribution of instructors and differences in an individual professor's class size. As we show below, increases in the proportions of assistant and associate professors teaching in the department lead to slightly smaller classes while increases in the proportion of full professors teaching in a semester leads to slightly larger classes.

The second type of outcomes we examine relate to student persistence. Topp (1984) suggests that large class sizes early in a student's academic career may alienate students leading to disassociation with the institution and consequently student withdrawal. Therefore, we measure the effects of the overall class size faced by a student during his or her first semester on the likelihood that the student drops out.

To estimate the effects of class size on these student-level outcomes, we estimate Equation (2):

(2) $$y_k = X_i\beta + f(class\_size_i, \lambda) + FE + \varepsilon_i$$

where the variables are defined as before. To estimate Equation (2), we aggregate the student-by-subject data to the student level.

The advantage of aggregating the data to the student level is that we can control for the combinations of courses and class sizes that students face. It may not be that the actual class size of any course matters but rather the minimum or maximum class size across students' schedules that matters, and our analysis tries different specifications to test for this. However, the disadvantage of aggregating the data by student is that is complicates our instrumental variables strategy. Just as we can aggregate the data to control for possible effects of class size throughout an individual's entire schedule, we can aggregate our instrument according to the changes in the faculty across all of the departments where the student took classes. Hence, instead of the instrument being movements away from the steady state level of staffing in one department, the instrument would be movements away from the steady state across all of the departments in which the student took a course. While this may provide a solution, it does not control for composition. The instruments will be based in part on the set of courses an individual took. Students who take a disproportionate number of courses in one department will have different variation in their instrument than students who took courses more heavily from other departments.

To control for this possibility, we insert a number of "portfolio" controls in the student-level regressions. They consist of a set of fixed effects which account for the combination of courses taken. For example, one "portfolio" could be the following set of courses during one semester: English, math, economics, physics, and biology. We insert a dummy variable for this portfolio identifying every student who took this combination of subjects. Another portfolio might represent the students who took courses only in English, math, and economics during the term. Given all the possible combinations, there are nearly 6,000 portfolios represented in our data. Especially in the context of the instrumental variables, the portfolio effects allow us to identify off differences across time between the class sizes and composition of faculty for students who take similar courses.

IV. Empirical Results

Table 1 shows some descriptive statistics for our sample. The top part of the table shows characteristics at the student level while the bottom section displays characteristics at the student-by-subject level. Among the students in our sample, about 85 percent are white and eight percent are black. The average ACT score is 22.4 and the average age of the students is slightly over 18. Similar to the national trends, there are more females than males at the colleges as only about 45 percent of the sample is male. The mean additional hours students complete in a given subject after their introductory course is close to six, and about six percent of students who take an initial course in a subject go on to major in the subject. Only approximately five percent of the courses in our sample are lab sections and about four percent of enrollments are in subjects in which students expressed interest as a potential major prior to college.

Tables 2 and 3 show the average lecture sizes for schools and subjects in our data. Table 2 shows the average lecture size for each campus while Table 3 shows the average lecture size by subject comparing both selective and non-selective campuses. The averages are computed at the student-level, so they reflect the average experience of entering, full-time freshmen. The average class sizes in first semester courses are quite high for most campuses. In no school is the average less than 50, and over one-third of campuses have average class sizes in first semester courses that exceed 100. The medians are much smaller reflecting the skewness of the class size distribution; however, more than half of the campuses still have median class sizes exceeding 50. At College G in particular, the median class size is extremely high.

As for departments (Table 3), English classes tend to have smaller classes (average around 25) than other subjects where the average lecture size of introductory courses often exceeds 100. Economics and Sociology introductory lectures tend to be especially large. Interestingly, the average class size in most subjects is very similar by the selectivity of the university.

Student-Level Results

Table 4 provides estimates of Equation 2 using OLS. The outcome variable in these regressions is an indicator variable for whether the dropped out. We define a dropout as someone who was not enrolled the fall semester the year after their initial enrollment (i.e. freshman to sophomore year persistence). Because we observe transfers within the state of Ohio, we do not count these transfer students as dropouts. We attempt multiple specifications to account for possible non-linear relationships between class size and dropout rates. All models include controls for student race, gender, and ACT scores. We also control for the campus and year of entry.

