Academic Achievement in Large Lectures: Analyzing the Effects of Attendance Rates and Class Motivation on Economics Exam Grades

Jill L. Caviglia-Harris*

ABSTRACT

Research on the impact of class size on student motivation and performance can be important for university instructors. This paper investigates the influence of class size on exam performance and attendance rates. Empirical analyses show that students in the larger section of microeconomic principles taught by the same instructor did not perform significantly worse on exams with the exception of the cumulative final. This is attributed to a significant drop in attendance at the end of the semester. Blinder-Oaxaca style decompositions of the residual effects reveal that the negative effects of large lectures can be greater for lower achieving students.

Introduction

The impact of class size on academic achievement has been the subject of education research for many years. However, research investigating the influence of class size on student achievement and motivation is not extensive in the economics or finance literature. Studies have found absenteeism to negatively impact grades in economics courses (Romer 1993, Devadoss and Foltz 1996, Chan et al. 1997, Marburger forthcoming, 2001), absences to increase with class size (Romer 1993, Devadoss and Foltz 1996) and that problems with motivation, incentives, and attention are more likely to occur in larger classes (McConnell and Sosin 1984). Together these results imply that students in larger classes may perform more poorly if the resulting reduced motivation leads to increased absenteeism. However, such issues have not been empirically tested.

Teaching is generally assumed to be a public good, however, as Bonesronning (2003) points out, there are also private good aspects. As class size is reduced, instructors have a greater chance to provide students with individual attention and can respond to the reduced class size by reallocating resources towards low-achieving students or by adopting teaching methodologies geared towards student needs (Brown and Saks 1987). The impact of class size on achievement can therefore be ambiguous, depending on the instructor's teaching method and student motivation.

Research on the impact of class size on student achievement has not focused on motivational issues or absenteeism but has shown that after accounting for student ability and other factors, class size does not impact student grades on multiple-choice exams or the ability to recall information in economic courses (Kennedy and Siegfried 1997, Mirus 1973, Lewis and Dahl 1972). However, Becker and Powers (2001) demonstrate that after accounting for missing student data due to student withdrawals, class size can impact student performance even on multiple-choice exams, while Becker and Johnson (1999) demonstrate that testing format can bias results. Studies have also shown that the higher level cognitive skills required to

* Jill L. Caviglia-Harris is an Assistant Professor of Economics, Salisbury University, Salisbury, MD, Email: jlcaviglia-harris@salisbury.edu. The author is grateful for suggestions by two anonymous reviewers.

answer essay questions can be impacted by class size (Raimondo et al. 1990, Siegfried and Fels 1979, Lewis and Dahl 1972, Crowley and Wilton 1974).

This paper seeks to investigate the influence of class size on exam performance and indirectly determine whether student motivation and attendance rates impact course grades with data that include performance on both multiple-choice and essay questions. This is achieved by first providing a description of the course, grading requirements, and data. The empirical analysis continues with the estimation of performance on exams, essay and multiple-choice components, and a breakdown of the cumulative multiple-choice questions found on the final, and concludes with an investigation of the indirect impact of student motivation and attendance rates on exam performance with Blinder-Oaxaca style decompositions of the residual effects. The paper concludes with a discussion of the results and suggestions for designing and teaching large lectures.

Institutional and Course Setting

In the Fall 2001 semester, the Economics Department at Salisbury University created a large lecture format for microeconomics principles to reduce the use of adjunct professors teaching lower level principles courses. One professor taught large (capped at 110 students) and small (capped at 30 students) sections in this semester. Class format included a mixture of traditional lecture (chalk-and-talk), games, discussion, and in-class exercises. Some games used by the instructor were modified for the large lecture size. These modifications were used in both sections to maintain consistency. The level of student participation was similar in both courses, and there were no attendance requirements. Two exams, a cumulative final and the highest seven of ten quizzes were averaged to evaluate student performance. Exams contained multiple-choice questions and essay questions (requiring students to provide graphs and/or numerical answers with explanations) that tested the same skills, topics and information for both classes. These exam questions were weighted 60 percent and 40 percent, respectively. The multiple-choice questions included on the final exam were developed by the author to evaluate student achievement in all principles courses and to assess the program through yearly evaluation and statistical analyses. They were designed to vary in difficulty and to represent material covered in all microeconomics courses taught in the department.[1]

The first and final exam contained identical essay and multiple choice questions for both groups while the second exam contained different essay questions designed to test the same skills. The second exam was made to appear different (by using different markets, examples, or numbers) to reduce the transfer of information and answers between students on the exam day. Since questions on the first and cumulative final exams were identical for the two groups, data on these exams are used in the empirical analysis.

