Order Number 9428027

The influence of class size on student achievement in principles of college economics: A production function approach

Cochran, Howard Henry, Jr., D.A.

Middle Tennessee State University, 1994

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

The Influence of Class Size on Student Achievement in Principles of College Economics: A Production Function Approach

Howard H. Cochran, Jr.

Doctoral Dissertation Submitted to the Faculty of
the College of Graduate Studies at Middle Tennessee State University
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Arts in Economics

May 1994

The Influence of Class Size on Student Achievement in Principles of College Economics: A Production Function Approach

Approved:

Graduate Committee:

Major Professor

Reuben Kyle

Committee Member

Bob Womack

Committee Member

John T. Lee

Chair of the Department of Economics and Finance

Mary Martin

Dean of the College of Graduate Studies

Abstract

The Influence of Class Size on Student Achievement in Principles of College Economics: A Production Function Approach

Howard H. Cochran, Jr., D.A.
College of Graduate Studies at Middle Tennessee State University
May 1994

The primary purpose of this study is to investigate the influence of class size on student achievement in principles of college economics. A sample is selected from the Test for Understanding of College Economics III (TUCE III) database. The TUCE III exams are used to measure student achievement. An educational production function is used to assess the influence of several independent educational inputs on the achievement outcome. Educational inputs are categorized into faculty capital, student capital, course organization and environmental factors.

Several methods are used to model the achievement variable: absolute achievement; absolute improvement; percentage improvement; and gap closing. Class size influence is an independent variable which is modeled as either a direct or indirect influence on achievement. Direct influences include the actual number of students in the class, the natural log of class size and class size squared. Indirect class size influences include: the range of student ability; the standard deviation of ability; minimum and maximum ability scores in a class; average ability; and kurtosis and skewness of ability. Dummy variables for class size ranges also are used in order to reduce the problem of aggregation. Ordinary least squares and Poisson regressions are employed to describe associations between independent and dependent variables.

Several policy implications are suggested by the statistical evidence of this study. First, a class size of 30 students or less will raise mean achievement. Second, managing controllable educational inputs effectively can raise average achievement. Third, class size can be increased, costs reduced and mean achievement raised when educational inputs are optimized. Fourth, an empirical method for evaluating teaching effectiveness relative to national standards is presented. Fifth, some practical suggestions are made for using the results in pay for performance schemes for faculty teaching principles of college economics.

Acknowledgments

to God - to the creator of all things for His many blessings

to family - to my parents Howard H. Cochran, Sr. and Hattie A. Cochran for their unconditional love, support, encouragement, sacrifice and prayers
to Hugh and Ana Paula Cochran for their kindness and care
to Harold, Celeste, Benjamin, Nathaniel and Josiah Cochran for their selfless example
to Bob and Sally Cochran for being my second family

to friends - to Buddy Tate who was always there at just the right time
to Casey Baluss for a empathetic listening ear
to Bill Bandy for saving my health
to many encouragers - Nellie Campbell, Jack Ross, Kelly Hienrich, Alan Pace, Fred MacKrell, Steve Simpler, Bob Simmons, Steve Felts

to mentors - to Joachim Zietz for his patience, generous support and call to excellence
to Reuben Kyle for his precision
to Duane Graddy for his enthusiasm in applying economics to everything
to Jack Arters, a noble teacher who inspires

to colleagues - to Jenny Carter Thomas for her friendship and support
to Gary Hodgin for his personal lending library
to Dick Lutz for allowing me the flexibility to reach a goal
to Lorena Edwards who encourages me to never stop learning
to Harry Hollis for never letting me lose faith

to Belmont - Jim Cotham, Cliff Eubanks, Larry Wacholtz, Jerry Warren, Wayne Brown, James Stamper, Bill Troutt, Pat Finney, Jim Godfrey, Louis Geresy and many others

to students - to all of my students past and present who have taught me how to teach
to my international students who have helped me see the world anew
to all who have taken the time to care

to others - to Billy Balch for keeping graduation within reach
to Chris Baroni for computer programming literacy
to H. Theodore Hoffman who gave himself to economics education

Table of Contents

List of Tables

Table Page

List of Appendices

Chapter I

Introduction

> Of all the beliefs held by American educators, few are more durable than the proposition that smaller classes will yield higher student achievement. It is one of the great 'givens' of education in this country. It may also be a very expensive excuse for instructional failure.
>
> Chester E. Finn, Jr.
> Former Assistant Secretary of Education
> in Mitchell et al. (1989)

> [Class size studies] are mostly dated, and more needs to be learned about this important question since it is one of the few determinants of learning that is easy (but not cheap) to control.
>
> Siegfried and Walstad (1990)

> The effect of class size on learning: I want to emphasize how important this issue is...the question is: do larger classes mean an inferior product, and, if so, how much inferior? A solid, scientific answer to this question would be of intense interest to deans, provosts and university presidents all over the country.
>
> Blinder (1991)

Class size has been a lingering policy issue in higher education. Ask a college professor to recommend an optimal class size, and a strong professional opinion will most likely be evoked. The consensus may suggest that students will learn more in small classes because small classes can be taught more effectively. This enduring proposition is several hundred years old. The small class or tutorial system at Oxford and Cambridge was the standard to be emulated until systems of mass higher education, such as the one in the United States, were developed. Bigger seemed to be better since more students could be educated at a lower cost per pupil. However, in recent years declining student achievement as measured by lower standardized exam scores or grade inflation has begun to question the "bigger is better" assumption.

Small class sizes are assumed to improve the teaching-learning process. The Tennessee legislature through the state department of education committed 12 million dollars to fund the Student/Teacher Achievement Ratio (STAR) project in 1984. The funds went primarily toward the costs associated with class size reduction in kindergarten through third grade during the four years of the project. Class sizes were reduced by approximately one-third, from an average class size of 23 to 15, in 75 schools across the state. Although some educators are convinced that class size should be reduced, Folger and Breda (1989) found insufficient statistical evidence to support the proposition that a small class size will improve student achievement in second and in third grade.

Balancing the cost and benefits of class size reduction or class size increase has not been adequately discussed. How should gains or losses in achievement due to class size policy changes be valued? One can assume the notion that any gain in student achievement is worth the cost incurred is as inappropriate as the idea that drastic cost reductions will have no secondary effects on student achievement. In a world of scarcity, achievement gains at any cost or budget savings at any price are inappropriate. Colleges should make efficient choices, not all or nothing choices. Thus, universities appear to be caught in a dilemma when trying to achieve increased student achievement, while simultaneously lowering cost. Class size policies appear to be in conflict with this goal.

The implicit assumption is that student achievement is a function of class size alone. Small class size advocates try to persuade others that class size is one of the single most important determinants of student achievement. Chester Finn argues that this obsession could "...be a very expensive excuse for instructional failure." Identifying the determinants of achievement may help prevent instructional failure.

Section 1: Purpose, Uniqueness and Limitations of the Study.

The primary purpose of this study is to investigate the influence of class size on student achievement in principles of college economics, as the quotes by economists Siegfried and Walstad (1990) and Binder (1991) at the beginning of the chapter indicate. This study takes up the challenge of providing "[a] solid, scientific answer..." (Blinder 1991) to the class size question for principles of college economics. In particular, this study seeks to determine if class size has a significant negative impact on student achievement in principles of college economics. Only by meeting several additional objectives can a scientific solution be found. This study will:

- identify a measurement instrument for student achievement that is both reliable and commonly accepted
- enumerate the inputs affecting the educational process of an introductory economics course
- analyze different achievement measures
- discover the determinants of student achievement other than class size using an educational production function
- find out how the determinants of student achievement affect higher and lower ordered cognitive skills
- assess the impact of direct and indirect class size influences on student achievement

- calculate the costs associated with class size reductions and savings associated with class size increases
- suggest policies which optimize student achievement while simultaneously lowering cost
- address the evaluation of teaching effectiveness
- provide suggestions for further research

This study is unique for two reasons. First, as Siegfried and Walstad (1990) suggest at the beginning of the chapter, class size studies are mostly dated. This investigation is the most comprehensive principles of college economics class size inquiry to date. Second, the sample for the empirical analysis is taken from the Test for Understanding of College Economics III (TUCE III) database, which is made available by the National Council on Economic Education (NCEE). The database contains 9,679 observations from 62 universities with feedback from 189 separate courses in either principles of macroeconomics or principles of microeconomics. To the author's knowledge any economics education research on class size using TUCE III data does not exist.

This research is also distinguished from its predecessors by improving the inquiry in the following ways: a wide range of class sizes is considered, specifically classes with 10 through 232 students; the sample size is larger than most previous inquiries; the economics course content can be identified as either principles of macroeconomics or principles of microeconomics; the influence of recitation or discussion groups on large classes can be measured; a common, easily accessible and universally accepted definition of achievement is used, the TUCE III exams; the primary question is the influence of class size on student achievement; an input-output model is used to assess student achievement outcomes; class size influence is theoretically modeled and tested; and educational inputs are selected ex ante from input categories known to influence achievement.

This study is limited by the following factors:

1. The definition of achievement is the TUCE III 30 exams. The conclusions may not be applicable when different outcome measures are used.

2. The largest class size in the sample selected from the TUCE III database for empirical analysis is 132 students. The influence of class size on student achievement may not be representative for classes above 132 students.

3. The TUCE III database put together from the voluntary submission of TUCE III exam scores, faculty questionnaires and student questionnaires. Therefore, not all of the students enrolled in the 189 introductory economics courses are sampled. Less conscientious students may have submitted incomplete exams or questionnaires. This study will exclude incomplete records; as a result, the better students may be oversampled.

4. Class size is measured by proxy. The number of students in each of the 189 course section files serves as a proxy for enrollment. This proxy does not include adjustments for students who either chose to withdraw from the course nor does the proxy include adjustments for students who chose not to complete the exams or questionnaire. The number of students withdrawing or not choosing to participate is assumed to be small.

5. This study may have a problem with misspecification. Extensive information on personal characteristics and sociodemographic influences is not available in the TUCE III database.

Section 2: Organization of the Study.

This study is organized into five chapters, including the introduction. Chapter two includes a review of the literature investigating the link between class size and student achievement in principles of college economics. It discusses at length the use of production functions to assess educational outcomes, and it presents several theories of class size influence on student achievement. The theories are grouped into either direct or indirect class size influences, but the direct and indirect theories have not been tested in economics educational research.

Chapter three describes the preparation of the data, the selection of the sample used for empirical analysis and the research methodology. The TUCE III data are available in 189 separate files that must be combined into one database before the selection of the sample. This chapter details the process of combining the files into an aggregate database, and it sets forth the method by which the sample of 2,942 observations is selected from the aggregate database. Finally, this chapter shows the research method employed decomposed into six steps.

Chapter four presents statistical evidence on the influence of class size on student achievement in principles of college economics. In presenting the evidence, this chapter: identifies the determinants of student achievement; compares OLS and Poisson statistical models; chooses a preferred method for describing student achievement; assesses the impact of direct and indirect class size influences; and uses class size ranges in order to reduce aggregation influences.

Chapter five enumerates several policy implications suggested by the statistical evidence. The evidence suggests that student achievement can be increased when the use of education inputs are optimized, that gains in student achievement can be realized when small class size enrollments are established, and that increasing class size is more economically efficient than reducing class size. This chapter also presents an empirical method for evaluating teaching effectiveness relative to national standards is presented, and it offers some practical suggestions for using the results in pay for performance schemes. Lastly, it proposes directions for future research.

Chapter II

Review of Related Literature

The professional literature discussing the influence of class size on student achievement in principles of college economics can be reviewed on the basis of three unifying themes: (1) research specifically investigating the link between class size and achievement; (2) the typical production functions used to estimate educational outcomes; and (3) the modeling of the class size and achievement relationship. The discussion that follows elaborates on these three themes.

Section 1: Research Investigating the Link Between Class Size and Achievement.

The purpose of this section is to report on the general findings regarding the influence of class size on outcome measures other than achievement and to focus on the literature specifically addressing the class size-achievement link.

1.1. Class Size Influence on Outcome Measures other than Achievement.

Class size is related to a number of different outcome measures other than achievement for principles of college economics courses. Raimondo, Esposito and Gershenberg (1990) find that enrolling in a large principles of macroeconomics course has a negative influence on the grade earned in intermediate macroeconomics. Adams and Becker (1990) and Dossugi (1992) find that class size does not significantly affect the decision to withdraw from a course. McConnell and Sosin (1984) discover that larger classes negatively impact student attitudes. DeCanio (1986) uncovers a negative relationship between class size and the student evaluation of teaching performance while

Mirus (1973) finds only a slight positive relationship. Lewis and Dahl (1972) detect a negative impact of class size on critical thinking skills. Card and Krueger (1992) are able to link a larger class size with decreased earnings. While class size appears to influence several educational outcomes, the question arises whether class size influences achievement.

1.2. Class Size Influence on Achievement.

Only seven widely published and cited empirical investigations that attempt to address this question for principles of college economics exist. Chronologically, the investigations are: Levin (1966); Lewis and Dahl (1972); Crowley and Wilton (1974); Williams, et al.; (1985); Lopus (1990); Myatt and Waddell (1990); and Gramlich and Greenlee (1993).

1.2.1. The Mixed Evidence of Class Size Influence on Achievement.

The empirical evidence on the effect of class size on student achievement in principles of college economics is mixed. Achievement is commonly defined as a final course grade or a score on an objective standardized test of general economic knowledge. A number of studies find small classes to have a positive impact on achievement. Crowley and Wilton (1974) find statistical evidence that decreasing class size improves scores on a standardized exam when the exam is proctored by an instructor and when the exam is counted as part of the course grade. When Crowley and Wilton (1974) gave the standardized exam under non-testing conditions (i.e., the exam was not proctored by an instructor and no credit for the exam was counted toward the course grade), the sign of the coefficient reversed but proved to be statistically insignificant. Lopus (1990) finds an inverse relationship between class size and scores on a standardized exam. Levin (1967) identifies improved final exam grades with smaller classes. Other studies, however, fail to detect a positive association. Lewis and Dahl (1972) and Levin (1967) do not find class size to be an influential predictor of a standardized exam score. Lewis and Dahl (1972) identify a negative but insignificant coefficient for class size. Williams et al. (1985), Gramlich and Greenlee (1993), Myatt and Waddell (1990) and Levin (1967) do not present any evidence that final grades in small classes are statistically different from final grades in large classes. Gramlich and Greenlee (1993) report a positive sign for the small class coefficient in principles of microeconomics and a negative sign for principles of macroeconomics, but neither result is statistically significant. Levin (1967) also finds class size to be an insignificant influence on a gain score definition of achievement.

Casual empiricism produces mixed sentiments on the influence of class size as well. A number of economics educators believe small classes have a positive influence on achievement. McKeachie (1986), Siegfried, et al. (1991) and Siegfried and Walstad (1990) essentially believe that professional judgment along with the dynamics of smaller classes point toward an inverse relationship between class size and achievement. However, Hanushek (1979) argues that the only guaranteed outcome would be a rise in cost when small classes are used uniformly.

1.2.2. Class Size Investigations Lack Common Foundations for Comparison.

The mixed empirical evidence of previous class size investigations may be due to the lack of common foundations in the literature. Although the results of class size investigations are widely cited, the conclusions are less clear when one tries to find a basis to compare the results. A variety of points can be used to make a comparison, including the following: the range of class sizes being tested; the size and source of the sample; the content of the economics class; the influence of recitation sections on larger classes; the definition and measurement of achievement; the primary question of the investigation; the underlying theoretical framework for modeling student achievement; the modeling of class size influence; approaches for selecting independent variables; statistical methods; and statistical results. Since these points have varied from one study to the next, comparison is more difficult.

1.2.2.1. The Range of Class Sizes being Tested.

A wide range of class sizes is being tested in the literature. Levin (1967) uses experimental large classes of 80 through 150 students and small class control groups of approximately 35 students. The average section size for Lewis and Dahl (1972) is 37 students with a range of 21 through 52 students, while the mean class size is approximately 30 for Lopus (1990). Small class sizes of 15 or fewer students occurs in about 10 percent of the observations for Gramlich and Greenlee (1993) when the variable is strictly used as a dummy to correct for attrition in poorly taught classes. Although class size ranges from 13 through 1,006, Williams, et al. (1985) do not give descriptive statistics on the classes which are considered introductory economics. No data on class size range is provided by either Crowley and Wilton (1974) or Myatt and Waddell (1990).

1.2.2.2. Sample Size and Source.

The sampling source and size varies widely as well. Lopus (1990) is the only investigator to use a national database. A 2,440 observation subset of the National Assessment of Economic Education (NAEE) survey database is used to profile high school students from across the nation taking the Test for Economic Literacy (TEL). All other investigators use sample sources from only one college. Levin (1967), Lewis and Dahl (1972), Crowley and Wilton (1974), Williams, et al. (1985), Myatt and Waddell (1990) and Gramlich and Greenlee (1993), respectively, draw a sample from the following institutions with the corresponding number of observations in parenthesis: University College of Rutgers (950); University of Minnesota (654); Queens University of Canada

(127); Brigham Young University (16,230); a small Atlantic Coast university (928); University of Michigan (5,066).

1.2.2.3. Content of the Economics Course.

Comparing the class size studies on the basis of course content can be difficult because the content may vary from one study to the next. Levin (1967) describes the classes as General Economics 101 and 102, while Myatt and Waddell (1990) simply refer to freshman principles of economics and Crowley and Wilton (1974) to freshman economics. Williams, et al. (1985) use the broadest definition by calling the class economics and by including it along with 14 other non-specific course categories. Only Lewis and Dahl (1972) identify the course as being principles of macroeconomics, and Gramlich and Greenlee (1993) identify either introductory microeconomics or macroeconomics. Finally, Lopus (1990) includes a high school, not college, economics course.

1.2.2.4. Recitation Section Influence on Large Classes.

A potential concern arises when students in large classes are enrolled in or allowed to attend recitation sections. Separate recitation or discussion sections along with the availability of teaching assistants could bias findings of large class size influence on student achievement. Some large classes are allowed to meet in small discussion groups (Levin 1967). Other classes meet all sections once a week for a mass lecture by senior faculty and then meet separately three times a week with teaching assistants (Lewis and Dahl 1972; Gramlich and Greenlee 1972). Still others do not specify if small class influences could exist in large classes (Crowley and Wilton 1974; Williams, et al. 1985; Lopus 1990; Myatt and Waddell 1990). None of the investigators tries to account for this potential bias in their statistical procedures. However, Myatt and Waddell (1990) did include a dummy variable to indicate the year in which departmental policy changed from a multisection teaching approach to a mass lecture approach. The mass lecture approach is intended to save manpower by meeting two hours per week in large sections then meeting one hour per week in small sections of about twenty. Although the dummy variable is insignificant, the possible influence of small sections on the investigation's conclusion is not explored.

1.2.2.5. Definition and Measurement of Achievement.

Achievement is commonly defined and measured in one of two ways. First, achievement is defined as a common final exam score (Williams, et al. 1985) or the numerical conversion of a final grade (Myatt and Waddell 1990; Gramlich and Greenlee 1993). One investigator collects data for both final exams and grades (Levin 1960). None of the studies indicates the comprehensiveness of the final exam nor the weight of the final exam, class activities or topics taught in the calculation of the final grade. The content and relative weight of final exam questions is not revealed. Moreover, none mentions whether a standard letter grading system or one with plus/minus notation was used.

Second, achievement is defined as a score on an objective standardized test of general economic knowledge. One study uses a TEL score (Lopus 1990) and does not mention if the test was given under exam conditions or was incorporated as part of the course grade. Two studies use a post-TUCE I score definition of achievement. Lewis and Dahl (1972) do not mention if the test was given under exam conditions or was incorporated as part of the course grade. Crowley and Wilton (1974) allowed students to be tested under both exam and non-exam conditions. No results report on the influence this condition had on performance. Additionally, these investigators modified post-TUCE I to better match the Canadian experience. Finally, Levin (1967) constructed an objective pre- and post-test by selecting questions suggested by Whitney (1960). The post-test was embedded as part of the final exam and was only given at the end of the second semester of General Economics. Levin (1967) is the only investigator to use several definitions of achievement, one of which is a gain score and is computed as the difference between post- and pre-tests.

1.2.2.6. Investigations' Primary Question.

Only Levin (1967) and Williams, et al. (1985) incorporate, as a primary question, the influence of class size on achievement. The former uses experimental large classes with small class control groups and then tests for a statistical difference between means on various achievement measures among the two types of classes. The latter uses class size and class size squared as the primary independent regressors on a common final exam score. The remaining studies include class size as an incidental rather than primary explanatory research variable. Consequently, the research question does not primarily focus on the influence of class size on achievement.

If the other researchers do not set out to primarily investigate class size influence, what is the primary purpose of their research design? Crowley and Wilton (1974) offer the broadest design by testing a wide range of independent variables which are thought to influence student performance. Class size is only one of several independent variables. Lewis and Dahl (1972) seek to explore the relationship between critical thinking skills and achievement. The first model presented regresses a post-TUCE I score on several independent variables, one of which is section size. Lopus (1990) primarily focuses on the effect that additional educational expenditures have on achievement in high school economics. Class size is one of several independent variables in the regression equation. Myatt and Waddell (1990) principally investigate the influence that high school variables have on the percentage grade of a first year college principles of economics course. Again, class size is one of several independent variables. Finally, Gramlich and Greenlee

(1993) primarily explore the effect that student evaluations of teaching performance have on student grades. In short, class size is not the primary focus for much of the literature that has been cited, although the literature is considered the most authoritative on the class size question.