In Column 1, we estimate OLS allowing the natural log of lecture size to enter the equation linearly.[2] In the OLS specification, the estimated effect is negative suggesting that larger classes reduce student dropout rates. This negative effect could arise from unobserved selection into large classes. For example, students who are more confident of their abilities may be more likely to take large classes. Column 2 allows lecture size to enter quadratically. The negative linear effect becomes larger but decreases in size as class size increases.

Starting in Column 3, the models include "portfolio" effects which account for the fact that different types of students take different combinations of classes and therefore face different potential class sizes. The inclusion of the portfolio effects reduces the negative relationship found between class size and dropout rates (specification 3). Columns 4 and 5 include additional measures of class size, and the main estimate does not change much.

While all of the OLS estimates are negative, they will be biased if there is unobserved heterogeneity across students who take similar courses. Table 5 estimates IV specifications where we treat class size regardless of its parameterization as endogenous. We instrument for class size by looking at variation in the proportion of assistant, associate, and full professors teaching in the given term relative to the departmental steady state. The first-stage regressions are reported in Appendix Table 1 and demonstrate a significant relationship between these deviations and class size.

The instrumental variables estimates provide a very different pattern of results. In Column 1, a linear specification, we find a positive but statistically insignificant relationship between class size and the dropout rate. When including a quadratic, the estimate becomes large

[2] In all specifications, we control separately for the size of labs and other types of non-lecture sections.

and negative. However, once controlling for differences in the portfolio of courses students had in their first semester, the results change sign. As shown in Column 3, students with larger average lecture sizes are found to be more likely to dropout; this negative effect declines with size as denoted by the quadratic.

One concern using the average class size as the chief explanatory variable is that it may mask some important heterogeneity. For example, suppose two students have an average class size of 100. One could have five classes with 100 per class while the other could have one class of 400 and four classes of 25. The student with 100 per class may have a very different experience with their class size. To control for this heterogeneity, we attempt some alternative specifications to allow for nonlinear effects of class size. For example, in Column 4 of Table 5, we allow the standard deviation of lecture size to enter the equation. Intuitively, if two students have similar average class sizes, the one with the larger standard deviation likely has more variation and likely a portfolio with some smaller classes than the other. When we use this specification, we fail to find any effects of the standard deviation (similar to the OLS results). Most of this is likely due to the inclusion of portfolio effects which may already control for some of this heterogeneity.[3]

In Column 5 of Table 5, we include the smallest class size that the student faced during the semester. This is meant to capture the possibility that a single intense interaction with a faculty may be sufficient to reduce a student's probability of dropping out. In this specification, we fail to find significant effects of the minimum class size. However, the linear effect of class size remains significant in this specification.

[3] When we run a similar specification without portfolio effects, we get a negative estimate on the standard deviation (i.e. larger standard deviation corresponds to less likelihood of dropping out).

In summary, once accounting for possible selection biases, the results suggest students in larger classes are more likely to dropout before their sophomore year. Comparing students with similar schedules and exploiting variation from short-run changes in the composition of a faculty, we find robust negative effects. The results are modest in size but statistically significant. We find no effect of the minimum class size nor do we find an effect on the standard deviation of class size on dropout rates.

In Table 6, we estimate a piecewise specification which allows the log of class size to have separate slopes for classes below certain cutoffs. The left side of the table uses 50 as the inflection point while the right side uses 100. A separate intercept for those classes over the cutoff is also included. In specifications 1 through 3, no special relationship is found below and above 50 students. However, students in courses above and below 100 may face different outcome likelihoods. Including portfolio effects, the IV estimates (specification 6) suggest that students in classes less than 100 do have an increasing likelihood of dropout as class size increases.

Student-by-Subject Level Results

The next part of the analysis focuses on student outcomes by subject. In this analysis, we examine the role of class size in a student's introductory course in a given subject to their subsequent involvement in that subject. Our sample size triples for this analysis reflecting that students in our data interact with, on average, three core academic subjects in their first semester.[4] In all of the estimates, we control for correlation across observations from the same student by clustering our standard errors at the student-level. Again, all models include controls

[4] We have restricted our analysis to core academic subjects excluding students' experiences in agricultural, physical education, and medical fields.

for student race, gender, ACT scores, campus, and year of entry. We also use faculty fixed effects in the specifications to control for selection of students into classes of different faculty members.