Data Description

Previous papers have analyzed the impact of class size on student performance (Kennedy and Siegfried 1997, Becker and Powers 2001) using data collected for the Test of Understanding in College Economics (TUCE), sponsored by Saunders (1994) for the National Council on Economic Education. The data set used in these studies includes information on the type of institution, instructor, student reported characteristics as well as performance on 30 multiple-choice questions taken before and after micro- and macroeconomics principle courses.

This paper uses student level data collected for one common course at the same institution. Such an approach has positive and negative aspects. On the positive side, data include student characteristics composed from university records as well as performance on all course exams. Student reporting errors and the provision of falsified information are avoided by using university records. Although such data reduces response errors, sample size is significantly smaller relative to other studies, reducing variability and degrees of freedom in the estimations.

Although the TUCE is widely recognized as an adequate measure of economic knowledge and used in many studies of student achievement (Rothman and Scott 1973, Kennedy and Siegfried 1997, Saunders and

[1] The author designed the questions with feedback and input from other department faculty, making them similar in terms of rigor and content to the questions on exams administered previously in the semester.

Saunders 1999, Finegan and Siegfried 1999), several shortcomings have been noted in the literature. Swartz et al. (1980) note that the exclusive use of the TUCE to examine student ability provides a downward bias on estimates. By evaluating both the difficulty index and discrimination index of TUCE and department developed questions, they find their own questions to be better discriminators of ability and better predictors of student achievement. In addition, O'Neill (2001) finds that students that have been tested using essay questions throughout the semester do significantly worse on the TUCE than those that are tested using the multiple choice format. And finally, department chairs have found internally developed measurements of student achievement designed to fit existing curricula to be more useful when assessing economic programs and courses (McCoy et al. 1994). For these reasons, and others, the TUCE questions were not used to examine student performance in this study. Instead, the economics faculty at Salisbury University are in the process of developing a model for outcomes assessment, including guidelines for the development of a common set of questions specifically designed to test the performance of students directly related to the topics addressed in the courses taught in a particular program.

Exam grades and student characteristics for the sample can be found in Table 1. As evidenced, the two course sections do not differ significantly in terms of student major, gender, prior economics knowledge (measured by the number of economics course taken prior to microeconomics principles), college year, grade point average (GPA), or verbal score on the SAT (Satverb). However, there are more transfer students and significantly higher math scores on the SAT (Satmath) in the larger section. Also evident is that student scores on exams do not differ significantly with the exception of the cumulative final exam. The larger class did significantly worse.

Table 1. Data Definitions and Overview

	Small Section (Enrollment Cap of 30)			**Large Section (Enrollment Cap of 110)**			
	Mean	Standard Deviation	Number of Obs.	Mean	Standard Deviation	Number of Obs.	t-stat
Large class (= 1 if enrolled in large section)	0	NA	30	1	NA	112	NA
Major (=1 for majors that require micro and macro principles)	0.700	0.466	30	0.714	0.454	112	0.152
Year (1-4 for freshmen - seniors, respectively)	2.233	0.430	30	2.107	0.411	112	-1.48
Gender (= 1 for females)	0.433	0.504	30	0.339	0.476	112	-0.95
Econ prior (number of economics courses completed prior to taking microprinciples)	0.100	0.403	30	0.116	0.349	112	0.217
Transfer (= 1 for transfers)	0.067	0.254	30	0.196	0.399	112	1.689*
Satmath (SAT math score)	546.429	65.162	28	572.917	56.006	96	2.121**
Satverb (SAT verbal score)	544.286	59.094	28	537.292	57.920	96	-0.56
GPA (cumulative GPA Before taking the course)	2.826	0.757	30	2.861	0.539	112	0.286
Exam1 (grade out of 100)	73.137	11.659	30	70.252	14.420	112	-1.01
Exam2 (grade out of 100)	68.296	13.859	30	68.747	13.211	106	0.163
Final (grade out of 100)	63.025	10.813	28	57.149	12.781	104	-2.226**
Classave (grade out of 100)	76.736	8.175	28	74.307	10.787	104	-1.108

*, **, *** indicate significance at the 10, 5, and 1 percent levels, respectively

To avoid any censoring that Becker and Powers (2001) report can occur due to student withdrawals, all enrolled students are included in the analysis. However, some observations are dropped from the sample where SAT scores are not available, reducing the sample size from 142 to 124. This occurs most often for transfer students since some transfers are not required to take the SAT for admittance. A dummy variable is included in the analysis to account for whether a student transferred from another institution or not.

Empirical Analysis

Estimations on the impact of class size on achievement can be made using the standard reduced form production function (Raimondo et al. 1990, Bonesronning 2003):

$$A_{it} = \alpha_0 + \beta_1 A_{it-1} + \beta_2 I_i + \beta_3 C_t + e_{it}$$

where the dependent variable, A_{it} is achievement in the course (represented by the final course number grade) for student i at time t, and is dependent on academic achievement prior to the current semester (prior GPA), A_{it-1}, a vector of student characteristics, I_{it}, (including SAT scores, gender, prior economics knowledge, transfer status and major) class size, C_t and a random error term, e_{it}.