1.2.2.7. Theoretical Modeling of Student Achievement.

The accepted theoretical framework for modeling the influence of economics educational variables is a typical input-output or production function approach. None of the class size studies specifically identify this framework, but simply use the approach previous investigations implied. Except for Williams, et al. (1985) all inputs reviewed are of the first order and non-interactive. None of the studies attempt to model the influence of specific significant inputs. Myatt and Waddell (1990), however, do attempt to specify a behavioral equation for attitude. Since attitude is unobserved, the equation is rearranged so that certain variables can serve as proxies for attitude in the estimated equation. The predicted signs of the variable coefficients are then compared to the expectations of the theoretical model. This is the only technique in the reviewed literature that attempts to use an a priori falsifiable scientific approach in modeling.

1.2.2.8. Theoretical Modeling of Class Size Influence on Student Achievement.

Since most studies are not designed to primarily investigate the class size influence, class size is not typically modeled. Most investigators simply include the number of students in a class as one of several independent variables in a standard regression (Lewis and Dahl 1972; Crowley and Wilton 1974; Lopus 1990; Myatt and Waddell 1990). While Gramlich and Greenlee (1993) use class size as an independent variable, it is included only as a dummy variable to indicate if the number of students in a class is fifteen or less. This dummy variable serves as a potential indicator of attrition of less capable students from sections that are poorly taught by graduate assistants. The lack of a strong theoretical model may explain the reason for sign switching in Crowley and Wilton (1974) along with the negative, although statistically insignificant, class size influence in Lewis and Dahl (1972). Williams, et al. (1985) is the only exception since class size influence is modeled by squaring the variable. Although the rationale is not discussed, this might be an attempt to try to account for the increased opportunity for interaction occurring in larger classes.

1.2.2.9. Selecting Independent Variables.

Three typical approaches to select independent variables have been used. One approach is to include variables applied in other studies to account for achievement in a principles course (Crowley and Wilton, 1974). A second approach includes variables that seem to make the most sense to the investigator in addressing a particular research question (Myatt and Waddell 1990). Most inquires incorporate this method even when broad categories are defined. The third approach seeks to identify broadly defined categories into which the educational input variables may fit. Lewis and Dahl (1972) identify three categories: student characteristics; instructor characteristics; and environmental variables. Gramlich and Greenlee (1993) list specific variables that they consider necessary to their research inquiry and fit these variables into two broad categories: various measures of student grades in economics classes; and student control variables. Lopus (1990) identifies approximately five categories: aptitude; socioeconomic factors of students; peer group effects; teacher attitude; and education expenditures.

The theoretical justification for categorization, behavioral or otherwise, is not explicitly identified in these studies that incorporate class size in their empirical design. The categorization varies across investigations and can also be vague. This lack of methodological unity may raise a question as to the comparability and usefulness of the findings along with prompting a discussion as to the influence of unobserved or omitted variables.

1.2.2.10. Statistical Methods.

Once the descriptive variables used to explain achievement are determined, a statistical method must be selected to sort out the relative influence of each variable identified in the input-output framework. Discovering which variables are influential usually consists of running linear regressions on the independent variables in order to determine significance. Most investigators use the method of ordinary least squares (OLS) to estimate their linear regression equations (Lewis and Dahl 1972; Crowley and Wilton 1974; Lopus 1990; Myatt and Waddell 1990; Gramlich and Greenlee 1993). Two investigations apply a standard stepwise multiple regression procedure (Lewis and Dahl 1972; Williams, et al. 1985). One study uses the statistical technique of comparing differences between mean grades and exam scores for large and small classes (Levin 1967). All models use a linear functional form except for Williams, et al. (1985), who estimate a log-linear form for their second equation.

1.2.2.11. Statistical Results.

There are numerous statistical results that can be compared. The class size results which can be compared from one study to the next include: the independent variables used in regression and their definitions; the coefficient signs of independent variables; the statistical indicators of model adequacy; and the identification of potential biases.

Table C.1 in Appendix C compares the independent variables, coefficient signs and statistical indicators of model adequacy for six of the seven class size studies. Levin (1967) uses a technique other than linear regression and is therefore not included in this comparison. Table C.1 attempts to identify the variables each of the six class size studies has in common. Independent variables that occur in two or more of the studies are the following: sex; cumulative college GPA; cumulative high school GPA; high school economics taken; college class (i.e., freshman, sophomore, etc...); grade in economics course being taken; class size; time spent studying; degree of instructor (i.e., B.A., M.A., etc...); number of years instructor has been teaching; and semester or year dummy variable. When the independent variables that occur in two or more of the studies are compared on the basis of whether or not the individual coefficients are significant at the five percent level or greater, for coefficients in each of the studies, the list shrinks dramatically to include these: sex; cumulative college GPA; cumulative high school GPA; class size; time spent studying. When the independent variables that occur in two or more of the studies are compared on the basis of whether or not the signs on the coefficient are the same, for coefficients in each of the studies, the list is greatly reduced to include: sex; college GPA; high school GPA; college class; class size; time spent studying; and years teaching. The only independent variables included in two or more of the studies that are significant at the five percent or greater level and that have like coefficient signs are the following: sex; college GPA; high school GPA; class size; and time spent studying. The negative influence of class size on achievement is present in two studies and is significant at the five percent or greater level. Three additional studies report a statistically non-significant class size coefficient. Two of those find a positive sign for the coefficient while the third finds a negative. Two studies do not report the class size sign when the coefficient is found to be statistically insignificant.

No consistent sign for the following independent variables used in two or more studies emerges: high school economics; course as requirement or as elective; class size; and the degree status of the instructor. The lack of consistent sign within the same study but in separate regressions is also apparent. Crowley and Wilton (1974) run separate regressions for the post-test given under exam and non-exam conditions. Six of the 10 variables tested switch signs. Gramlich and Greenlee (1993) run separate regressions for introductory microeconomics and introductory macroeconomics. When macroeconomic regressions are grouped as a category, the signs switch on the following variables: SAT math; class size; desire to take the class; state resident; and student age. The conclusion for either study may be that alternating signs from one regression to the next may indeed come from the significance on which category is run or may be a result of model misspecification.

When r-squared or adjusted r-squared is used as an indicator of model adequacy, four of the six empirical studies find the measure to be 0.40 or greater with the average being 0.45. Lewis and Dahl (1972) are able to achieve an adjusted r-squared of 0.54 when a step-down procedure is used. Gramlich and Greenlee's (1993) adjusted r-squared is lower for introductory microeconomics and Crowley and Wilton's is lower as well when the post-test is not given under exam conditions. While the r-squared is generally acceptable for studies of this nature, the noted exception to the adequacy of the models is Williams, et al. (1985). The investigators test their model for different class size ranges and generally find an adjusted r-squared of less than 0.01, r-squared of less than 0.02 and the F-statistic of less than 1.50 for essentially all models tested over various class size ranges. The exception to their findings are class size ranges of 13 through 20 and 30 through 40 where r-squared and the F-statistic are 2.25 and 0.03 along with 1.92 and 0.04 respectively for each range. One weakness of this particular study may be the use of only a limited number of independent variables. Looking at the studies with better explanatory power allows a few suggestions to emerge. Achieving better explanatory power may mean including in a research design the following: post-tests given under exam conditions; a varied number of independent variables; and identification of the content for the introductory course offered.

A potential limitation in comparing independent variables across different studies is the definition each study gives to the independent variable. In most cases where the variable is defined at all, the actual method used to measure the variable is not usually discussed. This ambiguity becomes apparent when one tries to interpret the cross-sectional significance of independent variables such as the following: the use of quarter or semester hours in the calculation of GPA; whether or not ACT scores were converted from SAT composite scores; the elements included in a pre-test, if not nationally standardized, such as TUCE; the type of English or math course being tested and whether or not the variable is discrete or a dummy and why; the expected grade in a class could be influenced by the amount of feedback given during the semester, because determining the amount of feedback given is difficult; the content and emphasis of the course; the categorization of college majors; the instrument used to measure various student and teacher attitudes; measurement of student study time; and the modeling of class size. These potential limitations make drawing conclusions more difficult and the task of replication more arduous.

Gramlich and Greenlee (1993) is the only study that presents a formal discussion of potential biases, which include the following: measurement error; omitted variables; attrition; reverse causation. The other studies do not substantively discuss the potential bias issue. Failure to identify, account and correct for potential biases, such as recitation sections, plus/minus grade scales, differing section ranges, late registration, withdrawal or absenteeism may exaggerate the conclusions of the investigation. None of the studies identifies sample selection biases, nor does any mention that the correction for such biases could be a possible issue. Finally, discussions of unobserved or omitted variables is not readily evident.

1.3. Summary.

Becker, et al. (1991) along with Saunders and Walstad (1990) state the general conclusions of some of the class size inquiries in economics education. The conclusions may be less clear when the cited literature is dissected. Interpreting the influence of class size is difficult when such methodological diversity exists among the cited literature. Future research designs, including the one of this study, will need to consider the implications of this literature review so that replications can be facilitated and extensions improved.

Section 2: Estimating Educational Outcomes Using Production Functions.

The production function is a basic model that is used to describe how educational outcomes are influenced. Many educational researchers have supported this approach, most notably: Allison (1982); Coleman, et al. (1966); Hanushek (1979); Manahan (1983); and Davisson and Bonello (1976). The model asserts that educational outcomes are a product of educational inputs. Production functions are useful because marginal productivities, elasticities and scale economies may be calculated. Additionally, expansions of the production possibilities frontier can be described on the basis of input changes.

Production functions can be developed from both a macro and micro educational perspective. In the macroeducational framework the entire school or university provides a basis for describing outcomes. Aggregate outcomes may include attendance rates or college continuation (i.e., dropout) rates. The inputs in this perspective may be broad based and typically include cumulative data on facilities, instructional services and administrative infrastructure. In the microeducational framework the individual unit within the school system is described--typically the classroom. Classroom research focuses on outcomes, such as achievement or attitude in a class, and on how inputs, such as ability, effort, socioeconomic variables, peer influences, teacher characteristics and quality of instruction, influence the outcome. Classroom or coursework research in economics education has generally followed the microeducational approach.

The most often cited microeducational production function is authored by Hanushek (1979, 363). This basic model follows the form:

$$A_{it} = f(B_{it}, P_{it}, S_{it}, I_i)$$

where,

A_{it} = Achievement at Time t

B_{it} = vector of family background influences cumulative to time t

P_{it} = vector of influences of peers cumulative to time t

S_{it} = vector of school inputs cumulative to time t

I_i = vector of innate abilities

i = the ith student

In estimating this type of equation, data are needed on each input for every time period until t. However, in the absence of longitudinal databases similar to those suggested by Ewell (1992) as well as Buckles and Freeman (1984), these equations cannot be estimated due to lack of cumulative data. Additionally, family or peer data, if available at all, may at best be incomplete or can only be observed by proxy. If the production model is to be used in empirical work, then variables that can be observed and measured must be used. Variables which cannot be measured or observed are omitted, even if they are potentially significant.

The use of production functions in research for economics education can be assessed by asking four questions: (1) How do production functions define achievement?

(2) Are educational production functions derived from a theory of learning? (3) How should achievement in economics education be tested empirically? (4) What educational inputs influence the level of achievement in principles of college economics? The discussion which follows attempts to answer these questions.

2.1. How do production functions define achievement?

Defining achievement is not easy since there may be multiple outcomes even for a principles of economics class. W. Lee Hansen (1986) believes economics majors are asked to demonstrate five proficiencies. These proficiencies include gaining access to economic knowledge, displaying command of economic knowledge, drawing out economic principles, exploring issues using existing economic knowledge and creating new knowledge. Judith Yates (1978) lists the following skills as a few of the many skills that can be potential outcomes from education: application; critical thinking; creativity; comprehension; understanding; leadership; communication; interpersonal relations; and vocational.

Since there are multiple outputs from a college class, measuring achievement with a simple test score may be difficult. Eric A. Hanushek (1986, 1154) confirms this difficulty but also acknowledges that "Nevertheless, performance on tests is being used to evaluate educational programs, and even to allocate funds, and there are some pragmatic arguments for the use of test scores as output measures."

Achievement in economics education has been primarily measured as either the final grade or a test score in a course. While use of a final grade may capture more of the multiple outputs of class activities, cross-sectional comparison and replication of investigations may be difficult due to differences in grading standards and activities that comprise the final grade. The probable use of non-discriminating test scores can make cross-sectional comparison and replication of investigations difficult as well, especially when the content of exams may differ.

In order to overcome these comparison and replication problems the National Council on Economic Education (NCEE) developed the Test for Understanding of College Economics (TUCE). TUCE is a 30 or 33 question nationally normed multiple-choice exam. The three question difference allows for optional testing in international economic concepts. Versions A and B for each class in principles of microeconomics and macroeconomics are available. The different forms allow for pre- and post-testing.

TUCE III has been improved from its two predecessors in that the exam was designed "...to discriminate across a broad range of intellectual ability and knowledge." (Examiners Manual 1991, 7) The test is not an attempt to measure everything that goes on in a classroom. However, TUCE III is the most reliable, valid, statistically justified, and widely accepted test of general economic knowledge (Examiners Manual 1991), despite potential criticisms (Swartz, et al. 1980). The utility of the TUCE III exam will allow for a more controlled design of scientific experiments. Use of the TUCE III should facilitate cross-sectional comparison and replication of investigations measuring achievement in the same way.

The number of correct responses on the TUCE III exam is a starting point for modeling the dependent achievement variable. Approximately four accepted ways to use a TUCE III score as an achievement measure are identified by Siegfried and Fels (1979). In their article in the *Journal of Economic Literature* (1979, 929), they write:

> The availability of two matched forms of the TUCE has led to pre- and post- course testing, which permits several forms of the output measure [the dependent achievement variable] to be specified: (1) absolute achievement--the post-test score; (2) absolute improvement--the difference between the post-test and the pre-test score; (3) percentage improvement--absolute improvement divided by the pre-test score; and (4) gap closing measure--absolute improvement

divided by the potential gain in score (which is the
difference between the perfect score and the pre-test score).

The absolute achievement score measures a stock of economic knowledge at a given point in time. The absolute improvement score identifies incremental learning occurring during an economics course. The last two measures acknowledge that the same difficulty of learning may not be continuous through a course. The percentage improvement measure suggests that poorer students have a harder time improving their scores while the gap-closing form suggests that students beginning with a larger amount of economic knowledge will have a more difficult time improving (Siegfried and Fels, 1979). Most researchers have defined achievement using definition one, although some have used definition two or a variation thereof. The variation uses the post-test score as the dependent variable with the pre-test as one of the independent variables. The inclusion of the pre-test achievement score as an independent variable may allow the influence of omitted or unobserved variables, such as innate ability, and family or social influences, to be captured. Essentially the pre-test becomes a stock variable. The benefit of including a pre-test independent variable may be to lessen the size of the bias of the estimated coefficients. Hanushek writes, "The importance of these omitted factors is lessened if the model is estimated in value-added form because any level effects have already been included through entering achievement and only growth effects of innate abilities have been omitted" (1986, 1157).

Another advantage to using TUCE III is that the Examiners Manual (1991) identifies the specific test questions which relate to different educational abilities. Bloom's Taxonomy of Educational Objectives (1956) identifies six educational objectives: Knowledge; Comprehension; Application; Analysis; Synthesis; Evaluation. TUCE III is capable of testing achievement in the first three categories.

TUCE III questions which are identified by an RU relate to the "Knowledge" objective. RU means that students are able to recognize and understand basic terms, concepts and principles. EA and IA relate to the "Comprehension" and "Application" categories. EA means the questions are designed for the explicit application of basic terms, concepts and principles. While IA stands for the implied application of basic terms, concepts and principles (Examiners Manual 1991, 3).

RU, EA and IA scores can be used to measure achievement in the absolute achievement, absolute improvement, percentage improvement and gap closing forms. Just using a post-TUCE III score could mask the achievement gained in "Knowledge" as well as the "Comprehension" and "Application" categories. If RU measures basic understanding in the principles course, this measure may more accurately reflect a student's mastery of basic knowledge. Alternatively, EA and IA scores could be combined to determine if critical thinking in economics has been achieved. Nevertheless, some evidence presented by Lewis and Dahl (1971) suggest that EA and IA questions should be cautiously interpreted when using them as a proxies to measure higher ordered skills such as critical thinking.

Finally, using a standardized test like TUCE III as a measure of achievement does not mean that the influence of other multiple outputs of economics education are ignored. Variables like attitude or critical thinking can be viewed as indirect outcomes of the introductory course and as such can be included as independent regressors in an estimable production function.

2.2. Are educational production functions derived from a theory of learning?

While production functions are theoretically appealing to economists, their application to the educational process is met with some reservations. According to Becker (1983A, 5), educational production functions do not address the appropriateness of the function or the behavioral assumptions underlying their construction. Becker (1983A, 5) writes:

> Production function studies are based on observed behavior. Observed behavior reflects more than a technical, static relationship between inputs and outputs. It also reflects some type of optimizing decision rule involving many variables, including the production function the students themselves elect or are forced to use...There is no reason to assume that students are or will be generating maximum learning for a given set of inputs or that observed values are mapping only the production frontier. As such, the specification and estimation of a production function relationship...continues to be questionable.

Very few theories that attempt to describe learning behavior in economics education have been presented. The theories that have been presented can be categorized as either professional judgments or as rational hypothetical deductions.

Theories based on professional judgment do not attempt to model relationships between learning determinants, although they usually discuss a principle that may facilitate learning. For example, information processing theory is concerned with how the mind processes information. The theory makes a recognition between the operation of short and long-term memory. Long-term memory codes information in hierarchies. Transferring any knowledge to long-term memory may require educators to hierarchically structure discipline content. While psychological confirmations for this assertion may exist, professional judgments in economics education tend to lack empirical proof. David Martin (1982) discusses the implications of developmental theory, information processing theory and operant conditioning for K-12 economics instruction, but does not offer empirical proof to justify the implications. Van Metre (1976) advocates matching method of instruction to learning type along with specifying behavioral course objectives, but does not offer evidence on the effectiveness of these changes. The appeal of professional judgment is directed more to intuition in professionals who should "know that these suggestions will work" rather than in those who seek empirical support.

Alternatively, the rational hypothetical deductive models attempt to reason from a set of assumptions, much like the development of consumer or producer theory. This approach is more in line with the traditional methods economists use to develop theory and also tries to answer Becker's (1983A) criticism by specifying behavioral assumptions. For example, McKenzie and Staff (1974) develop a model of student choice. The representative student chooses between achievement levels in fields of knowledge. The student is constrained by limited time devoted to effort and leisure. Consequently, the model predicts that a student will choose a combination of achievement in various courses which will maximize utility. These rational hypothetical deductive theories are partial rather than general equilibrium models because only some learning behavior relationships, not all possible interactions, are described. Partial equilibrium models attempt to describe either professorial choices (Hansen and Kelley 1973; Needham 1975; Becker 1982; Allen 1980; Lima 1981; Becker 1975; Becker 1979; McKenzie and Staaf 1974) or student choices (McKenzie and Staaf 1974; Kelley 1975). Most of these theories are based on utility models, comparative statics and optimizing behavior. Others are based on queuing theory (Mulligan 1984) or dynamic modeling through time (Wetzstein and Broder 1985). The typical behavioral relationships may describe trade-offs between variables like time, leisure, achievement and research. The difficulty with all of these models rests in the need for empirical confirmation. So far, few of these theories have been put to a test.

Kelley (1975) is the one of the few researchers to present an empirical assessment of student choice theory when course technology improves. A multivariate regression equation is used to determine if prepared lecture notes influence achievement. The coefficient of prepared lecture notes is negative and statistically significant which implies, ceteris paribus, that principles of macroeconomics is an inferior good since the improved technology actually reduces achievement and utility when educators would expect both achievement and utility to increase. The application of the student choice model is more descriptive and graphical than it is empirical. Becker (1983A), in using a partial equilibrium inventory model of student study time, reduces the theoretical model to an estimable equation. However, because the estimable equation lacks key variables known to influence achievement, estimation of the equation would be incomplete. Becker (1983A) does not discuss this concern, but this literature review shows that his estimable equation excludes inputs known to influence achievement. The practical consequence would be to include Becker's (1983A) estimable variables with other known influential variables. The problem here is that the influence of the additional variables is not predicted by the partial equilibrium model. Salemi and Tauchen (1982) attempt to model learning behavior using a production function approach. They identify three sub-models: (1) a model of the learning process; (2) a model of the differences in student aptitude; and (3) a model of the relationship between aptitude and test scores. Similar to Walstad (1987), Salemi and Tauchen (1982) use a two-stage least squares method to model both attitude and achievement. However, the model still does not describe how choices are made to optimize learning.