Table 7 displays the estimates for three measures of subsequent student interest. The first, as shown in Panel A, is whether the student took any additional courses in that department during later terms. In our linear specification including all departments (specification 1), we find that students in larger class were less likely to take additional courses in that subject. Similar effects are found when limiting the sample to only departments in the humanities (specification 3) or the sciences (specification 7), but only the coefficient for the sciences is statistically significant. However, when including a quadratic of log lecture size, our linear estimates become large and positive. Only the estimate for the humanities (specification 4) is statistically significant, and the quadratic suggests this effect declines as size increases.

The second panel repeats the analysis using total subsequent credits hours in the subject as the outcome. Again, larger lectures are associated with a lower likelihood of taking additional courses in the subject. The coefficient for the sample of science courses is also statistically significant. When a quadratic term is added, the sign of the main effect again reverses. The linear terms also become quite large, none are statistically significant.

The last panel examines the impact of class on major choice. While no effect is found in the sample of all subjects, class size is estimated to have a positive impact on major choice in the humanities and a negative effect amongst the sciences when using linear and quadratic terms (specifications 4 and 8).

In sum, we find that students in larger classes are less likely to show further interest in that subject. This appears to be especially true in the Sciences. However, the results show class size to have some positive effects in the Humanities.

V. Conclusions

In the academic literature focusing on class size and the production of education, researchers have focused almost exclusively on primary and secondary education. However, the results in this paper provide robust evidence that class size does matter in postsecondary settings as well. Students who take classes with large class sizes are more likely to dropout as a result of their experiences. Moreover, while we do not find that class size affects one's probability of majoring in a subject, students in larger classes tend to take fewer subsequent courses in that subject than other students who had smaller classes in the same subjects.

We estimated the results using an instrumental variables strategy that exploits variation in departments' faculties associated with discrete movements such as retirements, leaves of absence, and other changes. These changes in a faculty alter the proportion of courses being taught by assistant, associate, and full faculty members and as a result affect class size. The changes are likely uncorrelated with underlying subsequent student involvement and give us unbiased results.

The estimates suggest that large sections do impact student involvement and may suggest that an effective way to lower dropout rates may be to decrease class size. The results also provide some support for why the *U.S. News and World Report* includes class size amongst indicators of the quality of an academic institution. Interestingly, however, we fail to find that the specific metrics included in the *U.S. News and World Report* have significant explanatory power in our sample.

References

Angrist, Joshua D. and Victor Lavy. (1999) "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics*, 114 (2), pp. 533-75.

Becker, William E. and John R. Powers. (2001) "Student Performance, Attrition, and Class Size Given Missing Student Data." *Economics of Education Review*, 20 (4), pp. 377-88.

Dillon, Michael E. C. Kokkelenberg, and Sean M. Christy. (2003) "The Effects of Class Size on Student Achievement in Higher Education: Applying an Earnings Function." *Binghamton University Working Paper*.

Hoxby, Caroline M. (2000) "The Effects of Class Size on Student Achievement: New Evidence from Population Variation." *Quarterly Journal of Economics*, 115 (4), pp. 1239-85.

Kennedy, Peter E. and John J. Siegfried. (1995) "Does Pedagogy Vary with Class Size in Introductory Economics?" *American Economic Review*, 85 (2), pp. 347-51.

Kennedy, Peter E. and John J. Siegfried. (1997) "Class Size and Achievement in Introductory Economics: Evidence from the TUCE III Data." *Economics of Education Review*, 16 (4), pp. 385-394.

Krueger, Alan and Diane Whitmore. (2000) "The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR." *National Bureau of Economic Research Working Paper No. 7656*.

Mosteller, Frederick. (1995) “The Tennessee Study of Class Size in the Early School Grades.” *The Future of Children, Critical Issues for Children and Youths*, 5 (2), Summer/Fall.

Topp. Robert F. (1998) "Education of an Undergraduate: Action Before Reaction." *College and University,* 45(2), pp. 123-127.