These estimations are first performed using ordinary least squares (OLS) estimation of the determinants of exam grades and overall class average.[2] In addition, a Heckman model is run to account for censoring of the data occurring when student self select themselves from the sample by withdrawing, or dropping, the course (Becker and Powers 2001). The multiple choice and essay sections of each exam are also estimated using OLS and Heckman models to determine if class size influences student answers on multiple choice and essay questions differently as is found in previous studies (Raimondo et al. 1990, Siegfried and Fels 1979, Lewis and Dahl 1972, Crowley and Wilton 1974). Next, multiple-choice questions on the cumulative final exam are divided between material covered in the beginning, middle, and end of the course to further investigate differences on exam performance between the two classes. Finally, regressions are performed to investigate the determinants of the multiple-choice questions on the final exam, since differences in student performance related to class size are found only for these questions. Estimations are made for students in the large class as well as with the pooled data (including dummy variables and interaction terms). These results are used to decompose residual effects on class size and further examine the source of the differences in student achievement on the final exam due to class size.

Estimation of Student Performance

OLS estimations are made for the first exam, final exam and class average (Table 2). The most influential determinants of student performance on exam grades include cumulative GPA (measured prior to taking the course) and the SAT math score. These results are consistent with previous studies that find success in economics courses to be linked to mathematical ability (Fourneir and Sass 2000). In addition, the number of economics courses taken prior to microeconomic principles has a positive and significant impact on all exam grades and the class average. Verbal ability (indicating by SAT verbal score) also plays a role in increasing performance on the final exam and the class average, but on not the first exam.

The impact of class size on student grades is insignificant for the first exam and the course overall, but not the final exam. Students in the larger section did significantly worse on the final after accounting for differences in major, GPA, verbal and mathematic ability, and other factors. It is possible that the lower final exam grade may due to an indirect effect of class size on student motivation to attend class. Along similar lines, Marburger (forthcoming) suggests that an attendance policy has significant impacts on student attendance rates and performance on exams only later in the semester. In a comparison of attendance rates between a class with an attendance policy and one without, Marburger finds that student grades and attendance rates are statistically similar for students in early in the semester, but the differences gradually increase as the semester progresses.

In this study, it was found that overall attendance decreased for the large lecture, but not for the smaller section over the course of the semester, even though the smaller section was offered during “off-peak” evening hours and the large section was not.[3] For example, in the last weeks of class the attendance rate averaged 70 to 80 percent for the large section and between 80 to 100 percent in the small section

[2] All regression models are tested for multicollinearity, heteroskedasticity, and coefficient robustness.

[3] Using data from four universities and 400 students, Devadoss and Foltz (1996) find that students attending classes held during “prime” hours (i.e. between 10 a.m. and 3 p.m.) were less likely to miss class.

(compared to weekly averages of higher than 80 percent recorded earlier in the semester for both classes). And, on quiz days, at the end of the semester, 5 to 10 percent of students left after completing the quiz in the large section compared to none in the small section. These differences in behavior cannot be explained by students self selecting the large or small sections because, as Table 1 shows, students in these two sections did not differ in terms of GPA, prior economics knowledge, or SAT scores prior to taking the course. (Actually, students in the large section had significantly higher math SAT scores, which according the estimations increase, rather than decrease, exam performance). Rather, these differences may be attributed to class size. Students are more anonymous and may not think professors will notice or care about absenteeism in large sections, not recognizing that absenteeism can impact exam grades (Marburger 2001). And, students are more likely to have problems with motivation to attend, participate, and pay attention in larger classes (McConnell and Sosin 1984). These motivation and attention issues, and the tendency to leave class early, are in part attributed to the group mentality that can be fostered in large lectures.

Table 2. Estimation of Student Performance on Exams

	Exam 1 (n=124)	**Class Average** (n=118)	**Final Exam** (n=118)	**Final Exam With Selection Bias Correction** (n=118)
	11.535	12.296	-8.834	-7.784
Constant	(14.068)	(9.832)	(13.322)	(13.395)
	-3.616	-2.349	-5.951***	-5.452**
Large Class	(2.438)	(1.727)	(2.340)	(2.434)
	-1.553	0.199	1.075	0.940
Major	(2.381)	(1.649)	(2.234)	(2.194)
	-2.623	0.477	-1.633	-1.688
Year	(3.048)	(2.475)	(3.353)	(3.356)
	3.078	-0.871	-1.946	-2.734
Gender	(2.154)	(1.513)	(2.050)	(2.066)
	6.427**	4.889***	4.790*	4.165**
Econ Prior	(2.717)	(1.880)	(2.548)	(2.592)
	6.283	3.684	5.237	3.995
Transfer	(4.014)	(2.798)	(3.792)	(3.808)
	0.035*	0.036***	0.041**	0.040**
Satmath	(0.019)	(0.013)	(0.018)	(0.018)
	0.027	0.018	0.034**	0.044***
Satverb	(0.018)	(0.012)	(0.017)	(0.017)
	11.204***	11.210***	11.351***	9.623***
GPA	(2.214)	(1.571)	(2.128)	(2.221)
				-14.969***
Lambda				(5.530)
R-squared	0.37	0.50	0.40	0.44
Adj. R-squared	0.32	0.46	0.35	0.40