In summation, professional judgment and models based exclusively on a theory of learning that describes either student or professorial choices have not had strong empirical support. Models that are based on recognizing simultaneous relationships and that have been empirically tested lack a theory of learning. The question that arises, therefore, is how the development of learning theory should proceed. Hilgard and Bower (1975, 610) comment:

> applied science can't wait for the answers of pure science to come in: crops must be planted and gathered, the sick must be treated, and children must be taught with whatever tools and knowledge are available at the time... Skilled teachers

> contribute to educational advance, with students of psychology of learning sometimes bringing up the rear.

While the learning theory approaches are less than ideal for estimation in economic education, applied research using the traditional production function method can still help illuminate influential inputs. The discovery of these inputs and the way in which they are modeled may lead, in turn, to the development of an estimable theory of learning behavior. "Consequently, the model is based on 'a priori' knowledge and research, limited educational theory, and to a certain extent, intuition" (Walstad and Soper 1982). Known information and expected relationships between variables are the basis for model development in economic education.

2.3. How should educational outcomes be tested empirically?

The influence of various independent variables on achievement in economics has been tested using a generalized linear regression model described by Becker (1983B). This model follows the form:

$$Y_i = B_0 + \sum_k B_k X_{ki} + e_i$$

where,

Y_i = an accurately measured and continuous post-test score

X_{ki} = the kth regressor for the ith subject. May represent a dummy variable or any explanatory variable

B_k = parameter to be estimated

e_i = error term, assumed to be normally distributed, with mean zero, constant variance, zero covariance and no correlation with explanatory variables.

Testing variables using this general linear regression allows influences identified in previous investigations to be confirmed or denied. Further, new variables can be added to the estimation process to determine their significance.

In finding an acceptable method for thinking about which variables are important, it is useful to categorize potential influential inputs. Davisson and Bonello (1976) categorized inputs into three types. First is human capital, which may include pre-test scores and prior grades. Second are utilization rates that indicate the degree of access to or production of knowledge. Utilization rates may include time spent in study, class or research. Third is technology which may include teaching methods or equipment. Milkman and Tinkler (1990) identify influences on the learning process as related to faculty, administrative or student characteristics. Siegfried and Fels (1979) work with the categories of student human capital, faculty human capital, college environment, student effort and teaching methods. The advantage to categorizing potentially influential variables is that "...[this] taxonomy is superior to totally ad hoc specifications of learning production functions" (Becker, 1983A).

2.4. What educational inputs influence achievement in principles of college economics?

A variety of studies investigating the influence of faculty capital, student capital, technology and institutional factors on student achievement in principles of economics have been performed. Siegfried and Fels (1979) cited 179 references on economics education research. According to Siegfried and Walstad (1990) the number of citations could be updated to 250, but few of the newer studies would be considered "classics." This section identifies the following potential influences on student achievement as they are discussed in the literature: faculty capital; student capital; technology; and environmental factors.

2.4.1. Faculty Capital.

Faculty capital could be considered the set of personal characteristics, background experiences, utilization rates and teaching methods that influence student achievement. Teaching methods are most commonly measured by student evaluations of teaching (SET). SETs can be used as a measure of teaching effectiveness. The more effective the teaching, the better the achievement. Research indicates that students rate teaching more effective when the instructor exhibits the following attributes: clear presentation; enthusiasm; respect for student opinions; organization; signaling (i.e., verbal statements drawing attention to a point); active student involvement; rapport with students; and communication skills. Background experiences that are positively related to achievement include: years of teaching experience, instructor TUCE scores, graduate school grades and learning style. The influence of personal characteristics, such as gender, race or other sociodemographic variables, is not reported. One study finds that the age or maturity of the instructor positively influences achievement (Lewis and Dahl 1972). Another inquiry shows that different instructors influence achievement differently (Clauretie and Johnson, 1975). Utilization rates, such as office hours, availability, and research commitments, have not been discussed. Finally, graduate students are used as instructors (GSI) in many universities and appear to be just as effective at teaching as regular faculty, unless the GSI is a non-native speaker of English.

2.4.2. Student Capital.

Student capital may include the set of personal characteristics, innate ability and effort that influence student achievement. Personal characteristics which will likely improve achievement include: age; being male; years in school; being married; student interest. One personal characteristic that has been shown to reduce achievement is status as a freshman in college, although the variable "years in school" shows no relationship to achievement. Males tend to do better than females in introductory economics. However, males can be shown to enter the class with a greater stock of economic knowledge and, if this is accounted for, both males and females learn at the same rate given their relative starting points. Additionally, multiple choice exams, which are typically used in introductory courses or in pre/post-testing, have been shown to favor males. Student interest may be related to attitude. Siegfried and Walstad (1990) discuss the conventional belief that positive attitudes relate to improved achievement. However, their review of recent investigations find attitudes determined by the amount learned rather than the amount of learning determined by attitude. Personal characteristics that include sociodemographic variables, such as family income and parents' education, are usually non-influential. The innate ability of students is the stock of ability or stock of economic knowledge brought to the economics class. Higher SAT, ACT and introductory economics pre-test scores positively influence achievement. SAT verbal scores seem to be more predictive than the quantitative scores. An increasing GPA seems to improve achievement, while the number of previous college credit courses seems to be inconsequential. Having had a high school economics course does not impede students taking their first college principles course. One study found that the grade earned in high school economics can positively influence achievement in introductory college economics (Myatt and Waddell 1990). Mathematical preparation does not appear to be influential. As study effort increases, achievement has been found to improve. Study effort can be measured by the amount of time spent studying (Lewis and Dahl 1972) or by a proxy variable such as the final grade in economics divided by SAT (Prince, et al. 1981). Outside employment may also lead to lower achievement. No suggestions have been made regarding the type of inducements instructors can offer as an incentive to exert more study effort.

2.4.3. Technology.

Teaching technologies can include the use of computers, videos and various other methods to deliver course content. The computer has been used as a tool in computer aided instruction (CAI) and computer managed instruction (CMI). Games, simulation models and demonstration routines are typical applications of CAI. Review routines utilizing short quizzes are an application of CMI. Both CAI and CMI provide instant feedback, novelty and convenience. However, computerized instruction does not appear to enhance learning. A potential efficiency issue arises since conventional pedagogical methods probably cost less. For example, some games may be used without the use of computer technology, in which case computerization of the game may not be cost effective. The current generation of college age (i.e., 18 through 22 years old) students has been raised in a more visually oriented society. The use of television programming, televised lectures and videos would seem to complement the visual orientation of students. However, research findings indicate that television does not necessarily contribute to achievement. Live lectures may be better, and students appear to prefer the live presentation. Classroom teaching methods also involve other methods of delivery, such as the use of innovative pedagogies. Programmed instruction and personalized systems of instruction have been shown to significantly and efficiently improve learning. Pedagogies that have positively influenced achievement include the consistent reading of *The Wall Street Journal* by students (Lewis and Dahl 1972), while the use of case studies (Haley 1991) or designing a writing intensive curriculum (Milkman and Tinkler 1990) does not appear to significantly influence achievement. The availability of lecture notes to students negatively influences achievement by creating time savings which students can then use to engage in activities other than studying for an economics class (Kelley 1975). The impact of behaviorally anchoring course objectives on achievement is met with inconclusive evidence. Siegfried and Fels (1979) found research evidence to support both positive and negative effects of specifying student behavior for course objectives. Recent research on how the choice of a textbook or the text's readability influences achievement is not available. Additionally, the usefulness of classroom experiments involving group work, group grades, discussion format and frequency of assignments or exams has not been widely discussed in the professional literature. Many of the technological studies are dated and need to be revised, especially with the advent of new technologies such as CD ROM, satellite transmissions, advancements in personal computing and interactive multimedia.

2.4.4. Environmental Factors.

Environmental factors describe the variables that the university, rather than the student or instructor, controls. Course sequencing is controlled by the institution and there is evidence to suggest that students who take micro before macro principles perform significantly better, although students seem to like macro before micro. Additional institutional variables, such as the time and days the course is offered, the type of school (i.e., doctoral, comprehensive, liberal arts, community or otherwise), the use of recitation sections, the availability of university support services and the total quality management orientation of the administration, have not been thoroughly discussed in relationship to their respective impact on achievement in introductory college economics.

2.5. Summary.

Production functions used in economics education attempt to identify educational inputs that influence outcomes, such as grades or standardized test scores. In the absence of an adequate learning theory that lends itself to estimation, economics educators use empirical studies to help specify categories of influential inputs. The categories and variables reported here are the most commonly accepted. Some research on achievement has included class size as an input variable because of its expected effect or its inclusion in prior research. Class size is not usually included in research designs because of any theoretical model specifying the expected influence of class size. Researchers in economics education appear to have overlooked the potentially important contribution of modeling class size influence.

Section 3: Modeling Class Size Influence.

Research investigations which empirically try to determine the influence of class size in principles of introductory college economics usually employ a single representative independent variable. The variable is at times the actual number of students in the class or a dummy variable indicating whether or not the class is large. Only Williams, et al. (1985) use an equation containing a class size variant, which is class size squared.

Achievement can be objectively measured with a standardized exam like TUCE III. The number of individuals taking a class is fairly easy to measure. However, linking class size and achievement is not as obvious. Economics education research has assumed that class size has a direct influence on outcomes; this assumption may be too simplistic, however.

The review of the educational literature reveals that Mitchell, et al. (1992) detail six theories on the impact of class size on achievement. The investigators pose the question, "Why should we expect that removing some children from a classroom would cause the remaining ones to learn more?" (1992, 40). In answering this query they present competing models that can be classified as either direct or indirect. Although the application of these models was to data in grades kindergarten through third grade for Tennessee Project STAR (Student/Teacher Achievement Ratio), extensions for courses in higher education may exist. The models that follow are adapted from Mitchell, et al.

(1992) in order to provide insight into the potential link between achievement and class size in introductory college economics.

3.1. Direct Class Size Models.

The models that assume a direct influence between class size and achievement are identified as the following: Instructional Overhead; Interaction; Fixed Resource Distribution. Direct is defined as the actual number of students in a class can be used to predict the level of achievement.

3.1.1. Instructional Overhead.

In this model, achievement and class size are related linearly. The larger the class, the more time is spent on instructional administration. This theory recognizes that the opportunity cost for maintaining large orderly classes is that less time may be spent on assisting students with learning objectives. The simplest model could be constructed as follows:

$$A_i = K - B_1 C_{ij}$$

where,

A_i = Achievement for Student i

K = Constant or Amount of Knowledge to Be Acquired

B_1 = The Contribution of Class Size to Achievement

C_{ij} = Class Size for Student i in Class j

3.1.2. Interaction.

As class size grows, students and instructor have more opportunities to interact. The number of possible interactions for the group grows exponentially with the addition of another student. Interactions take time and therefore reduce opportunities for learning. A student interaction time theory model would suggest the following possible relationship (descriptions of previous variables will not be repeated):

$$A_i = K - 0.50B_1[C_{ij}^2 - C_{ij}]$$

multiplying out,

$$A_i = K - 0.50B_2C_{ij}^2 + 0.50B_3C_{ij}$$

where,

C_{ij}^2 = Class Size Squared for Student i in Class j

3.1.3. Fixed Resource Distribution.

Since every college instructor has a fixed set of resources available, as the number of students increases, the fraction of the resources available to each student becomes smaller. Because the slope of a natural logarithmic curve is asymptotic, this model may be specified as follows:

$$A_i = K - B(LNC_i)$$

where,

LNC_i = The Natural Log of Class Size for Student i

3.2. Indirect Class Size Models.

The indirect models are named: Heterogeneity; Instructional Pacing; and Ability Grouping. Indirect means that the composition of ability within classes, not the actual number of students, affects achievement through its influence on teaching efforts. Another way of saying this is that the achievement level is a function of the strata of ability in a class, while the strata of ability is a function of class size. As class size increases, the strata of ability increases and, as the strata of ability increases, a greater influx of less able and more able students will diffuse teaching efforts so that achievement declines. The strata or range of ability is positively related to class size, but negatively related to achievement.

3.2.1. Heterogeneity.

Not all students have the same abilities in a class. Hence, achievement may be related to the range of student abilities within a class or the standard deviation of ability for the class.

Ability may be measured for principles of economics in any number of ways; ACT, pre-TUCE, GPA. The range in ability for a class is the difference between the highest and lowest student score. If students randomly selecting a course come from a normally distributed population, then small classes should have a smaller range of ability since students will score closer to the population mean (Preece 1987). Therefore, teaching to a more homogenous group is possible with small classes. Small classes should then receive more focused teaching efforts, which would improve achievement for the entire class. Teaching efforts in large classes may be more varied, less focused, or marginally effective since the range of ability is greater. The consequence would be a reduction in achievement for large classes.

Although instructors may be able to cater to small numbers of low or high achieving students in large classes, which would seem to negate the influence of ability range, what matters is the degree to which the group is homogenous. Even large classes may be characterized by the same degree of homogenous ability as a small class. Homogeneity may be approximated by the standard deviation of ability for the class. A small standard deviation would result in a more homogenous group, which could be more effectively taught through precisely targeted teaching efforts. As the standard deviation of class ability falls, achievement rises. Alternatively, as the standard deviation of class ability increases, achievement declines due to the diffusion of teaching efforts across more heterogeneous abilities.

Both the range and standard deviation effects are indirect, but still a function of the class size. The heterogeneity model can be specified as follows:

$$A_i = K - B_1 R_j - B_2 S_j$$

where,

R_j = Range of Student Ability in Class $_j$

S_j = Standard Deviation of Student Ability in Class $_j$

3.2.2. Instructional Pacing.

Here the theory suggests that teaching efforts cater to the lowest ability student in the class, which in turn slows the class down. A first consideration of the pacing effect would be the lowest ability student in the class influencing the achievement of every class member. A second may be the degree to which the lowest ability student deviates from the class mean. A third may be the mean class ability. Finally, a fourth consideration may be the skewness of student ability within the class. Achievement is anticipated to be negatively influenced if teaching is paced to the low student or lower average ability students or classes which are characterized by low-ability. The instructional pacing effect may be modeled as follows:

$$A_j = K - B_1Lj - B_2D_j - B_3M_j - B_4Q_j$$

where,

L_j = Least Able Student Score in Class j

D_j = Deviation of L From Class j Ability Mean

M_j = Class j Ability Mean

Q_j = The Skewness of Class j Ability Distribution

3.2.3. Ability Grouping.

Ability grouping recognizes the tendency for instructors to assign or students to organize study groups. Although the grouping may be formal or informal, the groups usually contain students of like ability. The lower ability students may become discouraged by trying to keep pace with the highest achieving student in the class, or the lower ability students may work below their ability when seeing the pace of the least able student in class. Therefore, the most or least able student in a class may negatively influence achievement. Ability grouping may also refer to the degree to which class ability is clustered or spread out over a range of ability; this may be measured by the kurtosis of ability in the class. As the number of ability clusters rise, achievement declines. Modeling these ideas could take the following form:

$$A_{ij} = K - B_1H_j - B_2L_j - B_3Ku_j$$

where,

A_{ij} = The Achievement of Student i in Class j

Hj = The Highest Student Ability Score in Class j

L_j = Least Able Student Score in Class j

Ku_j = Kurtosis of Class j Ability Scores

3.3. Summary.

Class size may assert a negative influence on achievement through either direct or indirect variables. While these models of class size influence may provide a more accurate description of the dynamics of class size on achievement, they have generally not been tested in economics education research. In testing these models, one commonly acceptable method would be to include the variables in a production function category, such as environmental factors, since the size of the class is controlled by the university or department.

Section 4: Chapter Summary.

Very few research inquires exploring the link between class size and achievement in introductory college economics have been attempted. The inquiries usually cited are dated. Most of the studies have discussed the influence of class size as an incidental rather than a primary focus of the research design. Because of the methodological diversity found in the literature, the significance of the class size variable is not clear. The literature does agree on the use of production functions to model educational outcomes. While researchers disagree on whether learning theory should precede estimation or vice versa, the general approach has been to empirically test categories of suspected or known influential inputs as suggested by Becker (1983A). Inputs have not generally been tested on the basis of a theoretical model. Most empirical designs have allowed only class size to influence achievement as a first-order, non-interactive input. However, recent educational modeling of class size as suggested by Mitchell et al. (1992) would suggest that indirect as well as direct modeling of class size should be explored. A research design updating the class size literature in principles of college economics should replicate common foundations of previous work and extend the empirical analysis to include current theoretical modeling of class size impact on student achievement.

Chapter III

Data Preparation, Sample Selection and Research Methodology

The link between class size and student achievement in principles of college economics has not been tested using data from a national sample. However, national data are now publicly available from the National Council on Economic Education (NCEE). The data were gathered during the norming of TUCE III in the fall and spring terms of the 1989-90 academic year. Sixty-two separate institutions are represented in the data set covering 189 separate courses in either principles of macroeconomics or principles of microeconomics. There are 93 macro sections and 96 micro sections each representing 45 and 49 schools, respectively. NCEE offers the information in 189 files, one file per course section. The individual course section is the common unit used to assemble the 7-16-91 version of the TUCE III data. Faculty, student and institutional information is available for each student observation in a course section file. Preparing the data requires designing a database structure then appending each of the course section files into one aggregated file. A unified database allows for the selection of a cross-sectional sample. This chapter identifies the variables on which NCEE information is available, the preparation of the aggregate TUCE III database and the selection of the sample for this research inquiry. It also summarizes the methodology used to test the influence of class size on achievement.

Section 1: Identification of Variables.

Each observation in the raw data has approximately 208 descriptive variables prior to any variable transformations or additions. Information for each observation is compiled from faculty questionnaires, student questionnaires and student scores on the pre- and post-TUCE III. Essentially, each question on the questionnaires and each question on the TUCE III tests are descriptive variables. This section will identify variables in each of three categories: the institution offering the course; the instructor teaching the course; the student enrolled in the course.

1.1. Institutional Variables.

Data are collected from 62 institutions in 36 states. Institutional variables relate to either the college or the course taught. Institutional variables are administrative in nature and not instructional. Each observation is identified by five essential institutional variables: type of institution (i.e., doctoral, comprehensive, liberal arts, two year); selectivity rating of the institution (i.e., the estimated ability level of the student body compared to other colleges); term of a course (i.e., fall, spring); type of course (i.e., principles of macroeconomics, principles of microeconomics); and section code of the course. Additional institutional factors, taken from instructor questionnaires, indicate the number of meetings per week, number of minutes per meeting, number of weeks per term for each course section and whether or not the course section meets in a separate discussion or recitation group.

1.2. Instructor Variables.

Faculty variables attempt to outline the background of the course instructor, provide measures of teaching effectiveness when rated by the instructor, show how TUCE III was administered and describe the organization of the course. These variables are gleaned from instructor questionnaires.

Variables identifying the background of the instructor include the following: gender; ethnic origin; English as the instructor's native language; faculty status (i.e., regular tenure track, adjunct or part time, graduate student, other); publication of an article in a refereed journal within the last five years; highest degree earned (i.e., B.A., M.A., Ph.D.); the year the highest degree was earned and from what institution; the number of years teaching; the number of years teaching the present course; the percent of time spent in teaching, research, administration, consulting or other activities; and the number of classes as well as the number of separate courses the instructor teaches during the present term.

The instructor questionnaires ask each faculty member to rate their teaching effectiveness in comparison to other college professors. The teaching effectiveness variables measure, on a rank order scale, self-perception in: enthusiasm about teaching; preparation for class meetings; ability to speak English; rigor or grading standards; and overall teaching effectiveness.

Instructors are allowed to determine if students should take either pre-TUCE III, post-TUCE III or both tests. If the pre-test is given, NCEE suggests that it be given during the first day the class meets. However, some leniency in the administration of the post-test is permitted. Instructors can determine if the 30 or 33 question post-test would be given. The three questions that make up the difference relate to international economics. Presumably, the post-test matches the number of questions given on the pre-test. Instructors also are allowed to allocate a certain number of minutes for taking the post-test. Moreover, they may decide to count the post-test toward the course grade and to determine the percent the test counts toward that grade. Variables are available on each of these preceding points. There are no data available on course time allocations of the pre-TUCE III or the role the pre-exam may play in determining a course grade.

Five variables are available on the organization of the course: the lecture-presentation style; the intensity of use for ancillary materials; the components of the course grade; a profile of how tests are used in the course; and the textbook used in the course.

First, lecture-presentation style is assessed by the amount of time the instructor believes is spent in lecturing, responding to student questions, students responding to instructor questions and other activities. Second, the intensive use of ancillary course materials class is identified. Low intensity implies "recommended use but not graded" while high intensity implies "required use" and "graded." Ancillary materials include items such as the following: study guide or student workbook; homework problems or problem sets; non-text readings and handouts; computer exercises; television programs or movies; and other activities. Intensity is a scaled variable from one to four, with four being the highest intensity. The scale is available for each teaching ancillary used in the course. Third, components of the course grade are identified. Information is available on the number and weight of grade components, such as the following: quizzes; homework problems; mid-term exams; term papers; final exams; and "other activities" not identified by the previous components. The percent each item contributes to the course grade is also available. Fourth, "the use of tests" describes the percent of the final exam, which is comprehensive, and the percent the final counts over the last part of the course. Additionally, the weight of different question types (i.e., true or false, multiple choice, short answer, long answer, other) used on exams is also provided. Fifth, the instructors also provide the name of the author(s) of the text used in the course, and they state whether or not the textbook is the same one used in the previous semester.