Table 1: Descriptive Statistics

	Selective Four-year Colleges	Non-selective Four-year Colleges
Background Characteristics		
ACT Composite Scores (maximum score: 36)	22.93 (4.080)	20.89 (4.175)
White	.8667	.8084
Black	.0712	.1063
Male	.4459	.4648
Instate	.9987	.9979
Age	18.37 (.4997)	18.42 (.5587)
Course-Taking Behavior and Characteristics		
Lecture Size	4.140 (.9552)	4.098 (.9458)
Lab Section in Course	.0510 (.2201)	.0386 (.1926)
Course in Pre-college Major	.0412 (.1988)	.0379 (.1909)
Total Subsequent Hours Taken	6.161 (9.282)	5.041 (8.682)
Majored in Subject	.0647 (.2461)	.0551 (.2283)
Observations	30,319	9903

Table 2: Average Lecture Size of Courses Taken in First Semester by Institution

Institution	Mean	Median	Maximum
College A	87.19 (76.70)	53	287
College B	81.54 (97.24)	36	528
College C	50.27 (48.83)	34	258
College D	116.03 (142.35)	67	704
College E	69.05 (51.94)	47	231
College F	149.11 (156.25)	55	478
College G	223.80 (188.48)	223	454
College H	69.88 (66.20)	37	303
College I	95.58 (62.25)	74	199
College J	142.34 (103.55)	65	452
College K	59.83 (46.15)	41	296

Notes: Standard deviation reported in parentheses.

Table 3: Lecture Size of Courses in 1st Semester by Subject and Type of Institution

	Average Lecture Size	
	Selective	Non-selective
English	25.73 (20.35) [23]	25.06 (17.55) [24]
History	93.78 (53.53) [81]	156.99 (146.59) [359]
Philosophy	75.80 (53.50) [46.5]	72.36 (50.43) [44]
Economics	281.9 (167.80) [246]	253.13 (171.23) [72]
Political Science	95.67 (50.65) [81]	120.06 (107.00) [100]
Psychology	144.40 (137.01) [66]	148.50 (132.40) [195]
Sociology	300.86 (186.10) [285]	251.70 (180.92) [99]
Biology	195.76 (109.14) [180]	198.76 (114.26) [163]
Chemistry	143.99 (60.23) [134]	135.83 (59.11) [126]
Physics	173.09 (148.21) [157]	163.27 (144.00) [66]
Business	191.29 (197.07) [77]	139.88 (169.90) [39]

Notes: Standard deviation reported in parentheses. Median reported in brackets.

Table 4: OLS Estimates of the Impact of Class Size on Dropout Behavior

	Dependent Variable = Dropout Out Rate (mean= .138)				
	(1)	(2)	(3)	(4)	(5)
Log Lecture Size	-.0090*** (.0026)	-.0713** (.0179)	-.0490* (.0252)	-.0505* (.0252)	-.0488* (.0252)
(Log Lecture Size)^2		.0074** (.0021)	.0047 (.0030)	.0036 (.0031)	.0048 (.0030)
Lecture Size Standard Deviation				.0001 (.0001)	
Log of Minimum Lecture Size					-.0017 (.0057)
Course Portfolio F.E.	No	No	Yes	Yes	Yes
Observations	40,977	40,977	40,977	40,977	40,977

* Significant at the 10% level ** Significant at the 5% level *** Significant at the 1% level
Notes: In all specifications, we control separately for the size of labs and other types of non-lecture sections. Regressions also control for race, gender, age, and campus fixed effects.

Table 5: IV Estimates of the Impact of Class Size on Dropout Behavior

	Dependent Variable = Dropout Out Rate (mean= .138)				
	(1)	(2)	(3)	(4)	(5)
Log Lecture Size	0.0614 (0.0516)	-.3627* (.1991)	1.1323** (.5221)	1.127* (.5899)	1.1282** (.5641)
(Log Lecture Size)^2		.0433** (.0198)	-.1081** (.0483)	-.1077** (.0529)	-.1076 (.0551)
Lecture Size Standard Deviation				.00002 (.00116)	
Log of Minimum Lecture Size					-.0016 (.0795)
Course Portfolio F.E.	No	No	Yes	Yes	Yes
Observations	40,977	40,977	40,977	40,977	40,977

* Significant at the 10% level ** Significant at the 5% level *** Significant at the 1% level
Notes: In all specifications, we control separately for the size of labs and other types of non-lecture sections. Regressions also control for race, gender, age, and campus fixed effects.