Notes: standard errors in parenthesis; *, **, *** indicate significance at the 10, 5, and 1 percent levels, respectively

Performance on the final exam is also estimated with a two-stage model including all students enrolled in the course to correct for any bias resulting from censored data (Table 2). In the first stage, the probability that a student will remain in the course is estimated with a probit model.[4] In the second stage of the model,

[4] The only significant identifier is performance on the first exam. Also tested in the estimation are GPA, class size and major. It was expected that poorer students (as indicated by cumulative GPA) or students not enrolled in a major requiring microeconomic principles would be more likely to drop the course. However, the results do not indicate that any particular student attribute can

final exam score is estimated using the same variables as in the OLS regression.[5] In comparison, the results of the two estimations of the final exam score are similar: the same variables are significant and the coefficients have the same signs. Becker and Powers (2001) find the impact of class size is underestimated, because poorer students are more likely to withdraw from larger courses. In this case, the impact of the large class is found to be smaller in the corrected model. However, in both sections of the course, the same percentage of students (seven percent) withdrew, making the bias of these estimates not significantly different from zero.[6]

Multiple Choice and Essay Exam Estimations

Separate regressions are run on the multiple-choice and essay sections for the first and final exam administered in class to determine if class size influenced grades in the different answer formats (Table 3). As expected, similar variables are significantly related to exam grades compared to the estimation of total exam grades and therefore will not be discussed in detail. Interestingly, class size is insignificantly related to all test formats, with the exception of the multiple-choice section of the final exam. As suggested earlier, this result may also occur due to the indirect impact of greater levels of absenteeism in the larger class. And further, these results indicate that the negative impact of class size on this exam is evidenced by the multiple-choice section rather than the essay part (or both). This underlines the importance of exam design over question format. As Walstad (2001) reports, higher cognitive skills can be tested independent of test format, depending on question design. And, in this case it is evident from the OLS estimations that students in the large section found the multiple-choice section of the final exam to be relatively more challenging. Again, the Heckman correction models are run and the results of the two estimations of the final exam score are similar. The cumulative multiple-choice questions are divided between material covered in the first, second and third sections of the course (each section coinciding with an exam) to further investigate these results.

Table 3. Estimation of Student Performance on Multiple Choice and Essay Exam Formats

	Exam 1 (n=124)		**Final Exam** (n=118)		**Final Exam With Selection Bias Correction** (n=118)	
	Multiple Choice	**Essay**	**Multiple Choice**	**Essay**	**Multiple Choice**	**Essay**
Constant	6.781** (3.108)	-10.537* (5.984)	2.655 (4.172)	-12.919** (6.259)	2.998 (4.228)	-12.714 (5.980)
Large Class	-0.680 (0.539)	-1.032 (1.037)	-1.685** (0.733)	-1.706 (1.100)	-1.522** (0.769)	-1.608 (1.063)
Major	0.057 (0.526)	-1.490 (1.013)	0.985 (0.700)	-1.071 (1.050)	0.941 (0.691)	-1.097 (0.998)
Year	-0.308 (0.673)	-1.307 (1.297)	-0.254 (1.050)	-0.883 (1.575)	-0.272 (1.057)	-0.894 (1.509)
Gender	0.873* (0.476)	-0.003 (0.916)	-0.758 (0.642)	-0.115 (0.963)	-1.015 (0.651)	-0.269 (0.931)
Econ Prior	1.155* (0.600)	1.998* (1.156)	1.964** (0.798)	0.157 (1.197)	1.760** (0.817)	0.035 (1.157)

predict withdrawals accurately with the exception of performance on the first exam. All of the students that withdrew failed the first exam. A reduced model including the first exam score is used in the Heckman estimations.

[5] The two-equation procedure involves the estimation of a probit model of the adoption decision, calculation of the sample selection control function and incorporation of that control function (the inverse Mills ratio or lambda, λ) into the model of effort that is estimated with ordinary least squares (OLS). The inverse Mills ratio, sometimes referred to as the hazard rate, is based on the probability density function of the censored error term, and is used to normalize the mean of the error terms to zero. Consistent estimators are then calculated for α and β (Maddala 1983).