If the course has a separate recitation or group discussion section, the following variables are available on the recitation section discussion leader: gender; ethnic origin; English as the discussion leader's native language; and faculty status.

1.3. Student Variables.

Student variables attempt to provide information on student background, performance in the present economics course, assessment of teaching effectiveness, feedback on introductory economics and scores on TUCE III. The student variables are primarily gathered through data provided on the student questionnaires.

Variables on student background describe the following characteristics: gender; ethnic origin; SAT or ACT scores; college credit hours of coursework completed prior to the present term; GPA; college credit hours in calculus completed prior to the present term; college credit hours in economics completed prior to the present term; college credit hours in macroeconomics and/or microeconomics completed prior to the present term; completion of high school economics; type of high school course completed (i.e., macroeconomics, etc.); and duration of high school course (i.e., one year, semester, etc.).

Variables describing performance in the present economics course include the actual course grade along with the grade the student expected to earn. Other variables related to performance include: college credit hours being attempted this term; working in a job and hours per week working; hours per week studying and hours per week studying economics.

Students are asked to evaluate the instructor's effectiveness in comparison to other college instructors. The comparison is made on the basis of the instructor's enthusiasm about teaching, preparation for class, ability to speak English, rigor of grading standards and overall teaching effectiveness. Discussion leaders of recitation sections are also evaluated on each of these criteria. Teaching effectiveness is assessed by asking students what percent of the 100 percent of knowledge they gained from the class is attributed to the regular instructor, discussion leader and various ancillary learning aids. These ancillary learning aids include the following: textbook; supplemental workbook or problems; non-text readings and handouts; computer exercises; and other activities. The percent learned from term papers is also measured.

Students are asked to provide feedback on introductory college economics by ranking, in comparison to other college courses, the following: amount learned in the course; interest in the subject matter; importance of the subject matter; difficulty of the subject matter; and quality of the textbook.

Finally, scores are available for students who completed the pre-TUCE III, post-TUCE III or both tests. Responses to all questions are recorded. Scores indicate the number of correct responses on either the 30 or 33 question test. The three additional questions relate to international economics. The instructors determine which version of the test to administer. The number of correct responses for recognition-understanding (RU), explicit application (EA) and implicit application (IA) type questions is also recorded for each student. The Examiner's Manual (1991) for TUCE III identifies each question as either a RU, EA or IA type. In the 30 question pre and post-TUCE III there are 10 questions of each type.

1.4. Variable Additions and Transformations.

In order to improve the explanatory power of the variables used for statistical purposes, several variables in the NCEE raw data are either added or transformed. Nine points will be made regarding the manipulation of the raw data.

1. Since the TUCE III data contain public as well as private schools, a variable is added to indicate the status of the institution. The public-private dummy is set equal to one if the institution is private. Seventeen of the 62 institutions are private.

2. Institutional type is indicated by a single number in the original data. To facilitate interpretation, dummy variables are created for each institutional type. For example, the doctoral institution variable is set equal to one for doctoral institutions and zero otherwise.

3. The section code is transformed so that courses with only one section are denoted by a zero rather than as a blank.

4. The highest degree earned is transformed into a dummy variable with an earned doctorate being set equal to one.

5. The intensity of use for the ancillary teaching aids is a scaled number. Dummy variables are created for each ancillary aid and set equal to one if the intensity of use is greater than two. Greater than two indicates aids that are assigned but not graded, used frequently, required, collected or collected and graded.

6. All the variables for the direct and indirect class size models are added to the database. The direct variables of class size influence include: class size; class size squared; and log of class size. The indirect variables of class size influence include: maximum class ability score; minimum class ability score; standard deviation of class ability; average of class ability; range of ability within a class; deviation of minimum class ability score from average of class ability; skewness of class ability; and kurtosis of class ability. The pre-TUCE III 30 is used as the measure of ability that each student brings to the classroom.

7. Since not all students enrolled in a course section will have taken the pre-TUCE III, a variable is created to determine the percent of students taking the pre-test out of the total number of students in the course. The percent of students enrolled per course section completing the pre-TUCE III is found by dividing the number of students taking the pre-TUCE III by the total number of students in a course section.

8. The percent of time during a class that is devoted to instructor-student dialog is created and named RESPOND. RESPOND is found by adding the percent of time the instructor spends responding to student questions with the percent of time that students respond to instructor questions, then dividing this sum by the percent of time that the instructor spends lecturing. As the percent of time spent in instructor-student dialog grows, RESPOND becomes larger. The more time the instructor spends lecturing,

RESPOND becomes smaller. Ninth, dummy variables for various class size ranges are constructed. The reasoning for this addition is briefly discussed in the methodology section of this chapter and in more detail in section 4.2 of chapter four.

Section 2: Preparing the Aggregate TUCE III Database.

The TUCE III data were obtained from the NCEE in disaggregated form. Each observation in a course section file is a student taking that course section. When all students in all course sections for both macro and micro courses are combined there are 9,679 student observations in the aggregate database. This chapter section will describe the process of assembling the aggregate TUCE III database.

Robert J. Highsmith, the Vice President for Program and Research at the NCEE, has no knowledge of the NCEE data being combined into one database file. This is the first attempt to aggregate the NCEE files on a personal computer. The combined TUCE III files are easily accessible and represent the most comprehensive cross-sectional national database available for introductory college economics research.

The 189 separate files from NCEE arrive on diskettes. The data in each file are formatted in ASCII. There are 310 columns of information for each observation, although not all observations contain complete information. There are no commas or spaces used for field delimiters, only an end of record indicator. After reviewing several databases, DBase III Plus (1986) was chosen to combine the files. The first step is to define the database structure, which means breaking the 310 columns of information into the various variable ranges identified in the TUCE III Data File Codes (7/16/91) guide available from the NCEE. Initially, 208 variable fields were created. Each of the 189 course section files is appended to the aggregate database.

The next step in preparing the data is to create new variables in the aggregate database. The primary additions are the direct and indirect class size variables. Adding these variables requires the use of calculation operators for each course section. DBase III Plus (1986) is too cumbersome to accomplish this task, but Access for Windows (Microsoft 1992) can accomplish the task of adding these variables more simply. The TUCE III aggregate database is therefore downloaded to ASCII with commas as field delimiters and then read into Access for Windows (Microsoft 1992). The addition of the various dummy variables and the transformation of simple variables, such as RESPOND, is straightforward. However, the computation of the class size variables and the process of writing the results to each record requires more programming skill.

The number of students in a class is not an original variable obtained from NCEE. Class size is created by proxy. The number of students in each course section is unknown. However, the number of records in a class is known. A record exists for all students in a class who took the pre-TUCE III, the post-TUCE III or both. This study makes the assumption that all students will have taken either the pre- or the post-TUCE III test. Consequently, the number of students in a course section file will approximate the number of students enrolled in a class. Some students chose not to participate in the norming of TUCE III. Determining the number of students who did not take the pre- or the post-test is not possible. Students who did not answer all TUCE III questions were dropped from the file, and the number of students who were dropped was not recorded. Determining the number of students who withdrew from the course also is not possible. Although the number of students withdrawing from the class or choosing not to participate is assumed to be insignificant, the number of students in a course section file is not exactly equal to the number of students enrolled in the course section. Therefore, class size is measured by proxy.

Computing class size requires adding up the number of observations in each course section file, as this number is the proxy for class size. As each file is appended to the aggregate database, the number of observations in each file is recorded manually then entered into a separate indexed table. This table contains the course section identification for each of the 189 courses. The table is created manually and contains the identifiers for each course section file: term; school code; course type; section code; and class size. An "update" command is then used to add the class size variable to each of the 9,679 observations in the aggregate database. Once class size is available for each observation, the other direct class size variables that influence achievement, namely class size squared and the log of class size, can be computed easily.

The indirect variables of class size influence present a challenge similar to that of computing the proxy for class size. None of the indirect variables is included in the NCEE data. As a first step, a measure of ability must be chosen since ability is a key component to the calculation of the indirect variables. The four possible measures of ability within the database are: pre-TUCE III 30 score (8,238); GPA (5,360); SAT score (verbal 2,371; math 2,352; both math and verbal 2,351); or ACT composite score (1,825). (The numbers in parenthesis indicate the number of observations available for the potential ability measure.) Pre-TUCE III 30 is chosen since there are more observations and since the pre-TUCE III is probably a more accurate measure of economic ability when entering a principles of college economics course. Two possible pre tests, one that is 30 questions in length and one that is 33 questions in length, are available. The first 30 questions in the 33 question test are the same as in the 30 question pre-test. Although an equal number of students (8,238 students) take the 30 and 33 question pre-test, the pre-TUCE III 30 test is chosen as the ability measure since more post-test records match the pre-test ones. Matched records are needed for computations of different achievement measures. The study includes 7,312 post-TUCE III 30 observations, compared to 3,768 post-TUCE III 33 observations. Given the choice of the TUCE III 30 variable as the measure of ability, the indirect class size variables can be computed.

An index table with the course term, school code, course type and section code is created. A query is performed on the aggregate database so that records are grouped by the four indices. Once the grouping occurs, a calculation summary command is used to determine the maximum score, minimum score, standard deviation of scores and average scores of ability within each course section. These values then are written to the indexed table as separate variables. Two additional computations are made within this indexed table, the range of ability for each course section and the minimum ability score deviation from average ability in a course section. An update command is issued to write these new variables to the aggregate database for each observation. Skewness and kurtosis of ability require a greater degree of database manipulation in order to compute.

Skewness and Kurtosis are computed according to the formulas:

$$Skewness = \frac{[1/(N_j - 1)]\Sigma_{ij}(x_{ji} - \overline{X}_j)^3}{s_j^3}$$

$$Kurtosis = \frac{[1/(N_j - 1)]\Sigma_{ij}(x_{ij} - \overline{X}_j)^4}{s_j^4}$$

where,

i = The ith student observation

j = The jth course section

N_j = The number of observations in course section j

x_{ij} = The pre-TUCE III 30 score for student i in a course section j

$\overline{X}_j$ = The average pre-TUCE III 30 score for the course section j

s_j = The standard deviation of pre-TUCE III 30 Scores in course section j

These formulas must be programmed step by step since Access for Windows (Microsoft 1992) does not contain internal formulas for skewness or kurtosis.

All of the variables necessary for this research design have been appended, transformed, calculated or added to the aggregate TUCE III database.

Section 3: Sample Selection.

The 9,679 records in this database do not contain information for every variable taken from the TUCE III tests, instructor questionnaires or student questionnaires. Since providing the information was voluntary, not all instructors or students chose to fully participate. Because one criticism of previous class size investigations relates to the limited number or type of variables tested, an effort is made to maximize the number of variables while keeping the number of observations at a high level. In a first step, the essential variables needed to describe the institution, instruction or student are identified. In a second step, more variables are added depending on their impact on the sample size. What follows is a description of the sample selection process employed to extract the 2,942 student observations used in this research design from the aggregate TUCE III database.

Many variables in the sample have an expected range. Database queries use these expected ranges as search criteria when extracting the sample. Values outside of the expected range are assumed to be miscoded. Observations which include any miscoded variables are excluded from the sample. The expected range for each variable is described in the instructor questionnaires, student questionnaires and TUCE III Data File Codes (7/16/91). Variables that require a ranking, such as overall teaching effectiveness, have a limited scale. For example, the database query will include observations for which the overall teaching effectiveness variable has data in the range of 10 through 50. Table 3.1 gives the expected ranges.

In order to identify the origin of each record, course term, school identification code, course type and course section are needed for each record. All records in the database contain these identifiers. The next two essential components are the measures of achievement, which are the pre- and post-TUCE III 30 tests. A matched pre- and post-test is needed for each student in order to calculate various achievement measures. This requirement reduces the sample to 5,941 observations. Scores for the number of correctly answered RU, EA and IA questions are also included since they are a subset of the 30 question test. The four dummy variables indicating the type of institution and the dummy variable for public-private institution are also included in this sample since they are derived from the school identification code for every observation.

The next step is to identify categories of variables known to influence achievement. At this point the literature review can provide some guidance. Given the type of data, four broad categories of variables are identified. The categories comprise the following: faculty capital; course organization; student capital; and environmental factors. Faculty capital includes the instructor's personal characteristics, background experiences, research commitments, availability to students and evaluation of teaching performance. Course organization describes how the instructor organizes the course and may include components of the course grade, use of technology or use of ancillary teaching aids. Student capital includes the personal characteristics, innate ability and study effort of the student taking the course. Environmental factors include the type of institution, direct class size measures and indirect class size measures. Class size measures are included in environmental factors since neither the faculty member nor the student control the size of

Table 3.1 Expected Ranges for Variables in Sample of 2,942 Observations.

Variable	Minimum	Maximum	Variable	Minimum	Maximum
TTOT	0	30	GPA***	0	400
TRU	0	10	HSECON	0	1
TEAIA	0	20	JOB	0	1
PTOT	0	30	HRSTDY****	0	20
PRU	0	10	INTRST	10	50
PEAIA	0	20	IMPT	10	50
DOCI	0	1	ENTHU	10	50
COMPI	0	1	PREP	10	50
LIBI	0	1	STAND	10	50
TWOI	0	1	OTE	10	50
ENG	0	1	PUBPRI	0	1
CNT	0	1	DMACMIC	0	1
TCHPCT	*	99	CLASIZE	10	232
RSCHPCT	0	99	CLASQR	100	53,824
PHD	0	1	DIFSQR	**	**
YRSTCH	0	**	LOGCLAS	2.30	5.45
DSGWB	0	1	PREMAX	0	30
DHW	0	1	PREMIN	0	30
DRH	0	1	PRESTDV	0	**
COMPFNL	0	99	RANGE	0	30
PCTLEC	0	99	DEVFAVG	**	**
PCTINSR	0	99	SKEW	**	**
PCTSTUR	0	99	KURT	**	**
RESPOND	0	99	PCTPRE	50	100
GEN	0	1			

NOTES:

See Appendix B for definitions of the variables.

99 Coding indicator for 100 percent.

* Minimum is expected to be positive but non-zero.

** No minimum or maximum is expected beforehand. All variables are expected to be positive, except for skewness.

*** A 4.0 GPA scale is assumed. GPA is multiplied by 100 when coded.

**** Hours studying economics is assumed to be no greater than 20 hours per week.

a class. Table 3.2 identifies how each of the variables comprising the sample of 2,942 fit into this categorization. Using the variables of Table 3.2, the final sample of 2,942 observations can now be derived from the initial sample of 5,941.

The collection of all information on the faculty questionnaires was voluntary. Most faculty members chose to respond to all of the questions, although some did not furnish all of the requested information. As a consequence, some data are missing for certain variables. Adding variables, hence, will generally reduce the available sample size. Fortunately, the size reduction is not dramatic when faculty variables are added. Essential faculty capital variables include: English as a native language; number of years teaching; doctorate earned; percent of time spent teaching; percent of time spent researching; publication of an article in a refereed journal within the last five years; percent of time instructor spends responding to student questions; percent of time students spend responding to instructor questions, percent of time the instructor spends lecturing; number of years teaching; the post-TUCE III 30 counting toward the course grade; the percent of the final exam that is comprehensive; the three dummy variables representing use intensity for the study guide, homework and non-text reading-handouts; and identification of course recitation section. When these variables are required, the sample shrinks to 4,666 observations. Ancillary teaching aids do not include the intensity of use for computers, video or other technologically endowed methods. Inclusion of these variables when used with high intensity in the classroom would shrink the sample to zero observations. Therefore, these technology teaching aids will not be part of the 2,942 sample.

One key component of faculty capital is the evaluation of teaching performance. This information is provided on the student questionnaire. Although similar questions appear on the faculty questionnaire for measuring self-perception of teaching performance, the student assessment is the most common method used to evaluate teaching effectiveness. Teaching performance is measured by asking students to rank the preparation for class, rigor of grading standards and overall teaching effectiveness. Adding these variables reduces the sample from 4,666 observations to 3,827. The evaluation of teaching performance relates to the regular instructor and not to the recitation section discussion leader. Only 403 observations in the entire aggregate

Table 3.2 Categorization of Variables for Sample of 2,942 Observations

Dependent Variable:	Post-TUCE III
Faculty Capital:	English Native Language Number of Years Teaching Course Doctorate Earned Percent of Time Spent Teaching Percent of Time Spent Researching Article Published within Last Five Years in a Refereed Journal Regular Instructor Enthusiasm Regular Instructor Preparation Regular Instructor Grading Standards Regular Instructor Overall Teaching Effectiveness
Course Organization:	Separate Recitation Section Percent of Time Devoted to Instructor-Student Dialog (i.e., Respond) Does Post-TUCE III Count Toward Final Grade Percent of Final which is Comprehensive Use Intensity of Non-text Readings and Handouts Use Intensity of Study Guide or Workbook Use Intensity of Homework
Student Capital:	Student Gender Interest in Economics Importance of Economics Cumulative GPA Completed a High School Economics Course Working in a Job Hours a Week Studying Economics
Environmental Factors:	Institutional Type Public or Private Institution Type of Course Direct Class Size Variables Indirect Class Size Variables

database have indicated separate recitation sections. Since the number is so small, the evaluation of teaching performance of the discussion leader is not included in the 2,942 sample.

The last major category of educational inputs is student capital. This category of information is provided by the student questionnaires. Since filling out the questionnaires was voluntary, some students elected not to give some information. Consequently, larger gaps of missing information for student capital inputs than for faculty capital inputs are apparent. Essential student inputs include: gender; GPA; completion of a high school economics course; working at job; hours studying economics per week; interest in economics; and importance of economics. Requiring these inputs reduces the number of observations to 2,973.

No other variables are added since the sample size would be dramatically reduced. For example, when the following student inputs are sequentially added, the sample size is reduced to the number of observations indicated in parenthesis: actual course grade (2,701); ACT composite score (954); semester or trimester hours completed prior to the present term (730); semester or trimester hours of college calculus completed prior to the present term (660); and percent learned from the regular instructor (643). If SAT scores are converted into comparable ACT scores and if quarter hours are converted into semester hours, the sample size is reduced to approximately 1,000 rather than 643. No other variables are added to the 2,973 observation sample because the increase in number of explanatory variables does not appear to be worth the loss in observations.

The last step in selecting the sample from the aggregate TUCE database is to select only those classes which achieved more than a fifty percent participation rate on the pre-TUCE III 30 question test. Since many students elected not to take the pre-test or since instructors did not encourage full-participation from the class to take the pre-test, the number of students enrolled in the course does not always equal the number of students taking the pre-TUCE III 30. The percent of students taking the pre-test given the number of students enrolled in the course has ranged from a high of 100 percent participation to a low of 36 percent participation. Observations which came from classes that have less than 50 percent participation are not selected for the 2,942 sample. Students completing the pre-test for these lower participation rate classes are probably more able students. Consequently, these classes may oversample better students leading to some bias. When this final restriction is made, the sample of 2,973 observations shrinks to the 2,942 observation sample used in this study.

Additional variables which are added to the aggregate database but do not lead to a reduction in the number of sample observations are the indirect class size influence variables. These indirect variables for each class section are: maximum ability score; minimum ability score; standard deviation of ability; average ability; range of ability; the deviation of the minimum ability score from the average of ability; skewness of ability; and kurtosis of ability. Since each of these variables is the same for every class section completing the pre-TUCE III 30 tests, every student in the class has the same values for these variables. Because one of the first criteria is matched pre- and post-tests, the indirect class size variable restrictions are already included in the sample at this point and the number of observations is not reduced below 2,942 observations in the sample.

Table 3.3 contains descriptive statistics on the sample.

Section 4: Research Methodology.

Since the sample is selected, the research method used to investigate the influence of class size on student achievement in principles of college economics can be discussed. The methodology of this investigation attempts to eliminate, or at least reduce, the criticisms of previous class size investigations, which are identified in section 1.2.2 of chapter two.

Specifically, this research design improves upon its predecessors in the following ways: a wide range of class sizes is considered, specifically classes with 10 through 232 students; the sample size of 2,942 observations is larger than most previous inquiries; a national cross-sectional sample is used rather than a sample from one university; the economics course content can be identified as either principles of macroeconomics or

Table 3.3 Descriptive Statistics for Variables in the Sample of 2,942 Observations.