Table 6: Student-Level Estimates using a Piecewise Linear Specification

	Dependent Variable = Dropout Out Rate (mean= .138)					
	LECTURE SIZE <50			LECTURE SIZE <100		
	OLS	OLS	IV	OLS	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Slope Log Lecture Size	-0.0112	-0.0151	1.0270	-0.0174**	-0.0165**	0.7078*
(lecture size < X)	(0.0078)	(0.0101)	(1.2632)	(0.0041)	(0.0059)	(0.3907)
Slope Log Lecture Size	0.0049	-0.0001	0.2372	0.0215**	0.0040	0.0006
(lecture size > X)	(0.0044)	(0.0066)	(0.5116)	(0.0078)	(0.0121)	(0.0877)
I (lecture size > X)	-0.0797**	-0.0676	4.3429	-0.1859**	-0.0926	3.1252*
	(0.0333)	(0.0453)	(5.3925)	(0.0429)	(0.0649)	(1.7312)
Course Portfolio F.E.	No	Yes	Yes	No	Yes	Yes
Observations	40,977	40,977	40,977	40,977	40,977	40,977

* Significant at the 10% level ** Significant at the 5% level *** Significant at the 1% level

Notes: Standard errors are clustered at the student level. Regressions control for race, gender, age, and campus fixed effects.

Table 7: IV Estimates of the Effect of Class Size on Student Outcomes by Subject

	All Subjects		Humanities		Social Sciences		Science	
A. Dependent Variable: Any Subsequent Credit Hours								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Lecture Size	-0.4027***	2.4521	-0.5998	5.8130*	0.0173	9.1549	-0.2945***	1.7842
	(0.1270)	(1.8682)	(0.4591)	(3.2512)	(0.2594)	(8.2275)	(0.1087)	(1.6571)
Log $(\text{Lecture Size})^2$		-0.3081		-0.6183**		-0.9072		-0.2319
		(0.2018)		(0.3092)		(0.7977)		(0.1844)
B. Dependent Variable: Total Subsequent Credit Hours								
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Log Lecture Size	-5.0534**	47.3288	-4.8480	99.2365	-5.7972	144.6276	-3.1112**	42.5702*
	(2.1085)	(32.8241)	(6.6207)	(65.6673)	(3.8389)	(117.9153)	(1.5508)	(24.1811)
Log $(\text{Lecture Size})^2$		-5.6528		-10.0348*		-14.9350		-5.0970*
		(3.5231)		(6.0673)		(11.4307)		(2.7308)
C. Dependent Variable: Choose to Major in the Subject								
	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
Log Lecture Size	-0.0168	-0.8450	0.1715	4.3949*	-0.1105	-1.6343	0.0386	-4.1014**
	(0.0552)	(0.8918)	(0.1274)	(2.3219)	(0.1072)	(2.7888)	(0.0584)	(1.0291)
Log $(\text{Lecture Size})^2$		0.0894		-0.4072*		0.1513		0.4621**
		(0.0958)		(0.2110)		(0.2737)		(0.1147)

* Significant at the 10% level ** Significant at the 5% level *** Significant at the 1% level

Notes: All specifications include separate controls for the size of labs and other types of non-lecture sections. Regressions also control for race, gender, age, ACT Score, and campus fixed effects.

Appendix Table 1: First Stage in IV Regressions at Student Level

	Dependent Variable = Log of Lecture Size	
	(1)	(2)
Assistant Professor	-.506***	-.559***
	(.055)	(.053)
Associate Professor	.323***	.413***
	(.056)	(.059)
Professor	-.126***	.702***
	(.036)	(.037)
Course Portfolio Effects	No	Yes
Observations	40,977	40,977

* Significant at the 10% level ** Significant at the 5% level *** Significant at the 1% level

Notes: Standard errors are clustered at the student level. Regressions also control for whether the course had an accompanying lab, the lab size, race, gender, age, and campus fixed effects.