[6] The estimation precision of the Heckman, and other selection models, falls as sample size declines (Zuehlke and Zeman 1991). Therefore one must weigh the degree of censoring, or bias that results from the OLS estimations, versus the imprecision of the Heckman model to determine the preferred estimation technique under these circumstances.

	1.636*	0.432	0.400	3.662***	-0.005	3.420**
Transfer	(0.887)	(1.708)	(1.187)	(1.782)	(1.200)	(1.717)
	0.004	0.018**	0.009	0.016*	0.009*	0.016**
Satmath	(0.004)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)
	0.003	0.013**	0.010*	0.010	0.013***	0.012
Satverb	(0.004)	(0.008)	(0.005)	(0.008)	(0.005)	(0.008)
	1.901***	3.821***	3.096***	3.440***	2.532***	3.102***
GPA	(0.489)	(0.942)	(0.666)	(1.000)	(0.700)	(1.001)
					-4.886***	-2.924
Lambda					(1.771)	(2.432)
R-squared						
Adj. R-	0.27	0.30	0.36	0.23	0.40	0.24
squared	0.21	0.24	0.30	0.16	0.34	0.17

Notes: standard errors in parenthesis; *, **, *** indicate significance at the 10, 5, and 1 percent levels, respectively

Estimation of Final Exam Course Component Questions

The thirty multiple-choice questions included on the final were divided between material covered on the first, second and last exam to investigate the impact that falling attendance made on specific exam questions. (Nine questions covered material from the first third of the course, 11 from the second, and 10 from the last). Estimations of performance on these groupings of questions are made with OLS and Heckman selection models (Table 4). According to both OLS and Heckman selection models, the large class did significantly worse only on the questions related to the material covered in the first third of the course. (Other significant determinants are similar to previous estimations and are not discussed here as earlier estimations have greater explanatory power).

Explanations of these results are speculative, but presented here to suggest avenues for future research, and to motivate the decomposition analysis in the following section. One explanation is that it is possible that lower attendance rates experienced at the end of the semester impacted the retention of material covered earlier in the semester, especially since lectures held at the end of the semester included reviews of the final exam material. Attendance on the day of the formal in-class review was only 74 percent for the large section compared to 86 percent in the small section. It is also possible that some students learned a significant amount from the text and performed relatively well on previous exams (even though exams contained some material only discussed in lecture) (see Browne and Hoag 1995), expected similar results on the final. However, these students may not have anticipated the greater task of studying for a cumulative final. It is also possible that omitted variables related to student motivation, study habits, or attendance rates per student can better explain these differences. And a final possible explanation can be suggested based on the continuing study of students taking microeconomic principles course over the last four years (eight semesters) at Salisbury University (Caviglia-Harris and Kincaid 2004). In this study, we have found questions related to the material tested on the first exam to be the best discriminators of ability and predictors of student achievement. Specifically, questions about differences between demand and quantity demanded, supply and quantity supplied, demand and supply interactions and applications of these topics are some of the most difficult questions for the relatively poorer students to answer correctly. These are therefore some of the most important skills for student to focus on when studying for a cumulative final, and some of relatively important topics covered during the in-class final review.

Table 4. Estimation of Student Performance on Final Exam Multiple-Choice Questions

	OLS Estimations (n=118)			Heckman Selection Estimations (n=118)		
	Questions From First Third of Course	**Questions From Second Third of Course**	**Questions From Last Third of Course**	**Questions From First Third of Course**	**Questions From Second Third of Course**	**Questions From Last Third of Course**
Constant	2.803 (2.348)	-0.008 (2.910)	-1.093 (2.544)	3.270** (1.644)	0.430 (2.032)	-0.702 (1.987)
Large Class	-0.956** (0.407)	-0.787 (0.504)	-0.609 (0.441)	-0.694** (0.298)	-0.458 (0.367)	-0.371 (0.356)
Major	-0.361 (0.398)	-0.136 (0.493)	0.233 (0.431)	0.024 (0.270)	0.333 (0.336)	0.583* (0.331)
Year	-1.013** (0.509)	-0.953 (0.630)	-0.673 (0.551)	-0.265 (0.413)	-0.005 (0.514)	-0.002 (0.503)
Gender	0.138 (0.359)	0.005 (0.445)	-0.456 (0.389)	-0.088 (0.254)	-0.260 (0.316)	-0.668** (0.310)
Econ Prior	0.758* (0.454)	0.820 (0.562)	1.300*** (0.491)	0.394 (0.319)	0.382 (0.397)	0.984*** (0.387)
Transfer	1.009 (0.670)	1.284 (0.830)	-0.052 (0.726)	0.277 (0.468)	0.412 (0.583)	-0.694 (0.571)
Satmath	0.003 (0.003)	0.006 (0.004)	0.002 (0.003)	0.002 (0.002)	0.005* (0.003)	0.002 (0.003)
Satverb	0.002 (0.003)	0.005 (0.004)	0.005 (0.003)	0.002 (0.002)	0.005** (0.003)	0.005** (0.003)
GPA	1.310*** (0.370)	1.485*** (0.458)	1.562*** (0.400)	0.699*** (0.273)	0.775** (0.340)	1.057*** 0.333)
Lambda				-1.736*** (0.663)	-1.835*** (0.773)	-1.315* (0.766)
R-squared	0.25	0.25	0.28	0.25	0.25	0.28
Adj. R-squared	0.19	0.18	0.22	0.18	0.18	0.21