Variable	Mean	Std. Dev.	Minimum	Maximum
TTOT	14.86	5.46	3.00	30.00
TRU	5.18	2.14	0.00	10.00
TEAIA	9.66	3.94	0.00	20.00
PTOT	10.16	3.68	1.00	27.00
PRU	3.35	1.68	0.00	10.00
PEAIA	6.83	2.76	0.00	18.00
DOCI	0.25	0.43	0.00	1.00
COMPI	0.51	0.50	0.00	1.00
LIBI	0.12	0.33	0.00	1.00
TWOI	0.12	0.32	0.00	1.00
ENG	0.94	0.24	0.00	1.00
CNT	0.58	0.49	0.00	1.00
TCHPCT	58.53	22.18	5.00	99.00
RSCHPCT	23.16	20.73	0.00	80.00
PHD	0.67	0.47	0.00	1.00
YRSTCH	12.47	8.23	1.00	34.00
DSGWB	0.15	0.36	0.00	1.00
DHW	0.29	0.45	0.00	1.00
DRH	0.16	0.36	0.00	1.00
COMPFNL	57.62	40.24	0.00	99.00
PCTLEC	78.98	13.42	1.00	99.00
PCTINSR	11.26	6.54	0.00	33.00
PCTSTUR	8.63	7.36	0.00	33.00
RESPOND	0.44	2.13	0.00	30.00
GEN	0.57	0.50	0.00	1.00
GPA	290.02	60.71	0.00	400.00
HSECON	0.42	0.49	0.00	1.00
JOB	0.58	0.49	0.00	1.00
HRSTDY	3.27	2.29	0.00	20.00
INTRST	33.64	9.41	10.00	50.00
IMPT	37.62	7.81	10.00	50.00
ENTHU	42.33	7.16	10.00	50.00
PREP	43.71	6.51	10.00	50.00
STAND	39.60	7.44	10.00	50.00
OTE	40.63	8.13	10.00	50.00
PUBPRI	0.21	0.41	0.00	1.00
DMACMIC	0.51	0.50	0.00	1.00
CLASIZE	71.39	46.15	10.00	232.00
CLASQR	7,225.00	10,467.00	100.00	53,820.00
DIFSQR	7,153.60	10,423.00	90.00	53,590.00
LOGCLAS	4.09	0.59	2.30	5.45
PREMAX	18.49	3.22	11.00	27.00
PREMIN	3.91	1.59	1.00	9.00
PRESTDV	3.18	0.55	2.18	5.30
RANGE	14.58	2.94	7.00	23.00

Table 3.3 Continued.

Variable	Mean	Std. Dev.	Minimum	Maximum
DEVFAVG	6.07	1.33	3.22	9.54
SKEW	0.46	0.40	-0.51	2.07
KURT	3.23	1.10	1.45	8.57
PCTPRE	88.73	8.37	53.85	100.00

Notes: See Appendix B for definitions of the variables.

principles of microeconomics; the influence of recitation or discussion groups on large classes can be measured; a common, easily accessible and universally accepted definition of achievement is used, the TUCE III exams; the primary question of this study is the influence of class size on student achievement; an input-output model is used to assess student achievement outcomes; class size influence is theoretically modeled and tested; and educational inputs are selected ex ante from input categories known to influence achievement.

The research methodology incorporates these improvements into the following six steps:

1. The form of the production function is identified along with the expected signs of the educational inputs. The influence of the input variables is identified a priori rather than a posteriori. This methodology is preferred to totally ad hoc procedures.

2. An ordinary least squares (OLS) regression is run with all independent variables, except class size. The dependent variable is defined as absolute improvement since previous research efforts have generally used this measure. The results of the OLS regression will be used to determine if certain variables should be transformed or dropped.

3. After dropping or transforming certain variables, another OLS regression is run for each of the dependent variables of achievement, which are absolute achievement, absolute improvement, percentage improvement and gap closing. Class size is not included as an independent variable. Standard measures of model adequacy are used to

determine which of the dependent variables is most appropriate. In case the absolute improvement model with the pre-TUCE III 30 as an independent variable appears to be best, the discrete nature of the dependent variable would suggest choosing an estimator other than OLS. Poisson regressions are run and then compared to the OLS results to determine whether the more complex Poisson model should be used. The determinants of student achievement are identified.

4. The direct and indirect variables for class size are added separately to the preferred achievement model. Regressions are run and the results discussed. Regressions also are run with various class size ranges to determine whether there is an aggregation problem. Seven class size ranges for which the number of observations is approximately equal are identified. Dummy variables are created for each class size range, then inserted into the preferred achievement model. The influence of class size on student achievement is identified.

5. Using the most efficient class size model identified in the previous step, RU, EA and IA questions are used as a dependent variable. This time, instead of using matched TUCE III overall test scores, matched question categories are used to define student achievement for different levels of cognitive ability.

6. Policy recommendations are made regarding class size changes in principles of college economics. The recommendations take into consideration the impact that larger or smaller classes may exert on college budgets. The recommended class size policies consider the potential consequences that class size may have on student achievement. Suggestions are made to neutralize any unwanted consequences on either achievement or budgets.

Chapter IV

Empirical Analysis

The purpose of this chapter is to analyze several multivariate regressions in order to determine the influence of class size on student achievement in principles of college economics. The sample of 2,942 observations selected from the aggregate TUCE III database is used to supply the data for the production function input and output variables.

The chapter is organized into six sections, each identifying either a theoretical or empirical component needed to build the final model used to test the influence of class size. Section 1 describes the production function, educational inputs and the expected sign of the inputs. Section 2 identifies the determinants of achievement. Section 3 selects the achievement model used to test the influence of class size on achievement. Section 4 presents evidence regarding the influence of class size on absolute improvement. Section 5 discusses the influence of class size on various cognitive skills. Section 5 summarizes the empirical findings.

Section 1: Production Function Form, Education Inputs and Expected Signs of the Inputs.

The empirical analysis will use the most common form of an educational production function, the one by Becker (1983B). This function is preferred since it is simple and commonly accepted, making comparisons to previous studies and replication easier. The theoretical model described in section 2.3 of chapter two follows the form:

$$Y_i = B_0 + \sum_k B_k X_{ki} + e_i$$

Where, the X_k's in the equation represent K independent educational input regressors and the subscript 'i' represents an individual observation. The input variables used in this study are grouped into four broad categories identified in chapter three: faculty capital; course organization; student capital; and environmental factors. Table 3.2 of chapter three categorizes the input variables. The method describing the selection of the variables is included in section three of chapter three.

The expected signs of the coefficients B_k are identified in Table 4.1. The literature review provided a discussion of numerous education inputs that are known to be influential. Fifteen of the 42 variables identified in the literature review are included in this study. Column two in Table 4.1 identifies the direction of their influence according to the literature review. The influence of the 27 remaining input variables has not been evaluated in the previous literature. One goal of the TUCE III database is to provide information on numerous factors that may influence achievement, but whose influence has not been investigated before. Variables that have not been evaluated in professional journals are denoted by a question mark or an asterisk in column one of Table 4.1. Although not reviewed in the literature, intuition and professional judgment would suggest certain expectations on the signs of the 27 remaining variables. These intuitive expectations are identified in the third column of Table 4.1.

An explanation will clarify why the intuitive sign expectations are either positive or negative. The two-year college (TWOI) is the base institution and thus is excluded from the regressions. Compared to TWOI, universities granting doctorates (DOCI), comprehensive universities (COMPI) and liberal arts colleges (LIBI) may have higher achievement scores than TWOI because institutions other than TWOI may have more resources to assist in the teaching-learning process and may admit better students.

Table 4.1 Expected Signs of Independent Variable Coefficients.

Variable	Sign Identified in Literature Review	Intuition	Variable	Sign Identified in Literature Review	Intuition
DOCI	?	+	HRSTDY	+	
COMPI	?	+	INTRST	+	
LIBI	?	+	IMPT	+	
PTOT	+		ENTHU	+	
SEPRSX	?	+	PREP	+	
ENG	+		STAND	?	+/-
CNT	?	+	OTE	+	
TCHPCT	?	+	PUBPRI	?	+
RSCHPCT	?	+/-	DMACMIC	?	+
ART	?	+/-	CLASIZE	+/-	
PHD	?	+	CLASQR	- *	
YRSTCH	+		LOGCLAS	- *	
DSGWB	?	+	DIFSQR	- *	
DHW	?	+	PREMAX	- *	
DRH	+		PREMIN	- *	
COMPFNL	?	+	PRESTDV	- *	
RESPOND	?	+	PREAVG	- *	
GEN	+		RANGE	- *	
GPA	+		DEVFAVG	- *	
HSECON	+/-		SKEW	- *	
JOB	-		KURT	- *	

NOTES:

\+ Positive influence on achievement
\- Negative influence on achievement
+/- Inconclusive
? No expectation
* Theoretical expectation

Student achievement may rise because better students may respond more favorably to teaching efforts and may be more motivated as well as more coachable. Separate recitation sections or small group discussions (SEPRSX) are designed to provide more direct communication between the student and the teacher. SEPRSX should improve achievement since questions regarding learning objectives may be clarified. Counting the post-TUCE III (CNT) toward the final course grade should raise achievement since counting the exam will most likely motivate students to care about their performance on the test. The percent of time a professor spends teaching (TCHPCT) is expected to positively influence achievement, while the percent of time devoted to research

(RSCHPCT) is expected to have the opposite effect on achievement. Instructors who spend a greater percent of their time teaching should improve their craft. A rising TCHPCT should result in improved achievement. Spending more time in research results in an opportunity cost of less time spent teaching. Rising RSCHPCT should result in declining achievement, although some professionals believe that research encourages academic alertness and therefore better teaching. The influence of RSCHPCT may be inconclusive. Making the effort to have an article published in a peer reviewed journal at least once in the last five years (ART) could result in either a positive or a negative effect similar to RSCHPCT. Therefore, the influence of ART may be inconclusive. Having an earned doctorate in economics (PHD) should positively influence achievement since the instructor is better prepared academically. The intensive use of the study guide or workbook (DSGWB) and the intensive use of homework (DHW) should positively influence achievement since practice, repetition and correction of mistakes provides students with feedback. Comprehensive final exams act as an extrinsic motivator. As the comprehensiveness of the exam increases, so does the motivation to retain a greater amount of economic knowledge. Consequently, COMPFNL is expected to have a positive influence on achievement. RESPOND is a ratio that measures the percent class time devoted to instructor-student interaction relative to the percent of class time that the instructor spends lecturing. RESPOND should positively influence achievement since greater amounts of feedback allow the instruction-learning process to become more focused. Private schools (PUBPRI) are more likely to positively influence achievement since private schools may tend to admit better students, have more resources available per student and provide for more personal interaction with faculty members. PUBPRI is coded 1 for private schools and 0 for public. The rigor of a professor's grading standards (STAND) may positively or negatively influence achievement since the rigor of grading standards could either motivate or de-motivate students. Principles of microeconomics

(DMACMIC) may be more structured than principles of macroeconomics due to the theoretical disarray in macroeconomics. Consequently, students should be expected to perform better in micro than in macro. DMACMIC is coded 1 for principles of microeconomics and 0 for principles of macroeconomics. DMACMIC should be positive. Theoretically, both the direct and indirect class size model variables expect larger classes to exert a negative influence on achievement. Since these variables have not been tested in economics education research, the expectation is based on theory rather than on intuition or on previous empirical results.

Section 2: Identification of the Determinants of Student Achievement.

This section identifies the determinants of student achievement in principles of college economics. The production function is constructed using the absolute improvement definition of student performance, which is a typical research definition of achievement. A preferred way to define the absolute improvement measure is to take the difference between a post-test and a pre-test. However, an alternative is to use the post-test as the dependent variable and to include the pre-test as an independent variable. Hanushek (1986) believes that scale effects will be picked up in this way. The latter method is employed in this study for two reasons: (1) the absolute achievement and absolute improvement measures are commonly used to identify achievement determinants; and (2) Hanushek (1986) believes this model to have more explanatory power. Table 4.2 presents the first models of the determinants of achievement.

A potential concern is the cross-influence of one independent variable on another. Slight multicollinearity is confirmed by a condition number of 44.86, which is above the critical value of 30 (Belsley et al., 1980). A correlation matrix is used to help determine

Table 4.2 The Determinants of Student Achievement in Principles of College Economics Using TTOT as the Dependent Variable.

Variable	ONE TTOT Estimated Coefficient	TWO TTOT Estimated Coefficient
Constant	-3.8798	-4.9823
	(-4.3)	(-6.3)
DOCI	0.9533	0.9269
	(3.0)	(2.9)
COMPI	-0.5402	-0.5129
	(-1.7)	(-1.6)
LIBI	0.4809	0.6395
	(1.2)	(1.6)
PTOT	0.6075	0.611
	(29.7)	(29.6)
SEPRSX	0.2218	0.3402
	(0.6)	(0.9)
ENG	0.4618	0.5093
	(1.4)	(1.5)
CNT	0.9635	1.0584
	(4.9)	(5.6)
TCHPCT	0.0069	
	(1.5)	
RSCHPCT	0.0087	0.0029
	(1.6)	(0.6)
ART	-0.3509	
	(-1.6)	
PHD	1.2812	1.0239
	(4.7)	(4.2)
YRSTCH	-0.0032	-0.0091
	-(0.3)	-(0.8)
DSGWB	-0.0239	0.2024
	-(0.1)	(0.8)
DHW	0.3776	0.3846
	(2.1)	(2.2)
DRH	0.6313	0.6174
	(2.9)	(2.9)
COMPFNL	0.0122	0.0114
	(6.1)	(5.7)
RESPOND	-0.0033	0.0015
	(-0.1)	(0.1)
GEN	0.9838	1.1191
	(6.4)	(7.4)
GPA	0.0247	0.0248
	(17.3)	(17.3)

Table 4.2 Continued

Variable	ONE TTOT Estimated Coefficient	TWO TTOT Estimated Coefficient
HSECON	-0.0210	-0.023
	(-0.1)	(-0.2)
JOB	0.0934	0.1311
	(0.6)	-(0.8)
HRSTDY	-0.0673	-0.0656
	(-2.1)	(-2.0)
INTRST	0.0735	
	(7.3)	
IMPT	0.0033	
	(0.3)	
INTIPT		0.0437
		(8.6)
ENTHU	0.0208	
	(1.5)	
PREP	-0.0362	
	(-2.4)	
ENTPRP		-0.0041
		(-0.6)
STAND	-0.0385	
	(-3.5)	
OTE	0.0229	
	(1.7)	
PUBPRI	1.9042	1.8708
	(7.4)	(7.3)
DMACMIC	0.2212	0.2189
	(1.4)	(1.4)
Observations	2,942	2,942
R-squared	0.50	0.49
Adjusted R-squared	0.49	0.48
Regression F-test	95.45	116.26
[DF]	[30; 2,911]	[24; 2,917]
Prob of F-test	0.00	0.00
Breusch-Pagan	49.65	46.82
[DF]	[30]	[24]

Notes: The numbers within the parentheses are the t-ratios derived from White's (1980) heteroscedasticity corrected variance-covariance matrix. TWO above the second column represents regression ONE after deleting or transforming certain variables.

the variables that may be collinear. Table 4.3 lists variables that have correlations of 0.50 and greater, and it can help identify which variables may need to be transformed or deleted.

Table 4.3 Correlations of 0.50 and Greater for First Regression

Correlated Variables	Correlation Coefficient
DOCI - COMPI	-0.59
PHD - TWOI	-0.52
RSCHPCT - TCHPCT	-0.62
PHD - ART	0.50
IMPT-INTRST	0.59
PREP - ENTHU	0.62
PUBPRI - LIBI	0.56
OTE-ENTHU	0.63
OTE - PREP	0.61

The dummy variables identifying doctoral (DOCI), comprehensive (COMPI) and liberal arts (LIBI) institutions have to be interpreted in reference to the two-year (TWOI) variable, which serves as the base institution. No transformation is performed on these variables. The negative correlation between PHD and TWOI is not a problem because only PHD appears in the regression, while TWOI is omitted.

The negative correlation between RSCHPCT and TCHPCT is to be expected: the more time spent on research, the less time will be available for teaching. Since there are several other variables evaluating different aspects of teaching, TCHPCT is dropped. The positive correlation between PHD and ART is anticipated. ART is dropped since RSCHPCT may better represent the time requirements for the publication process. PHD is kept because of the assumption that faculty members with terminal degrees are more likely to be engaged in some form of research that could lead to professional publications. Keeping ART with PHD may duplicate information on research efforts. ART does not indicate the number of peer reviewed publications in excess of one during the last five years. It, therefore, may be a weak indicator of faculty productivity.

Student interest in economics (INTRST) and student judgment as to the importance of economics (IMPT) may be contributing the same information to the explanation of achievement. The difference between INTRST and IMPT is not explained in the student questionnaires. The correlation of 0.59 may indicate that students viewed both categories similarly. This problem is addressed by summing IMPT and INTRST.

Previous research inquiries have found evidence that instructor enthusiasm (ENTHU) and preparation (PREP) positively influence achievement. The correlation of 0.62 may indicate that instructors who are more prepared for class have reason to be more enthusiastic. Here again this problem is dealt with by adding PREP and ENTHU together. Students were asked to assess the overall teaching effectiveness (OTE) as a composite indicator of the instructor in the course. OTE is the last question on the student questionnaire in the section evaluating teaching performance. Since it may duplicate information being captured by other variables already in the regression that assess specific aspects of teaching performance or teaching methodology, OTE is dropped. The positive correlation between private colleges (PUBPRI) and liberal arts institutions (LIBI) is to be expected since liberal arts institutions are more likely to be private. PUBPRI and LIBI are both retained because dropping them from the regression would result in a loss of significant information.

Regression two in Table 4.2 incorporates these variable transformations and deletions. Regression two has reduced multicollinearity as confirmed by a condition number of 26.52, which is below the critical value of 30. Regression two identifies several determinants of student achievement. However, the adequacy of the absolute improvement model should be compared to the four other forms that can be used to model student achievement.

Students are also asked to evaluate the grading standards of the instructor (STAND). STAND is probably the most subjective variable in the student assessment of teaching performance. Since STAND is a more subjective standard, both in terms of student judgment and institutional expectations, the variable is dropped from the regression.

Section 3: Selection of the Achievement Model.

This section will identify a preferred model of student achievement in principles of college economics. The determinants of achievement in Section 2 have been identified after correcting for influential input correlations. The determinants from regression two in Table 4.2 are used in this section as the input variables for each of the five models of student achievement. The adequacy of the models will be compared, and a preferred model will be selected.

Hanushek (1986) identifies four ways to construct the dependent variable: absolute achievement (TTOT); absolute improvement (VA); percentage improvement (PCTVA); and gap closing (GAPC). A variation of the absolute improvement model is to include the pre-test as an independent variable (T/PTOT), a variation which is used for regressions one and two in Table 4.2 of the preceding section. The various achievement models are discussed in detail in section two of chapter two. Table C.2 in Appendix C compares the five separate achievement definitions.

The five models appear to predict similar signs and significance for most of the determinants of achievement. Two exceptions are found in the PCTVA model. Both the RESPOND and DMACMIC variables are negative and significant. In the alternative models these two variables are positive and insignificant in the case of RESPOND and positive and significant in the case of DMACMIC. The alternating signs and significance in these models may be due to the dependent variable measuring different outcomes. The PCTVA model measures the difficulty that students of lesser ability have in trying to improve their scores while the GAPC model measures the difficulty that more competent students have in trying to improve their scores. VA measures incremental learning and T/PTOT measures gains in achievement when scale effects are included in the regression. The five achievement models in Table C.2 may not be directly comparable since different outcomes are measured from one model to the next. The most complicated definition of achievement is GAPC, which predicts signs and significance of the coefficients that are similar to the other models. However, policy implications for improvement in gap closing scores may be more difficult to interpret.

The absolute improvement model with the pre-TUCE III 30 as an independent variable has the best overall model adequacy results as measured by: r-squared; adjusted r-squared; significant F-statistic; number of correctly predicted coefficient signs; number of significant t-statistics; simplicity of model; ease of interpretation; comparability to previous research results; and practicality for developing administrative or instructional policies. As the dependent variable the post-TUCE III 30 represents count data.

Ordinary Least Squares (OLS) may be inefficient for count data. A Poisson estimator may be more appropriate. OLS is less efficient because the estimator does not fully capture the characteristics of the dependent variable. The post-TUCE III 30 score only takes on integer values, which would make Poisson more efficient. The basic formulation of the Poisson regression model for a discrete random variable Y, and observed frequencies, y_i, i=1,...,N, where $y_i >= 0$ and regressors x_i (Greene 1992, 539):

$$prob(Y = y_i) = e^{-\lambda_i} \lambda_i^{y_i} / y_i!, \quad y = 0, 1, \ldots,$$

where,

$$\ln \lambda_i = \beta' x_i$$

and where λ_i is the mean and variance of y_i

When the Poisson model is used, tests for overdispersion should be made in order to test whether the mean of y is actually the same as its variance as assumed by the model. The t-ratios used to test for overdispersion are based on the tests identified by Cameron and Trivedi (1990). The t-ratios are 2.694 and 1.415 respectively. Only the first t-ratio is statistically significant, the five percent level. Consequently, there may be a slight problem with overdispersion. The assumption of the mean being equal to the variance in the Poisson model may be violated. This problem may be addressed by using a negative binomial regression model, which allows the variance to differ from the mean. The model differs from Poisson because an error term is included in the determining equation for λ_i.

$$\ln \lambda_i = \beta' x_i + \varepsilon$$

When a negative binomial regression is run on the absolute improvement model, the t-statistic for the additional parameter of the negative binomial relative to the Poisson is not significant. Hence, overdispersion does not appear to be a problem for the Poisson model.