Notes: standard errors in parenthesis; *, **, *** indicate significance at the 10, 5, and 1 percent levels, respectively

Decomposition of the Residual Effects of Class Size

To further examine the possible impacts of class size on performance on the final exam, Blinder-Oaxaca style decompositions (Blinder 1973, Oaxaca 1973) of the residual effects are performed based on the framework presented in Jackson and Lindley (1989). Essentially, the impact of class size is divided between the endowment and residual effects, and the residual effects are decomposed into the constant and coefficient effects. This method allows for the partial isolation of the sources of disparity with a joint testing of the significance of the two components of the residual effects, and a more complete and accurate interpretation of group differences (Jackson and Lindley 1989). The endowment effect measures differences in exogenous variables such as intelligence. If this value is negative and large, this implies that differences in exam performance by students in the larger class can be attributed to lower initial endowments of those variables impacting exam grades. The constant effect is that portion of the total difference between group means that cannot be attributed to the endowment effect or those differential responses due to different initial characteristics. We would expect the constant effect to be negative and significant if there is a clear impact of class size on final exam performance. The coefficient effect measures differences between group responses in the dependent variable due to changes in the independent variables. If the coefficient effect is negative, this further supports the supposition that students in the smaller class perform relatively better on exams due to class size effects or different individual choices resulting from class size.

Table 5. Estimation of Student Performance on Final Exam Multiple-Choice Questions and Calculation of the Decomposition of the Residual Effects

	Large Class (n=91)	**Large and Small Classes Pooled** no class size dummy (n=118)	**Large and Small Classes Pooled** with class size dummy (n=118)	**Large and Small Classes Pooled** with class size dummy and interaction terms (n=118)
	0.943	1.274	2.655	0.447
Constant	(4.588)	(4.209)	(4.172)	(9.394)
	0.973	1.138	0.985	0.704
Major	(0.766)	(0.710)	(0.700)	(1.914)
	-0.265	0.270	-0.254	-0.694
Year	(1.304)	(1.045)	(1.050)	(1.996)
	-1.387*	-0.722	-0.758	0.322
Gender	(0.730)	(0.654)	(0.642)	(1.453)
	2.083**	1.923**	1.964**	0.707
Econ Prior	(0.926)	(0.813)	(0.798)	(1.805)
	1.583	-0.019	0.400	-5.263
Transfer	(1.270)	(1.196)	(1.187)	(3.416)
	0.002	0.006	0.009	0.021**
Satmath	(0.007)	(0.006)	(0.006)	(0.011)
	0.010*	0.011**	0.010*	0.018
Satverb	(0.006)	(0.005)	(0.005)	(0.012)
	4.476***	3.147***	3.096***	0.360
GPA	(0.782)	(0.679)	(0.666)	(1.312)
				0.269
BUAD*LARGE				(2.059)
				0.428
Year*LARGE				(2.377)
				-1.710
Gender*LARGE				(1.623)
Econ				1.376
Prior*LARGE				(2.025)
				6.846*
TRANS*LARGE				(3.640)
				-0.019
Satmath*LARGE				(0.012)
				-0.008
Satverb*LARGE				(0.014)
				4.116***
GPA*LARGE				(1.523)
			-1.685	0.497
LARGE			(0.733)	10.436)
Residual Sum of Squares	775.0094	1129.257	1076.546	927.6489
R-squared	0.41	0.32	0.36	0.44

Constant effect = 0.497 (coefficient of interactive model)

Coefficient effect = -2.19 (-1.693-0.497; Residual-Constant)

Mean of the Dependent (large class) = 20.780

Mean of the Dependent for large class given no class size effects= 20.603

Mean of the Dependent (small class) = 22.296

Total Effect = -1.516

Endowment Effect = 0.177

Residual Effect = -1.693

F* for residual effect = 2.4148** F* for coefficient effect = 2.0063**

Notes: standard errors in parenthesis; *, **, *** indicate significance at the 10, 5, and 1 percent levels, respectively

In the analysis of the endowment and residual effects of class size for this sample, results reveal that the endowment effect is 0.177, and the constant and coefficient effects are 0.497 and -2.189, respectively (Table 5). The endowment effect is relatively small and positive, implying that students in the large class were actually likely to get 0.177 more correct answers due to relatively greater initial endowments (as measured by prior GPA, the number of economics courses taken prior to microeconomics principles, and SAT scores). Both the residual and coefficient effects are found to be significant at the 5% level; however the constant effect is not significant. These results suggest that final exam score differentials cannot be attributed to differences in the intercept. However, the coefficient effect is found to be significant and negative, implying that differentials in class performance may be attributed to diverse choices made by students in the large lecture (such as attending less often) that resulted in lower grade or by bias created by the larger lecture format.