The coefficients from OLS and the coefficients from Poisson are not directly comparable since Poisson is a non-linear regression. However, marginal effects from the Poisson model can be computed and directly compared to the OLS results. The marginal effects for the Poisson coefficients are calculated in the following way (Greene, 708):

$$E[y_i | x_i] = Var[y_i | x_i] = \lambda_i = e^{\beta' x_i}$$

so,

$$\frac{\partial E[y|x]}{\partial x} = \lambda_i \beta$$

Table 4.4 compares the marginal effects of the OLS with those of the Poisson regression. Because no significant difference between the two estimation methods is apparent, the signs and significance of the regression variables are essentially the same. The results from both OLS and Poisson show that most of the signs of the statistically significant independent variables conform to a priori assumptions regarding the signs identified in Table 4.1.

HRSTDY is the only exception of a statistically significant independent variable where the sign is opposite of the expectation. HRSTDY is expected to be positive, indicating that the more hours per week a student spends in the study of economics, the greater the expected gain in achievement. HRSTDY may be negative because the hours spent studying during the week were not hours spent studying for the TUCE III, but hours spent studying for specific learning objectives identified by the instructor. The learning objectives may not coincide with the questions on TUCE III. Consequently, a negative relationship may exist between HRSTDY and TUCE III achievement. If instructors were allowed to teach to the TUCE III during the semester, a positive correlation could have resulted. HRSTDY may show a direct relationship when a course grade is used as the achievement measure. Instructors teach to specific learning objectives of the course grade and motivate students to gear their study efforts for achievement in these course objectives. The TUCE III, by contrast, measures overall aptitude in economics and not how well a student has mastered the course objectives emphasized by particular instructors. The negative relationship of HRSTDY also may be explained by the belief that less able students might have to study more hours than better students. The more able student possibly receives a larger marginal increment to achievement from pre- to post-exam with fewer hours of study relative to the less able student while the less able

Table 4.4 Marginal Effects of OLS and Poisson Regressions on Absolute Improvement

Variable	OLS	POISSON
Constant	-4.9823	19.4990
	(-6.3)	(24.2)
DOCI	0.9269	1.0008
	(2.9)	(3.2)
COMPI	-0.5129	-0.2801
	(-1.6)	(-0.9)
LIBI	0.6395	0.6886
	(1.6)	(1.8)
PTOT	0.611	0.5191
	(29.6)	(26.7)
SEPRSX	0.3402	0.3656
	(0.9)	(1.0)
ENG	0.5093	0.7422
	(1.5)	(2.2)
CNT	1.0584	0.9582
	(5.6)	(5.2)
RSCHPCT	0.0029	0.0001
	(0.6)	(0.0)
PHD	1.0239	0.9638
	(4.2)	(4.0)
YRSTCH	-0.0091	-0.0123
	(-0.8)	(-1.2)
DSGWB	0.2024	0.2992
	(0.8)	(1.4)
DHW	0.3846	0.4333
	(2.2)	(2.7)
DRH	0.6174	0.0301
	(2.9)	(2.7)
COMPFNL	0.0114	0.0112
	(5.7)	(5.8)
RESPOND	0.0015	0.0059
	(0.1)	(0.2)
GEN	1.1191	1.1142
	(7.4)	(7.6)
GPA	0.0248	0.0254
	(17.3)	(20.1)
HSECON	-0.023	-0.0454
	(-0.2)	(-0.3)
JOB	0.1311	0.1810
	(0.8)	(1.2)
HRSTDY	-0.0656	-0.0718
	(-2.0)	(-2.3)

Table 4.4 Continued.

Variable	OLS	POISSON
INTIPT	0.0437	0.0438
	(8.6)	(8.7)
ENTPRP	-0.0041	-0.0066
	(-0.6)	(-1.1)
PUBPRI	1.8708	1.5724
	(7.3)	(6.9)
DMACMIC	0.2189	0.2174
	(1.4)	(1.4)
Observations.	2,942	2,942
R-squared	0.49	
Adjusted R-squared	0.48	
Regression F-test [DF]	116.26 [24; 2,917]	
Prob of F-test	0.00	
Breusch-Pagan [DF]	46.82 [24]	
Correctly Predicted Coefficient Signs	17	
Statistically Significant Coefficients	14	
Correct Sign & Statistically Significant	11	
Chi-squared		3,158.8
G-squared		3,213.4

Notes: Both the OLS and Poisson model use the post-TUCE 30 as the dependent variable. The numbers within the parentheses are the t-ratios. The OLS t-ratios are derived from White's (1980) heteroscedasticity corrected variance-covariance matrix.

student possibly receives a smaller marginal increment to achievement from the pre-to post-exam with more hours of study relative to the better student.

The statistically significant coefficients in the OLS and Poisson regressions in Table 4.4 show that average student achievement in principles of economics can be improved when the post-TUCE III 30 counts toward the final grade, when the instructor has an earned doctorate in economics, when homework is used intensively, when non-text readings are used intensively and when comprehensive final exams are given. In addition, average achievement is higher for students attending a doctoral institution and for students attending a private school. The results confirm previous empirical work by showing that mean achievement is positively influenced when students enter the class with a higher stock of economic knowledge as measured by the pre-TUCE III 30, when the instructor's native language is English, when the student is a male, when a student enters the class with a higher cumulative GPA, when the student enters the class more interested in economics and when the student believes in the importance of the course.

The absolute improvement achievement model with the post-TUCE III 30 as an independent variable is the preferred model to investigate the influence of class size on student achievement principles of economics. The preferred model is estimated using OLS with the t-ratios being derived from White's (1980) heteroscedasticity corrected variance-covariance matrix. The OLS method is preferred over Poisson for its simplicity, replicability, comparability and ease of interpretation.

Section 4: The Influence of Class Size on Absolute Improvement.

In this section the influence of class size on absolute improvement is tested. Both direct and indirect class size influence models are compared. Class size ranges are used with the instructional overhead model in order to reduce the problem of aggregation.

4.1. The Influence of Direct and Indirect Class Size Variables.

The direct and indirect models of class size influence were discussed in section three of chapter two. The independent variables representing class size influence are added to the variables of the absolute improvement model. PREMIN is not included in the instructional pacing model since the choice of the student ability variable, pre-TUCE III 30, results in perfect correlation with DEVFAVG and PREAVG (PREMIN equals DEVFAVG minus PREAVG). PREMIN, not DEVFAVG or PREAVG, is dropped since PREMIN is included as a variable in the ability grouping model. Table C.3 compares the direct models, and Table C.4 compares the indirect models. Both tables are in Appendix C.

None of the indirect or direct class size influence models predicts as expected. All of the coefficients are positive except for KURT. Also, all of the coefficients that are positive, except for DEVFAVG, are significant. If KURT is a possible reason for decreased achievement, then KURT should be negative when regressed as the only class size proxy. However, KURT may be correlated with other variables in the regression. Table 4.5 lists class size influence variables which have a correlation of 0.40 and greater. Due to the large number of correlated variables, each of the direct and indirect class size variables should be tested alone in order to isolate the true influence of the variable. This procedure would not be at odds with Mitchell, et. al., (1989) since the theoretical equations presented in section three of chapter two merely attempt to categorize the class size variables according to type of influence. The equations can be rewritten so that achievement is a function of only one class size variable rather than several.

Table 4.5 Correlated Class Size Influence Variables of 0.40 and Greater

Correlated Variables	Correlation	Correlated Variables	Correlation
CLASIZE-CLASQR	0.9543	PREMAX-KURT	0.4178
CLASIZE-LOGCLAS	0.9410	PREMIN-PREAVG	0.7011
CLASQR-LOGCLAS	0.8049	PRESTDV-PREAVG	0.7052
RANGE-LOGCLAS	0.4039	PRESTDV-RANGE	0.6613
PREMAX-PREMIN	0.4176	PRESTDV-DEVFAVG	0.6121
PREMAX-PRESTDV	0.7473	PREAVG-RANGE	0.4235
PREMAX-PREAVG	0.7326	PREAVG-DEVFAVG	0.5255
PREMAX-RANGE	0.8703	RANGE-DEVFAVG	0.6770
PREMAX-DEVFAVG	0.4995	RANGE-KURT	0.4838
PREMAX-SKEW	0.4419	SKEW-KURT	0.7880

Table C.5 in Appendix C compares the regression results of the absolute improvement model when only one of the direct or indirect class size variables is used as the only class size regressor. The regressions are run in order to eliminate the potential cross-influence of the class size variables. For example, when KURT is regressed without the other indirect class size variables, it turns out to be both positive and statistically significant. In fact, all indirect and direct class size variables are positive, and all are significant except for SKEW. The logic of these results would imply that principles of economics should be taught in concert halls or sports stadiums--the greater the number of students in the class, the greater the improvement in achievement. These results, however, are not logical, and five possible reasons for the empirical results to be opposite to the expectation can be offered.

First, CLASIZE is the proxy variable derived from the number of observations in each course section file in the raw TUCE III data. The number of observations in a file include students who only took the pre-TUCE III and some who only took the post-TUCE III. Observations may not have contained matched pre and post-exams. The number of observations in a course section file does not account for students who withdrew from the course or completed less than the number of required questions on the exam. Therefore, the number of observations in a course file is also a proxy for enrollment.

Using a proxy for course enrollment may have biased the data toward the selection of better-performing students. Since students answering less than the required number of questions on the pre-TUCE III or post-TUCE III are excluded from the database, the more conscientious and academically able students may be more likely to complete the exams. The information contained in the database may have been collected from achievement oriented students. Also, the pre-TUCE III was instructed to be given on the first day of the course. Every student enrolled in a course may not have shown up for the first class meeting. Less conscientious students may begin attending on the second or third class periods when they perceive the preliminaries of the course to be finished. Less conscientious students may not have planned ahead with their course schedules and consequently added the course to their schedules after the first course meeting. Further, some instructors may count the post-TUCE III toward the computation of a student's final grade. Students in courses that count the exam may have an incentive to perform better on the exam which, in turn, would bias the observations toward greater achievement. Consequently, better students will tend to do better on exams, regardless of class size. The empirical results may bear this out with the coefficients of the various indirect and direct class size models being positive.

Second, 37 percent of the 2,942 observations in the sample are enrolled in classes of 90 students and greater. Only 10% of the 2,942 observations in the sample are enrolled in classes of 30 students and smaller. The number of students for each class size can be found in Table C.2 of Appendix C. Since all students will improve from the pre-TUCE III to the post-TUCE III, having more observations of students in the upper class sizes means that a bigger group of students in larger classes are improving their achievement. Given this fact, the empirical results incorrectly show that larger classes seem to contribute more toward achievement, when the truth is that more students from larger classes are sampled and that these students will see guaranteed gains in achievement due to the nature of the measurement device. When this is compounded with an oversampling of more conscientious students from point number one, achievement could very well increase with class size.

Third, each of the direct class size coefficients are positive because of the reason mentioned in points one and two. ClASIZE, CLASQR and LOGCLAS are all positive and significant because there are more observations in the upper class size ranges who are better students. CLASQR and LOGCLAS depend directly on CLASIZE, which is the proxy for the number of students enrolled in a class. Although CLASIZE is negative in the interaction model, theory predicts the difference between CLASQR and CLASIZE to be negative. In other words, CLASQR is expected to be negative and CLASIZE is expected to be positive, the opposite of the empirical results.

Fourth, most of the indirect class size coefficients are positive because they are a function of class size. The reasons mentioned in points one and two cause most of the indirect variables to mirror the influence of the direct variables. The oversampling of large classes and the bias toward better students seem to cause the indirect class size influences to be positive. Since the influence of class size is less obvious in the indirect models, a brief explanation for RANGE, PRESTDV, DEVFAVG, PREMAX and PREMIN follows. The influence of SKEW and KURT is discussed in point five.

As class size increases RANGE and PRESTDV is expected to increase. When RANGE and PRESTDV increases, more students of lower ability have a better chance of improving their score. Increasing achievement from a score of five on the pre-TUCE III to a score of 10 on the post-TUCE III is easier than trying to achieve a five point increase from 20 to 25. Less able students may be able to see greater gains in their score from pre- to post-TUCE III than more able students. Lower ability students can improve more since moving from a lower to higher score is easier than moving from high to even higher scores. If large class sizes are oversampled, then there are greater RANGES in bigger classes with larger PRESTDV. The result is that a larger number of students are guaranteed, from a statistical perspective, to improve from the pre to post-TUCE III.

If RANGE increases as class size increases, then the PREAVG should fall. When the PREAVG falls there are more students of lower ability who have a better chance of improving their score. If large classes are oversampled, then a lower PREAVG of the larger classes will result in a larger number of students being guaranteed an improved score. The nature of the test instrument means that all students will improve from pre- to post-exam. Lower ability students may have an easier time improving if the TUCE III is used for achievement measurement. Also, a greater range among larger classes suggests that the upper end of the range contains higher pre-test scores than in smaller classes. This would result in a greater number of students in larger classes who are naturally inclined to excel from the pre to post-exam.

Larger classes are more likely to have larger ranges of ability in the class. Larger classes would have lower minimum ability (PREMIN) scores and higher maximum ability scores (PREMAX). Intuition and experience would suggest that one poor or one gifted student would not significantly influence the achievement of all other students in principles of economics. However, minimum and maximum ability scores can be linked to RANGE and since RANGE is associated with class size, an oversampling of larger classes would result in PREMIN and PREMAX being positively correlated with achievement.

Table 4.6 generally supports the assertions made in point four. Oversampled large classes result in larger ranges, lower minimum pre-test scores, higher maximum pre-test scores, lower pre-test averages and greater deviations of the minimum pre-test score from the average pre-test.

Table 4.6 Average Value of Indirect Class Size Variable for Small and Large Classes

Variable	CLASIZE 30 and Less	CLASIZE 100 and Greater
PREMAX	18.30	19.17
PREMIN	5.47	3.35
PRESTDEV	3.40	3.23
PREAVG	11.13	10.50
RANGE	12.83	15.81
DEVFRMAVG	5.67	7.15

Fifth, SKEW and KURT may be inappropriate proxies for class size when using a measure of ability, such as the pre- and post-TUCE III. Skewness is positive when the tail of the ability distribution is on the right. A positive skew would suggest that low ability scores are bunched closer to the mean and that high values extend far above the mean. A positive skew for pre-TUCE III indicates that a class has students of lower ability bunched together. This bunch of students with lower scores will be able to improve by a greater percentage from the pre- to post-exam. Students in the tail of the positive skew will naturally improve because of their higher ability. Alternatively, skewness is negative when the tail of the ability distribution is on the left. A negative skew suggests that high ability scores are bunched closer to the mean and that low values extend far below the mean. A negative skew indicates that a class is composed with more able students who will tend to excel. Students in the tail of the negative skew have a better percentage chance at improving their score than students who naturally tend to excel. Given these considerations, SKEW should be positively correlated with achievement when a pre- and post-test is used to measure achievement.

Kurtosis measures the extent to which ability is spread out or clustered together. The higher the kurtosis, the more scores are clustered together. Since both the sample of 2,942 observations and the TUCE III database may oversample good students, ability scores would be clustered at higher levels. Students of greater ability will naturally perform better on a post-exam resulting the kurtosis being positively correlated with achievement. The higher the kurtosis, the more focused teaching efforts toward students of similar ability can be, which could improve achievement. When kurtosis is lower and ability scores are more spread out, the possibility that the lower achieving students are guaranteed a greater gain in achievement due to the nature of the testing variable is higher; at the same time, the higher scoring students will naturally tend to excel. Using the pre-TUCE III 30 as a dependent variable measure could result in a positive correlation with achievement, regardless of the magnitude of kurtosis because of the improvement bias associated with using a pre, post-test measure of achievement.

4.2. Correcting for the Oversampling of Large Classes by Using Class Size Ranges.

The oversampling of large classes causes an exaggerated effect on the direct and indirect class size coefficients. Class size may be positive because of the significant number of students in large classes who are more conscientious and who will always do better from the pre- to post-test. In retrospect, this consequence could have been predicted if identified beforehand. The influence of the oversampling may be reduced by the use of dummy variables that represent a limited range of class sizes. The use of range dummies would allow classes within similar class sizes and with similar characteristics to influence achievement. The influence of class size ranges would mean that observations within the upper class size ranges have essentially the same probability to influence achievement as do observations in the smaller class size ranges. In effect, the issues involved with oversampling large classes and the problem of aggregation may be greatly reduced.

Table C.6 in Appendix C identifies the total number of observations for each class size in the sample of 2,942 observations. Class sizes range from 10 through 232 students in principles of economics courses. Class sizes increase by one and two students per class size through class size 64, but then class size increments by a greater number of students per class size. The most noticeable gap can be seen in class size range 132 through 193, limits in which no observations exist. One class of 194 and one class of 232 is used.

Table C.6 is useful in identifying the ranges for the class size dummy variables. All class size ranges should have a similar number of observations to reduce the oversampling and aggregation effects. Also, a class size range should be wide enough so that the observations in the range are representative of the characteristics of the sample. For example, the range needs to be wide enough to include a variety of institutional types as identified by DOCI, COMPI, LIBI, TWOI or PUBPRI. Table 4.7 identifies class size ranges based on the preceding requirements.

Table 4.7 Observations per Class Size Range

Variable Name	Class Size Range	Observations in Range
CLAS0	0 - 30	307
CLAS1	31 - 40	627
CLAS2	41 - 50	457
CLAS3	51 - 75	429
CLAS4	76 - 100	602
CLAS5	101 - 132	343
CLAS6	133 - 232	177

CLAS0 is the base class size range to which other class size ranges are compared in the regression. The coefficients of the class size ranges will indicate if larger classes add to or take away from achievement as compared to classes of 0 through 30 students. The six class size ranges are dummy variables, which are coded 1 and 0. The dummy variable is coded 1 if the student observation is enrolled in a class within that range; otherwise, it is coded 0. The dummy variable cannot be the actual number of the direct or indirect class size variables since the actual number is biased. The reasons for the bias are mentioned in section 3.1 of this chapter.

CLAS6 will not be used for regression purposes since only two schools are represented in this entire range. One class size consists of 194 students and one of 232 students. No more than one instructor, one class and one institutional type is represented by each class size. While CLAS6 is a wide range, the characteristics of each observation within the range is narrow. These two classes could be considered outliers; therefore, CLAS6 will not be used in the regression.

Table 4.8 compares three OLS regressions. The first regression focuses on the determinants of student achievement, the second on the instructional overhead model.

Table 4.8 The Influence of Class Size Ranges

Variable	DETRMNTS Coefficient	INSTRUCT Coefficient	RANGE Coefficient
Constant	-4.9823	-4.9490	-3.1762
	(-6.3)	(-6.2)	(-3.7)
DOCI	0.9269	0.5394	0.9058
	(2.9)	(1.6)	(2.7)
COMPI	-0.5129	-0.6258	-1.1613
	(-1.6)	(-2.0)	(-3.4)
LIBI	0.6395	0.5686	0.0859
	(1.6)	(1.4)	(0.2)
PTOT	0.611	0.6155	0.6144
	(29.6)	(30.0)	(28.9)
SEPRSX	0.3402	0.2219	0.3198
	(0.9)	(0.6)	(0.8)
ENG	0.5093	0.2639	0.3281
	(1.5)	(0.8)	(1.0)
CNT	1.0584	1.1730	1.2934
	(5.6)	(6.2)	(6.6)
RSCHPCT	0.0029	0.0000	-0.0018
	(0.6)	(0.0)	(-0.4)
PHD	1.0239	0.9688	1.8785
	(4.2)	(4.0)	(6.7)
YRSTCH	-0.0091	-0.0211	-0.0320
	(-0.8)	(-1.9)	(-2.6)
DSGWB	0.2024	0.3896	0.4036
	(0.8)	(1.6)	(1.6)
DHW	0.3846	0.4285	0.3666
	(2.2)	(2.4)	(2.0)
DRH	0.6174	0.7308	0.7902
	(2.9)	(3.4)	(3.6)
COMPFNL	0.0114	0.0106	0.0114
	(5.7)	(5.2)	(5.4)
RESPOND	0.0015	0.0138	0.0123
	(0.1)	(0.5)	(0.5)
GEN	1.1191	1.1178	1.0979
	(7.4)	(7.4)	(7.1)
GPA	0.0248	0.0246	0.0237
	(17.3)	(17.1)	(16.4)
HSECON	-0.023	-0.0017	0.0894
	(-0.2)	(0.0)	(0.6)
JOB	0.1311	0.1730	0.2492
	(0.8)	(1.1)	(1.6)
HRSTDY	-0.0656	-0.0631	-0.0778
	(-2.0)	(-2.0)	(-2.3)

Table 4.8 Continued.

Variable	DETRMNTS Coefficient	INSTRUCT Coefficient	RANGE Coefficient
INTIPT	0.0437	0.0452	0.0426
	(8.6)	(8.9)	(8.1)
ENTPRP	-0.0041	-0.0071	-0.0091
	(-0.6)	(-1.1)	(-1.4)
PUBPRI	1.8708	1.9776	1.8011
	(7.3)	(7.7)	(6.7)
DMACMIC	0.2189	0.1951	-0.1011
	(1.4)	(1.2)	(-0.6)
CLASIZE		0.0092	
		(4.9)	
CLAS1			-0.9381
			(-3.3)
CLAS2			-1.0827
			(-3.3)
CLAS3			-0.6315
			(-2.0)
CLAS4			-0.7709
			(-2.4)
CLAS5			-1.2450
			(-3.4)
Observations.	2,942	2,942	2,765
R-squared	0.49	0.49	0.51
Adjusted R-squared	0.48	0.48	0.51
Regression F-test [DF]	116.26 [24; 2,917]	113.50 [25; 2,916]	98.54 [29;2,735]
Prob of F-test	0.00	0.00	0.00
Breusch-Pagan [DF]	46.82 [24]	47.21 [25]	50.55 [29]

Notes: The numbers within the parentheses are the t-ratios derived from White's (1980) heteroscedasticity corrected variance-covariance matrix. The determinants of achievement regression (DETRMNTS) and the direct instructional overhead regression (INSTRUCT) are included for comparison to the class range regression (RANGE). The base range in RANGE is 0 through 30 students per class.

The third regression presents the class size range model. The signs of the statistically significant variables in the DETRMNTS regression remain unchanged in the RANGE regression.[1] The RANGE regression confirms the direction of influence that the significant input variables have on achievement when first identified in the DETRMNTS regression. Each class size range is negative and significant. Each class size range greater than 30 students will reduce the mean post-TUCE III achievement score by the coefficient amount. The RANGE regression suggests that increasing the size of a class above 30 students per class will reduce the average post-TUCE III 30 measure of achievement in principles of collegiate economics courses.

Section 5: The Influence of Class Size on Differing Cognitive Skills.

Class size appears to diminish mean achievement when class size grows beyond 30 students. An advantage of using TUCE III is that one can identify the effect of class size on lower level and higher level cognitive skills. Using the aggregate TUCE III 30 score as the dependent variable may mask the influence of class size on different levels of cognitive skills. When the achievement variable can be defined for different cognitive skills, the impact of class size on achievement can be described with more precision.

TUCE III is designed "... to emphasize the application of basic concepts and principles." (Examiners Manual, 1991) Two-thirds of the questions in the exam are classified as application questions. The application questions "... require students to go beyond memorization and recall." (Examiners Manual, 1991) Application questions ask students to apply basic terms, concepts or principles in economics either explicitly or

---

[1] YRSTCH still remains negative but is significant in the RANGE regression. As more experience is gained teaching, instructors are expected to develop more effective pedagogies. However, improvement in communication and organization skills may plateau after a certain number of years. After the plateau year is reached, improvement in student achievement due to teaching skill will also cease to increase. Also, as professors gain more experience, the greater the likelihood that the professor will teach lessons gleaned from professional and life experience rather than the objectives of the TUCE III. Large classes taught by professors who have a significant number of years teaching would appear to reduce achievement on the post-TUCE III 30. Large classes as opposed to small classes have more observations that can influence outcomes. Consequently, after class size is taken into account, the number of years teaching could be negatively related to achievement.

implicitly in order to find the best solution. The explicit application (EA) questions and the implicit application (IA) questions attempt to test for the third category in Bloom's (1956) taxonomy--application. EA and IA questions may be used as measures of higher ordered cognitive skills, such as critical thinking. The other one-third of the questions in the exam assess a basic understanding of terms, concepts and principles. The recognition and understanding (RU) category of questions attempt to test for the first two categories in Bloom's (1956) taxonomy--knowledge and comprehension. RU questions may be used as a measure of lower level cognitive skills, such as identifying definitions.

The dependent variable achievement can be redefined in the class size range regression as (1) the number of correctly answered RU question on the post-TUCE III 30 and (2) the number of correctly answered EA and IA questions on the post-TUCE III 30. The post-test RU variable is labeled TRU, while the EA plus IA variable is labeled TEAIA. There are 10 RU and 20 EA plus IA questions in the post-exam. The pre-TUCE III is included as an independent scale variable. Since pre-test questions are comprised of the same category type, the pre-test variable can be redefined as PRU for the 10 pre-test RU questions and PEAIA for the 20 pre-test EA plus IA questions.

Table 4.9 compares the results of the class size range regression model developed in section four but defines achievement using RU skills and EA plus IA skills. As expected, each class size range is negative and significant for each class size range above base range. When class size rises above 30 students, some erosion in achievement occurs. In general, each t-ratio is less and each parameter is smaller in the RU regression for each class size range when compared to the full class size range model where achievement is defined as the number of correctly answered questions out of 30. Each t-ratio is greater but coefficients are smaller for the EA plus IA category when compared to the full range

Table 4.9 Comparison of Class Size Influence on Differing Cognitive Skills

	DETRMNTS	RANGE	RU	EA + IA
Variable	Coefficient	Coefficient	Coefficient	Coefficient
Constant	-4.9823	-3.1762	-0.5726	-1.7795
	(-6.3)	(-3.7)	(-1.5)	(-2.8)
DOCI	0.9269	0.9058	0.0893	0.9240
	(2.9)	(2.7)	(0.6)	(3.6)
COMPI	-0.5129	-1.1613	-0.6054	-0.6618
	(-1.6)	(-3.4)	(-3.9)	(-2.5)
LIBI	0.6395	0.0859	-0.1699	0.2576
	(1.6)	(0.2)	(-0.9)	(0.8)
PTOT	0.611	0.6144		
	(29.6)	(28.9)		
PRU			0.3702	
			(17.4)	
PEAIA				0.4759
				(22.2)
SEPRSX	0.3402	0.3198	-0.1380	0.2842
	(0.9)	(0.8)	(-0.7)	(0.9)
ENG	0.5093	0.3281	0.3860	0.1495
	(1.5)	(1.0)	(2.4)	(0.6)
CNT	1.0584	1.2934	0.2595	0.9266
	(5.6)	(6.6)	(2.9)	(6.2)
RSCHPCT	0.0029	-0.0018	-0.0002	0.0014
	(0.6)	(-0.4)	(-0.1)	(0.4)
PHD	1.0239	1.8785	0.6491	1.3778
	(4.2)	(6.7)	(5.0)	(6.4)
YRSTCH	-0.0091	-0.0320	-0.0168	-0.0197
	(-0.8)	(-2.6)	(-2.9)	(-2.1)
DSGWB	0.2024	0.4036	-0.0748	0.3667
	(0.8)	(1.6)	(-0.7)	(1.9)
DHW	0.3846	0.3666	0.2015	0.1578
	(2.2)	(2.0)	(2.4)	(1.2)
DRH	0.6174	0.7902	0.2748	0.5103
	(2.9)	(3.6)	(2.8)	(3.1)
COMPFNL	0.0114	0.0114	0.0039	0.0073
	(5.7)	(5.4)	(4.0)	(4.6)
RESPOND	0.0015	0.0123	-0.0064	0.0396
	(0.1)	(0.5)	(-0.5)	(2.1)
GEN	1.1191	1.0979	0.4447	0.8121
	(7.4)	(7.1)	(6.3)	(7.0)
GPA	0.0248	0.0237	0.0093	0.0166
	(17.3)	(16.4)	(14.3)	(15.6)
HSECON	-0.023	0.0894	0.0150	0.0507
	(-0.2)	(0.6)	(0.2)	(0.4)

Table 4.9 Continued.

Variable	DETRMNTS Coefficient	RANGE Coefficient	RU Coefficient	EA + IA Coefficient
JOB	0.1311	0.2492	0.1681	0.1423
	(0.8)	(1.6)	(2.3)	(1.2)
HRSTDY	-0.0656	-0.0778	-0.0111	-0.0793
	(-2.0)	(-2.3)	(-0.7)	(-3.1)
INTIPT	0.0437	0.0426	0.0165	0.0284
	(8.6)	(8.1)	(6.8)	(7.0)
ENTPRP	-0.0041	-0.0091	-0.0012	-0.0093
	(-0.6)	(-1.4)	(-0.4)	(-1.9)
PUBPRI	1.8708	1.8011	0.7750	1.2748
	(7.3)	(6.7)	(6.8)	(6.3)
DMACMIC	0.2189	-0.1011	-0.4384	0.5290
	(1.4)	(-0.6)	-(6.1)	(4.3)
CLAS1		-0.9381	-0.3137	-0.7688
		(-3.3)	(-2.5)	(-3.5)
CLAS2		-1.0827	-0.2749	-0.9820
		(-3.3)	(-1.9)	(-3.9)
CLAS3		-0.6315	-0.2141	-0.5938
		(-2.0)	(-1.5)	(-2.4)
CLAS4		-0.7709	-0.3528	-0.6633
		(-2.4)	(-2.5)	(-2.7)
CLAS5		-1.2450	-0.4647	-0.7665
		(-3.4)	(-2.8)	(-2.8)
Observations.	2,942	2,765	2,765	2,765
R-squared	0.49	0.51	0.33	0.45
Adjusted R-squared	0.48	0.51	0.33	0.44
Regression F-test	116.26	98.54	47.82	77.34
[DF]	[24; 2,917]	[29;2,735]	[29; 2,735]	[29; 2,735]
Prob of F-test	0.00	0.00	0.00	0.00
Breusch-Pagan	46.82	50.55	41.20	43.60
[DF]	[24]	[29]	[29]	[29]

Notes: The numbers within the parentheses are the t-ratios derived from White's (1980) heteroscedasticity corrected variance-covariance matrix. The determinants of achievement regression (DETRMNTS) and the class size range model (RANGE) are included for comparison to the cognitive skills regressions. RU represents the recognition and understanding cognitive skill of the pre-TUCE III 30 and EA + IA represents the sum of the explicit understanding and implicit understanding categories of the pre-TUCE III 30. The definitions of the class size ranges remain the same as the original RANGE regression.

model. However, the coefficients in the EA plus IA model are significantly greater than the parameters in the RU model. The coefficients in the full range model appear to equal the sum of the other two cognitive models. For example, the RANGE regression indicates that students in the largest class size range, CLAS5, will see their post-TUCE III 30 score decline by 1.2450 points. When TUCE III is decomposed into RU and EA plus IA categories, CLAS5 reduces the post-RU score by 0.4647 points and lowers the post-EA plus IA score by 0.7665 points. Students in class size range CLAS5 will have their EA plus IA learning more negatively affected than their RU learning. When the two coefficients are summed together they total -1.2313 (-0.4647 - 0.7665), which is approximately equal to the corresponding class size coefficient -1.2450 in the RANGE regression. The additivity would suggest that the influence of class size on student achievement has been appropriately decomposed. The general conclusion would be that increasing class size negatively impacts higher ordered cognitive skills more than lower ordered cognitive skills once class size rises above 30 students.

The determinants of achievement other than class size appear to influence TRU achievement differently than TEAIA achievement. The most surprising difference among the regression comparisons is the RESPOND variable. While RESPOND does not significantly influence TRU, the variable does make a positive and statistically significant contribution toward TEAIA. The interpretation of this result could be that higher ordered cognitive skills are influenced by increasing amounts of instructor-student feedback in the classroom. Typically, the coefficients and the t-ratios for the EA plus IA regression are larger than for the RU regression. The educational inputs contribute more toward higher ordered cognitive skills than lower ordered cognitive skills. The signs of the statistically significant inputs are the same in both the full range, RU and EA plus IA models, with the exception of DMACMIC. This dummy variable is coded 1 for principles of microeconomics. Apparently, application questions in microeconomics are easier, while recognition and understanding of basic concepts is harder. The coefficients in the full range model appear to be additive of the other two cognitive models. For example, the intensive use of homework (DRH) contributes 0.7902 points to the mean post-TUCE III 30 score. When TUCE III is decomposed into RU and EA plus IA categories, DRH contributes 0.2748 points to the post-RU score and 0.5103 points to the post-EA plus IA score. DRH contributes more toward EA plus IA learning than to RU learning. When the two coefficients are summed together they equal 0.7851 (0.2748 + 0.5103), which is approximately equal to the corresponding class size coefficient 0.7902 in the RANGE regression. The additivity would suggest that the influence of class size on student achievement has been appropriately decomposed. The general conclusion would be that education inputs other than class size impact higher ordered cognitive skills more than lower ordered cognitive skills.

Section 6: Chapter Summary.

This chapter has demonstrated that student achievement in principles of college economics is influenced by several categories of educational inputs, which include the following: faculty capital; student capital; technology; environmental factors. This chapter also has identified several statistically significant input variables. The empirical results confirm and expand the findings of previous research on the determinants of student achievement other than class size. In all, the empirical results present comprehensive evidence that class size significantly influences student achievement.

Class size and student achievement in principles of college economics are negatively related. Once class size rises above 30 students, statistically significant evidence reveals that average achievement deteriorates. Class size does not seem to diminish mean achievement as much when achievement is defined as lower level cognitive skills, such as the RU questions on TUCE III 30. However, class size reduces mean achievement more significantly when achievement is defined as higher level cognitive skills like the EA plus IA questions on the TUCE III 30.

Student achievement in principles of college economics is also influenced by statistically significant factors other than class size. Student achievement is positively influenced by: counting the post-TUCE III toward the course grade; receiving instruction from a faculty member with an earned doctorate; using teachers who speak English as a native language; using homework intensively; using non-text readings and handouts intensively; giving comprehensive exams; starting a course with a higher stock of economic knowledge; having a higher cumulative GPA; being male; attending a private school; increasing the amount of instructor-student feedback; and having a greater interest in economics. In general, higher ordered cognitive skills seem to be more significantly influenced by educational inputs than do lower ordered cognitive skills.

Chapter V

Policy Implications and Suggestions for Further Research

Increasing the size of a principles of college economics class above 30 students per class <u>**will**</u> reduce mean student achievement, ceteris paribus. However, class size is not the only variable that may influence achievement. This chapter explores how class size interacts with a variety of educational inputs. In light of constraints on college budgets, policy makers should be interested in the impact that small or large introductory college economics classes may impose on university budgets. In addition, this chapter reviews several policy implications suggested by the empirical findings. It recommends class size policies that seek to maximize achievement while minimizing cost. It also suggests ways to neutralize any unwanted consequences on either achievement or budgets. Further, this chapter makes several suggestions for further research..

Section 1: Policy Implications.

Economics departments are confronted with a dilemma. On the one hand, principles of economics class sizes of 30 students or fewer are likely to see gains in mean achievement. However, reducing class size leads to an increase in the cost for the additional faculty needed to teach the same number of students. On the other hand, increasing class sizes above 30 students per course will reduce mean achievement. But, increasing class size will reduce the number of faculty needed to teach introductory economics and will reduce the costs for contract faculty salaries and benefits. Increasing or decreasing class size may lead to unwanted negative consequences on either budgets or achievement. Any class size policy which is implemented must weigh the gains in mean achievement with the costs incurred. The solution to the achievement-cost dilemma is discussed in this section in four parts. First, controllable educational inputs that can be used to improve student achievement are identified. Second, evidence showing that student achievement in classes with enrollments of larger than 30 students can be outperformed by classes in which 30 or fewer students are enrolled is introduced. Third, the apparent achievement-cost dilemma associated with changing the size of a class is resolved. The empirical results suggest that class size and mean student achievement can increase, while simultaneously saving scarce budget funds. Fourth, specific course design and class size policies are recommended.

1.1. Improving Student Achievement Using Controllable Education Inputs Other than Class Size.

Teachers are essentially managers of a classroom. They exhibit all of the following management functions: planning; organizing; directing; and controlling. Principles of management teaches that some factors can be controlled by a manager and that other factors are uncontrollable. Teachers can control the variables directly related to the internal environment--the classroom. The instructor cannot directly control external environmental variables occurring outside of the classroom. Professors and administrators need to identify which inputs of the educational process can be controlled or at least influenced to improve student achievement and which inputs cannot be controlled or influenced. This section explores the educational inputs other than class size that can be controlled in order to improve student achievement in principles of college economics.

Table 3.2 in chapter three identifies the variables selected from the aggregate TUCE III database for the sample of 2,942 observations. Several variables that may be considered uncontrollable in the sample are identified. The institutional type and the public-private status of the institution are in this category. They cannot be influenced by the instructor. The instructor is incapable of altering student gender, the cumulative GPA of a student prior to enrolling in the economics class or the employment status of the student. The interest that a student has in economics or the students perception of the importance of economics may largely be determined by sociodemographic variables outside of the professors control. However, instructors may have the potential of influencing student interest in economics and student perception of the discipline's importance during the course of a semester. How the instructor objectively or tangibly influences student interest or importance is less certain. Professors can influence the comprehensiveness of the final exam or the intensity by which ancillary teaching aids are used. University administrators can require terminal economics degrees of principles instructors or change the percent of time faculty devote to research efforts. Policies attempting to improve student achievement should focus on educational inputs that can be controlled.

The regression in Table 4.4 of chapter four identifies the contribution that each of the educational inputs makes toward the post-TUCE III 30 measure of achievement. Each ordinary least squares (OLS) coefficient represents the marginal effect of the educational input. Twelve of the 25 variables are statistically significant, while five of the 12 can be controlled. Uncontrollable influences that affect performance but which an instructor cannot change include: institutional type (DOCI); the stock of economic knowledge with which a student enters a principles course (PTOT); the gender of the student (GEN); the cumulative GPA of the student (GPA); the hours a student spends studying economics (HRSTDY); the amount of student interest in economics (INTIPT); and the public or private status of the university (PUBPRI). The instructor can influence the remaining five variables and thereby improve achievement. Incorporating the following inputs into a principles of economics course design could contribute the following points to the mean post-TUCE III 30 score, assuming the input is not an existing part of the course design:

- Counting the post-TUCE III 30 (CNT) toward the final course grade can add 1.06 points to the mean post-TUCE III 30.
- Instructors who hold terminal degrees in economics (PHD) and who teach principles of college economics can add 1.02 points to the mean post-TUCE III 30.
- The intensive use of homework (DHW) will add 0.38 points to the mean post-TUCE III 30. DHW implies homework that is required and graded.
- The intensive use of non-text readings and handouts (DRH) will add 0.62 points to the mean post-TUCE III 30. DRH implies non-text readings and handouts that are required and graded.
- Making the final exam 100 percent comprehensive, can add 1.14 to the mean post-TUCE III 30. The 1.14 is found by multiplying the 0.0114 coefficient, which is the percent of the final exam that is comprehensive (COMPFNL), by 100. Some final exams may only be 60 percent comprehensive with the remaining 40 percent of the exam testing on material covered in the last part of the course.

Table 5.1 shows the cumulative marginal effect that the five controllable inputs have on achievement. A Wald test is used to sum the marginal effects of the individual variables. Table 5.1 uses the same inputs as the OLS regression in Table 4.4 of chapter four. The OLS regression in Table 4.4 excludes class size as a determinant of student achievement. The CONTROL column shows the combined marginal effects of requiring CNT, PHD, DHW, DRH and a 100 percent COMPFNL in courses where the inputs were not previously required. Incorporating these variables into a principles of economics course design will add 4.23 points to the mean post-TUCE III 30. Courses already using one of the inputs will not see the full 4.23 points added to the mean post-TUCE III 30 and courses already using all of the inputs will not see the increase in the mean post-exam.

Table 5.1 The Cumulative Marginal Effects of Controllable Inputs on Achievement - Class Size Excluded.

Variable	CONTROL	RU	EA + IA
CNT	1.0584	0.1520	0.8474
	(5.6)	(1.8)	(5.9)
PHD	1.0239	0.4669	0.5803
	(4.2)	(4.3)	(3.1)
DHW	0.3846	0.1992	0.2240
	(2.2)	(2.5)	(1.7)
DRH	0.6173	0.2201	0.4130
	(2.9)	(2.3)	(2.5)
COMPFNL	0.0114	0.0040	0.0071
	(5.7)	(4.3)	(4.7)
WALD	4.2287	1.4376	2.7683
	(11.2)	(8.4)	(9.6)

Notes: The numbers within the parentheses, except for Wald, are the t-ratios derived from White's (1980) heteroscedasticity corrected variance-covariance matrix. WALD is the Wald coefficient for the cumulative effects for the variables listed in the column. CONTROL indicates the variables which can be influenced by the course instructor or university administration. The cumulative marginal effects are computed assuming a 100 percent comprehensive final exam and 25 percent of class time in spent in dialog between the instructor and student in the classroom.

The marginal effects for each of the variables in the CONTROL column can be decomposed into the marginal effects that each variable has on either lower ordered cognitive skills or higher ordered cognitive skills. RU assesses basic recognition and understanding of economic principles. RU represents lower ordered cognitive skills, while EA assesses explicit application of economic concepts while IA assesses the implicit application of economic knowledge. EA plus IA are summed since both represent higher ordered cognitive skills. Essentially, the marginal effects of RU and EA plus IA are additive. CNT, DRH and COMPFNL are more statistically significant in the EA plus IA regression than in the RU regression. These controllable variables seem to add more to higher ordered cognitive skills. Improving higher ordered cognitive skills is most effective when: non-text readings and handouts are used intensively; the post-TUCE III is counted toward the final course grade; and final exams are 100 percent comprehensive. The importance of the 4.23 incremental score cannot be underestimated. Table C.7 in Appendix C indicates the national percentile rank of scores for students completing the post-TUCE III 30. The average macro score on the post-TUCE III 30 for the sample of 2,942 observations is 14.16 and for micro the average score is 15.53. The difference in achievement between principles of macroeconomics and principles of microeconomics is statistically insignificant. If a department of economics requires that the controllable elements identified in Table 5.1 be incorporated into a principles of economics course design where the course design did not previously incorporate the variables, then approximately four points can be added to the mean TUCE III 30 measure of achievement. This four point increase translates into a gain of approximately 20 percentile points.