The class size interaction terms included in the last regression presented in Table 5 can shed light on the source of the differences, and help to differentiate whether they are student or class format driven. The interaction term that is most interesting in the analysis is the interaction between GPA and class size. The coefficient is positive and highly significant, indicating that students with higher GPAs in the large class are more likely to answer the multiple choice exam questions correctly. In other words, students with lower GPAs are impacted by the large class size to a larger extent. This suggests that the effect on exam performance results from student choices, and that the large lecture format plays a role in these decisions, most likely by providing the perceived anonymity that may reduce incentive to attend class on a regular basis and possibly by providing additional distractions that may make lecture less valuable. The decomposition analysis allows the differentiation of these exogenous and choice effects, essentially isolating the negative effect of class size to a subpopulation of the class (the relatively poorer students). A similar impact was not found for the smaller class, suggesting that the needs of individual students may be better addressed in this format by allocating greater resources to lower achieving students or providing an environment more conducive to student participation.

Discussion and Conclusions

Using the data collected from students enrolled in a large and small section of microeconomic principles, the empirical analysis performed in this paper shows that after accounting for GPA, SAT scores, and other factors, students in the smaller section preformed significantly better on the final exam but no different than students in a larger lecture on the first exam or the class overall. Rather, the major determinants of course grade are found to be cumulative GPA, prior economics knowledge, and the SAT math score, all of which are exogenous to course design, and suggest that instructors have a limited role in addressing student achievement through direct departmental policy or different course design measures. These results are somewhat different from previous studies that have shown class size to be insignificant. In this case, like Marburger (2001), it is believed that student attendance, found to be relatively lower in the larger class at the end of the semester, impacted grades on the final exam. However, the lower performance on the final exam was not great enough to impact class average overall. Students were much more likely to miss class and leave shortly after taking quizzes (administered during the first 15 minutes of class) in the large lecture, especially at the end of the semester, compared to the smaller section taught by the same professor. And, further analysis finds that students in the large lecture did significantly worse on questions related to material covered in the first third of the course, suggesting that attendance may impact student retention.

Although attendance rates were significantly lower for the larger of the two courses reviewed in this paper, they were greater than (or similar to) what has been found in the economics literature. For example, Romer (1993) finds an average of 33 percent absenteeism in economics courses, while Marberger (2001) finds an average of 19 percent and Sheets et. al (1995) calculate an average of 24 percent. The relatively high rate of absenteeism in the large class most likely cannot be attributed to the instructor as there is a history of relatively high levels of attendance, some of which can be attributed to positive evaluations by students (Sheets et al. 1995, Devadoss and Foltz 1996). Therefore the results suggest that the atmosphere created in the large classroom created motivation problems, at least with class attendance, and this may have impacted scores on the final exam. Furthermore, Blinder-Oaxaca style decompositions of the residual effects reveal that the negative effects of large lectures can be greater for lower achieving students.

This inevitably leads to the discussion of attendance requirements for large courses (see Marburger forthcoming). Since different behavior patterns were found for students in the large class, polices that encourage student participation and attendance, and reduce disruptions (such as walking out after a quiz) should be considered in course design. It is important that instructors are aware of these differences in behavior and develop course policy to reduce problems arising from class structure. Attendance policies are one way to address such issues. However, there is debate over whether such policy is warranted in an adult setting (Browne and Hoag 1995, Chan et al 1997). On the other hand, instructors may be able to address these issues by increasing the number of quizzes administered, including class projects, and developing critical thinking activities (Devadoss and Foltz 1996).

Although the sample used to investigate class size is small relative to previous studies that have utilized data collected from multiple universities, the nature of the data is improved in that individual characteristics of students and scores on different test formats are available. Even so, the results of the study should be taken as preliminary. Economic and finance departments cannot determine optimal class size by accounting for GPA, prior economics knowledge, or the SAT math score. However, this study suggests that lecturers of large classes carefully consider attendance policies and motivational issues when designing courses.

References

Becker, William E. and John R. Powers. 2001. "Student Performance, Attrition, and Class Size Given Missing Student Data," *Economics of Education Review* 20: 377-388.

Becker, William E and Carol Johnson. 1999. "The Relationship Between Multiple Choice and Essay Response Questions in Assessing Economics Understanding," *The Economic Record* 75 (231): 348-357.

Blinder, A. S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources* 8: 436-455.

Bonesronning, Hans. 2003. "Class Size Effect on Student Achievement in Norway: Patterns and Explanations," *Southern Economic Journal* 69: 952-965.

Brown, Byron W. and Daniel H. Saks. 1987. "The Microeconomics of the Allocation of Teachers' Time and Student Learning" *Economics of Education Review* 6: 391-332.

Browne, M. Neil and John H. Hoag. 1995. "Can We (in Good Conscience) Require Attendance in Our Classroom?" *The Journal of Economics* 21: 116-119.

Caviglia-Harris, Jill and Joel Kincaid. 2004. "Developing Outcomes Assessment Tools to Evaluate and Improve Economics Programs" working paper, Salisbury University.

Chan, Kam C., Connie Shum, and David J. Wright. 1997. "Class Attendance and Student Performance in Principles of Finance," *Financial Practice and Education* Fall/Winter: 58-65.

Crowley, R. W. and D. A. Wilton. 1974 "A Preliminary Report on the Efficiency of Introductory Economics Courses," *Journal of Economic Education* 5 Fall: 103-111.

Devadoss, Stephen and John Foltz. 1996 "Evaluation of Factors Influencing Student Class Attendance and Performance," *American Journal of Agricultural Economics* 78: 499-507.

Finegan, T. Aldrich and John J. Siegfried. 1999. "Do Introductory Economics Students Learn More if Their Instructor Has a Ph.D.?" *The American Economist* 42 (2): 34-46.

Fournier, G. M. and T. R. Sass. 2000 "Take My Course, *Please*: The Effects of the Principles Experience on Student Curriculum Choice. *Journal of Economic Education* Fall: 323-39.

Jackson, John D. and James T. Lindley. 1989 "Measuring the Extent of Wage Discrimination: A Statistical Test and a Caveat," *Applied Economics* 21: 515-540.

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

Lewis, D. and T. Dahl. 1972. "Critical Thinking Skills in the Principles Course: An Experiment," in *Research Papers in Economic Education*, A Welsh (Editor) New York: Joint Council on Economic Education.

Maburger, Daniel R. forthcoming. "Does Mandatory Attendance Improve Student Performance?" *Journal of Economic Education.*

Marburger, Daniel R. 2001. "Absenteeism and Undergraduate Exam Performance," *Journal of Economic Education* Spring: 99-109.

McConnell and Sosin. 1984. "Some Determinants of Student Attitudes Toward Large Classes," *Journal of Economic Education* Summer: 181-190.

McCoy, James P., Don Chamberlain and Rob Seay. 1994. "The Status and Perceptions of University Outcomes Assessment in Economics," *Journal of Economic Education* Fall: 358-366.

Mirus, R. 1973. "Some Implications of Student Evaluations of Teachers," *Journal of Economic Education,* 5: 35-46.

Oaxaca, R. 1973. "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review* 14: 693-709.

O'Neill, Patrick B. 2001. "Essay Versus Multiple Choice Exams: An Experiment in the Principles of Macroeconomics Course," *The American Economist* 45 (1): 62-70.

Raimondo, Henry J, Louis Esposito, and Irving Gershenberg. 1990. "Introductory Class Size and Student Performance in Intermediate Theory Courses," *Journal of Economic Education* Fall: 369-381.

Rothman, Mitchell P and James H. Scott, Jr. 1973. "Political Opinions and the TUCE," *Journal of Economic Education* Spring: 116-123.

Sanders, P. 1994. The TUCE III Data Set: Background information and file codes (documentation, summary tables and five 3.5 inch double sided high density disks in ASCII format), National Council on Economic Education: New York.

Saunders, Kent T. and Phillip Saunders. 1999. "The Influence of Instructor Gender on Learning and Instructor Ratings," *Atlantic Economic Journal* 27 (4): 460-473.

Sheets, Doris F., Elizabeth E. Topping and John Hoftyzer. 1995. The Relationship of Student Evaluations of Faculty to Student performance on a Common Final Examinations in the Principles of Economics Courses," *The Journal of Economics* 21: 55-64.

Siegfried, J. and R. Fels. 1979. "Research on Teaching Economics: A Survey," *Journal of Economic Literature* 17: 923-969.

Swartz, Thomas R., Frank J. Bonello, and William I. Davisson. 1980. "The Misuse of the TUCE in Explaining Cognitive Achievement," *Journal of Economic Education* Winter: 23-33.

Walstad, William B. 2001. "Improving Assessment in University Economics," *Journal of Economic Education* (Spring) 281-293.

Zuehlke, Thomas W. and Allen R. Zeman. 1991. "A Comparison of Two-Stage Estimators of Censored Regression Models," *Review of Economics and Statistics* 73 (1) 185-88.