The students and the university can benefit by requiring principles of economics courses to be designed with the variables identified in Table 5.1. The students benefit on average because their achievement will increase. The university benefits because the increase in mean achievement can be accomplished with little or no increase in cost. Counting the post-TUCE III toward the final grade (CNT), requiring a 100 percent comprehensive final exam (COMPFNL) and intensively using teaching ancillaries (DHW, DRH) may be added to the course design with potentially no increase in cost to the university. The intensive use of ancillary teaching aids does not necessarily mean that faculty time requirements for grading would increase. Homework, handouts and non-text readings can be assessed though incorporation into existing testing procedures. Final exams could simply be made more comprehensive rather than less comprehensive. The potential increase in cost could come from requiring professors to hold terminal degrees in economics or from an increase in faculty time requirements due to workload increase. Universities who use teaching assistants or adjunct professors to teach principles of economics may need their budgets increased in order to staff the classroom with terminally degreed professors. Instructors who are not able to include the intensive use of ancillaries into existing testing procedures may need help to grade the assignments, if academic advisement, committee work, research requirements and community service time commitments are not reassessed.

1.2. Improving Student Achievement with Small Class Enrollments.

The professional judgment and intuition of some college professors may suggest that a small class enrollment is a better teaching environment for the professor and better learning environment for the student. On average, students in a small class should achieve more than students enrolled in a large class. Economics professors may tend to agree with this belief, especially since the complexity of economics relative to other subject matter may make the subject more difficult to learn.

Table 4.8 in chapter four helps confirm this well entrenched belief by concluding that increasing the size of a class above 30 students will reduce achievement in principles of college economics. Table 5.2 reproduces the marginal effects of the class size ranges found in Table 4.8. Each class size range reduces achievement by approximately one point.

Table 5.2. Marginal Effects of Class Size Ranges

Variable Name	Class Size Range	Coefficient
CLAS1	31 - 40	-0.9381 (-3.3)
CLAS2	41 - 50	-1.0827 (-3.3)
CLAS3	51 - 75	-0.6315 (-2.0)
CLAS4	76 - 100	-0.7709 (-2.4)
CLAS5	100 - 132	-1.2450 (-3.4)

Students enrolled in a class of more than 30 students will achieve approximately one point less, on average, on the post-TUCE III 30 than students enrolled in a class of 30 or fewer students. The average post-TUCE III 30 for principles of macroeconomics is 14.09, while the average post-TUCE 30 for principles of microeconomics is 15.48. The averages are computed for classes of 10 through 132 students in the sample of 2,942 observations. Using Table C.7 in Appendix C, a one point reduction in the post-TUCE III 30 from the mean score could reduce student achievement by approximately six percentile points when compared to the national performance of all students. Alternatively, reducing class size to 30 students or fewer will improve achievement by approximately one point and increase the national ranking by approximately six percentile points. Reducing principles of economics class sizes to 30 or fewer students will increase mean achievement.

While the reduction in class size may improve mean achievement, the improvement has a cost. The following hypothetical example will illustrate the incremental cost issue associated with class size reduction:

> Suppose a college has 16 sections of introductory economics classes being taught in both the fall and spring semester. The classes are three semester hours each. The total enrollment in all sections for each semester is 720 students. Introductory classes are taught by full-time associate professors who teach 12 credit hours each. Some of the class sizes are fewer than 30 students, while the enrollment of many classes exceeds 30. The economics department chair believes that reducing class size for all classes to 30 or fewer students will improve mean achievement. How many faculty are needed to teach the 16 classes now offered? How many faculty are needed to teach the 720 students in classes when enrollment is required to be 30 or fewer students? What will the increase in faculty cost the college?

In the hypothetical example there are an average of 45 students per course section (720 students / 16 sections). Four full-time faculty members with teaching loads of 12 credit hours each are needed per semester to teach the 720 students (16 sections/ 4 sections per professor). Lowering the enrollment limit to 30 students per class will require a minimum of 24 (720 students/ 30 students per section) sections to be taught. Twelve semester credit hour teaching loads translate into four classes of three semester credit hours each. Six (24 sections/4 sections per professor) full-time associate professors are needed to teach the 24 sections, an increase of two full-time associate professors. The cost to the university requires some elaboration.

The cost of a faculty member consists of the nine-month salary contract plus benefits. According to Linda Cooper in the Office of Human Resources at Middle Tennessee State University in Murfreesboro, Tennessee, a conservative estimate of benefits attached to a professor's contract amounts to an average of 35 percent of the nine-month salary contract. Benefits include group insurance, employer social security contributions, employer retirement contributions, sick leave, study or sabbatical leave and tuition reduction for continuing education or the education of immediate family members. Table 5.3 identifies the average salary for economics professors by academic rank. The source of the salary numbers is a salary survey published by The American Assembly of the Collegiate Schools of Business Statistical Service, which nationally surveys faculty salaries at business schools. The salary numbers do not include benefits. Table 5.3 calculates benefits at 35 percent of a nine-month contract salary, in addition to calculating the total amount of a nine-month compensation package.

In the hypothetical situation the two additional associate professors needed for both the fall and spring semesters in order to reduce class size to improve average achievement will cost the university approximately \$129,000 (\$64,500 x 2 associate professors). The cost of improving mean student achievement in principles of economics by approximately six national percentile points is \$21,500 (\$129,000 / 6 percentile points) per percentile.

Table 5.3 Estimate of Contract Salary Plus Benefits for Economics Faculty at AACSB Schools in 1992 - 1993.

Rank	9-Month Contract Salary	Benefits at 35% of Salary	Salary plus Benefits
Professor	$64,500	$22,600	$87,100
Associate	47,800	16,700	64,500
Assistant	42,900	15,000	57,900

1.3. The Economic Efficiency of Increasing Class Size.

The benefit of reducing principles of economics class size is an increase in mean achievement. The cost of reducing class size is an increase in wages and benefits required for additional faculty that must be hired. The dual dimension of this problem would suggest that increasing class size would reduce mean achievement and lower cost. The decision maker is apparently caught in a dilemma, either to reduce class size and increase cost, or to increase class size and reduce mean achievement. Either choice appears to have an unwanted consequence. The point is to make the most efficient choice, class size reduction or class size increase. One part of this efficiency problem centers on how to value a one percentile increase in performance. Is $21,500 per percentile point a bargain or unreasonably expensive? Fortunately the dilemma does not exist. The choice does not have to result in mean achievement being lowered or cost being raised. Average student achievement can be raised while cost is lowered at the same time.

The hypothetical example continues as follows:

> The economics department chair believes that by increasing class size to 90 students per section and that by controlling selected educational inputs, costs can be reduced while achievement improved. How many faculty are needed to teach the 720 students in classes when enrollment is required to be 90 or more

> students? Which educational inputs should be controlled in the introductory courses? If controlling the education inputs has a zero marginal cost, how much can the college save with the reduction in faculty? What will be the net increment to achievement?

When the change in course enrollment occurs, only 8 (720 students / 90 students per section) sections are needed to teach the 720 students. Two full-time faculty members with teaching loads of 12 credit hours each are needed per semester to teach the 720 students (8 sections / 4 sections per professor). The controllable educational inputs that should be required of the course are identified in section 1.1 of this chapter: CNT; PHD; DHW; DRH; and COMPFNL. The reduction in faculty will be 2 associate professors (4 before class size change - 2 after class size change). This faculty reduction will save the university approximately $129,000 ($64,500 x 2 associate professors) per year. The savings are based on estimated salaries for economics professors at AACSB business schools as identified in Table 5.3. The net increment to mean achievement can be discussed in the context of Table 5.4.

Table 5.4 presents the cumulative marginal effect that the five controllable inputs have on average achievement. A Wald test is used to sum the marginal effects of the individual variables. The variable labeled WALD is the cumulative marginal effect without the class size component. The 'W' preceding the class size range variables indicates the cumulative marginal effects of the educational inputs on average student achievement when class size falls within the range of each class size variable. The class size ranges are identified in Table 4.8 of chapter four. Class size has a negative influence on average achievement, but managing the educational inputs appropriately can actually increase achievement. In the hypothetical example, the 90 students per class are in the CLAS4 range. If class size is increased to 90 students and no other factors change, mean

Table 5.4 The Cumulative Marginal Effect of Class Size and Controllable Inputs on Achievement.

Variable	CONTROL	RU	EA + IA
CNT	1.2934	0.2595	0.9266
	(6.6)	(2.9)	(6.2)
PHD	1.8785	0.6491	1.3778
	(6.7)	(5.0)	(6.4)
DHW	0.3666	0.2015	0.1578
	(2.0)	(2.4)	(1.2)
DRH	0.7902	0.2748	0.5103
	(3.6)	(2.8)	(3.1)
COMPFNL	0.0114	0.0039	0.0073
	(5.4)	(4.0)	(4.6)
CLAS1	-0.9381	-0.3137	-0.7688
	(-3.3)	(-2.5)	(-3.5)
CLAS2	-1.0827	-0.2749	-0.9820
	(-3.3)	(-1.9)	(-3.9)
CLAS3	-0.6315	-0.2141	-0.5938
	(-2.0)	(-1.5)	(-2.4)
CLAS4	-0.7709	-0.3528	-0.6633
	(-2.4)	(-2.5)	(-2.7)
CLAS5	-1.2450	-0.4647	-0.7665
	(-3.4)	(-2.8)	(-2.8)
WALD	5.4692	1.7729	3.6988
	(13.5)	(9.5)	(12.0)
WCLAS1	4.5312	1.4592	2.9300
	(9.1)	(6.3)	(7.7)
WCLAS2	4.3865	1.4980	2.7169
	(8.5)	(6.3)	(6.9)
WCLAS3	4.8378	1.5588	3.1050
	(10.0)	(6.9)	(8.3)
WCLAS4	4.6984	1.4201	3.0356
	(9.6)	(6.2)	(8.1)
WCLAS5	4.2243	1.3082	2.9324
	(8.8)	(5.8)	(8.1)

Notes: The numbers within the parentheses, except for Wald, are the t-ratios derived from White's (1980) heteroscedasticity corrected variance-covariance matrix. WALD is the Wald coefficient for the cumulative effects for the variables listed in the column. CONTROL indicates the variables which can be influenced by the course instructor or university administration. The cumulative marginal effects are computed assuming a 100 percent comprehensive final exam and 25 percent of class time in spent in dialog between the instructor and student in the classroom. The 'W' preceding the class size range coefficient is the cumulative effects for the variables when the size of a class falls within the indicated range.

achievement would decline by 0.77 or approximately one point on the post-TUCE III 30 when compared to students in classes with enrollments of 30 students or less. However, by requiring the incorporation of the controllable variables in course designs where the inputs were not previously required, an average class size of 90 could see and average improvement of 4.70 points when compared to classes of 30 students or less. An approximate 5 point gain in the mean post-TUCE III 30 translates into approximately a 25 percentile improvement in performance when compared to the national ranking in Table C.7 of Appendix C. The average macro score and the average micro score are the same as the example used with Table 5.1. Courses already using one of the inputs will not see the full 4.70 points added to the mean post-TUCE III 30, and courses already using all of the inputs will not realize the increase in the mean post-exam.

The marginal effects from CONTROL can be decomposed into lower and higher ordered cognitive skills. Once again RU and EA plus IA appear to be additive. The variable coefficients are generally more significant in the EA plus IA regression than in the RU regression. The controllable educational inputs seem to contribute more toward explicit and implicit application skills rather than basic recognition skills of economic principles.

Although the cost to the university may decline when class size is increased, the assumption of zero marginal cost may need to be relaxed. Faculty teaching larger classes may see an increase in workload, especially if teaching ancillaries are used intensively. If faculty time allocations to academic advising, committee work, research requirements and community service are not reassessed, then faculty may not want to teach large classes. At some universities where the workload for large classes increases, an increase in cost may occur.

The potential increase in cost is best illustrated by returning to the hypothetical example. When class size is increased to 90 students per course section, the number of required associate professors is reduced to two. Although the university will realize reduced costs, these professors will see an increase in their personal marginal cost if their workloads increase. The professors ask the department chair for help. The department chair assigns each professor two graduate assistants. The graduate students are assigned to work with two principles of economics classes. There are four graduate assistants who each work 20 hours per week for a total of 80 hours. Each graduate student spends 10 hours per week per course section for each of the two assigned sections. The 10 hours is spent grading assignments and being available for student questions. Two hours per day, five days a week may be made available to meet with principles students and grade assignments. The increase in personal marginal cost to the professor has been reduced; however, the university now realizes an increase in marginal cost.

The College of Graduate Studies at Middle Tennessee State University estimates the cost of a nine month graduate assistant stipend to be $4,550 plus fees. Fees at MTSU include tuition and other registration expenses. The average of tuition for in and out-of-state students plus other fees is approximately $2,025 per semester and approximately $4,050 for one academic year. MTSU does not incur benefit expenses for graduate assistants. The average cost per graduate assistant per academic year is $8,600 at MTSU. Assuming that costs for graduate assistants are similar at other state universities, the four students needed in the hypothetical example have a marginal cost to the university of $34,400. A net savings of $94,500 can be realize after four graduate students are employed to offset the increase in workload for faculty members.

An alternative to the use of graduate assistants would be to offer those professors desiring to teach the large classes overload pay to compensate them for the increase in workload. The overload pay could approximate the funds required for graduate assistants. In the hypothetical example, each of the two associate professors could see their contract salary with benefits increase by approximately $17,000. As is the case in the graduate assistant example, the university can realize a net savings by offering large principles of economics classes.

The results from Table 5.4 show that the negative influence on achievement of class sizes larger than 30 students can be offset by managing educational inputs effectively. When class size increases are accompanied by controlled use of influential inputs, gains in achievement can be realized. The benefits of large class management are actual and real. The model predicts there will be a gain in mean student achievement when certain critical inputs not previously incorporated into a course design are made part of an introductory economics course. Actual university expenditures can decline due to reduction in faculty contract compensation both when marginal costs are zero and under assumptions of either graduate labor substitution or faculty overload remuneration. Effectively managing large principles of economics classes is the most efficient choice that economics professors and university administrators can make.

1.4. Policy Recommendations.

The foregoing discussion on the implications of the empirical results leads to three policy recommendations for principles of college economics.

1. The RANGE equation in Table 4.8 can be used to predict achievement scores on the post-TUCE III 30 for principles of economics courses. TUCE III can best be used when the test results are compared to national performance. Individual economics instructors and economics departments can gauge their effectiveness more accurately when comparing actual mean TUCE III 30 achievement scores with the mean scores predicted by the RANGE equation in Table 4.8.

The assessment of student achievement in principles of college economics using TUCE III is more complex than looking at mean post-test scores because many educational inputs influence achievement. Comparing a mean post-TUCE III 30 score from one class or department to another class or department is inappropriate. A lower mean score does not necessarily indicate that something is wrong with the instructor or the department; the lower mean could have resulted from less able students or from a larger number of students in a class, rather than faculty incompetence. An economics department that calculates a mean score of 15 on the macroeconomics post-TUCE III 30 is not necessarily better than other schools when the average national macroeconomics post-TUCE III 30 score is approximately 14. If the economics department enrolls more capable students and requires economics principles to be taught by faculty with a doctorate, then the department has an advantage that must be taken into consideration. The RANGE equation in Table 4.8 can account for advantages and disadvantages in assessing a mean post-TUCE III 30 score. These ideas can best be illustrated by a hypothetical example:

> Professors A and B each teach principles of macroeconomics at State University. The professors do not teach any other classes and are not engaged in academic research. The economics department chair will give a $1,000 award to the more effective teacher. The TUCE III 30 is used as a measure of student achievement and as a proxy for overall teaching effectiveness. Professor A's mean post-test score is 17, while professor B's mean post-test score is 13. Based on these results should professor A or B receive the $1,000 award for teaching effectiveness?

If the award is based on mean student achievement, then professor A should receive the $1,000. After all, professor A appears to be able to increase mean student achievement more than professor B. However, the advantages and disadvantages each instructor must manage should be considered. Table 5.5 identifies the mean values for the education inputs in the classes that professors A and B teach. Both professors are similar in that they: teach at a doctoral institution (DOCI); have the same English communication skills (ENG); count the post-TUCE III 30 toward the final course grade (CNT); are not engaged in academic research (RSCHPCT); have earned doctorates in economics (PHD);

Table 5.5 Hypothetical Example Comparing Actual Post-TUCE III 30 Results with Predicted Results for Two Instructors with Different Mean Input Values - Before and After Requiring Controllable Educational Inputs in the Course Design for Principles of College Macroeconomics.

		MEAN INPUT BEFORE CHANGE		MEAN INPUT AFTER CHANGE	
Variable	Coefficient	INSTRUCTOR A	INSTRUCTOR B	INSTRUCTOR A	INSTRUCTOR B
Constant	-3.1762				
DOCI	0.9058	1.0000	1.0000	1.0000	1.0000
COMPI	-1.1613	0.0000	0.0000	0.0000	0.0000
LIBI	0.0859	0.0000	0.0000	0.0000	0.0000
PTOT	0.6144	12.0000	8.0000	12.0000	8.0000
SEPRSX	0.3198	0.0000	0.0000	0.0000	0.0000
ENG	0.3281	1.0000	1.0000	1.0000	1.0000
CNT	1.2934	1.0000	1.0000	1.0000	1.0000
RSCHPCT	-0.0018	0.0000	0.0000	0.0000	0.0000
PHD	1.8785	1.0000	1.0000	1.0000	1.0000
YRSTCH	-0.0320	8.0000	8.0000	8.0000	8.0000
DSGWB	0.4036	1.0000	1.0000	1.0000	1.0000
DHW	0.3666	0.0000	0.0000	1.0000	1.0000
DRH	0.7902	0.0000	0.0000	1.0000	1.0000
COMPFNL	0.0114	0.0000	0.0000	100.0000	100.0000
RESPOND	0.0123	25.0000	25.0000	25.0000	25.0000
GEN	1.0979	0.7500	0.2500	0.7500	0.2500
GPA	0.0237	350.0000	250.0000	350.0000	250.0000
HSECON	0.0894	1.0000	1.0000	1.0000	1.0000
JOB	0.2492	0.0000	0.0000	0.0000	0.0000
HRSTDY	-0.0778	4.0000	4.0000	4.0000	4.0000
INTIPT	0.0426	25.0000	25.0000	25.0000	25.0000
ENTPRP	-0.0091	25.0000	25.0000	25.0000	25.0000
PUBPRI	1.8011	0.0000	0.0000	0.0000	0.0000
DMACMIC	-0.1011	0.0000	0.0000	0.0000	0.0000
CLAS1	-0.9381	0.0000	0.0000	0.0000	0.0000
CLAS2	-1.0827	0.0000	0.0000	0.0000	0.0000
CLAS3	-0.6315	0.0000	0.0000	0.0000	0.0000
CLAS4	-0.7709	0.0000	0.0000	0.0000	0.0000
CLAS5	-1.2450	0.0000	1.0000	0.0000	1.0000
PREDICTION		18.7916	12.1700	21.0004	14.4669
ACTUAL		17.0000	13.0000	17.0000	13.0000

have taught the same number of years (YRSTCH); intensively use the study guide or workbook (DSGWB); engage the class in instructor-student dialog 25 percent of the class time (RESPOND); have students who completed high school economics (HSECON); have students who are not working in a job (JOB); have students who study an average of four hours a week (HRSTDY); have students with the same interest level in economics (INTIPT); and have equal student ratings for enthusiasm and preparation (ENTPRP). With so many similar characteristics, comparison of mean post-TUCE III scores would seem appropriate. But the real question is not to what extent the teaching environments are similar but how the teaching environments differ between the classes of the two professors and whether the differences come from controllable or uncontrollable factors.

The primary differences in this hypothetical example come from uncontrollable factors. Professor A's macroeconomics classes enroll better students. Professor A's students enter macroeconomics with a higher average pre-TUCE III 30 scores and a higher cumulative GPA. Professor A also teaches more male than female students. In contrast, professor B's students are less capable and predominately female. If professor B cannot control these characteristics, then professor B should not be held responsible for the adverse impact on achievement that these disadvantages exert. Additionally, professor B teaches classes enrolling more than 100 students, while professor A teaches classes enrolling 30 students or less. The large class size negatively influences achievement. When class sizes may be determined by scheduling or other considerations, individual professors may not be able to control the size of a class they will teach.

When the mean values for professor A and professor B are computed using the coefficients from the RANGE regression in Table 4.8, the predicted post-TUCE III 30 score for professor A is 18.80 and for professor B is 12.17. Professor A's actual mean score of 17 is below the national prediction for schools and students exhibiting the same characteristics. Professor B's actual mean score of 13 is above the national prediction of 12.17. Professor B is a better teacher on average when compared to the national average. Professor A's teaching effectiveness is below the national average. Based on this analysis professor B should receive the $1,000 award for teaching effectiveness. Hence, use of mean scores on TUCE III can be very misleading unless used in the context of multiple regression analysis to predict scores for either departments of economics or individual professors.

2. Economics departments and economics instructors should use the RANGE equation in Table 4.8 to identify controllable educational inputs for inclusion in an introductory economics course design. The input inclusion will raise mean student achievement when compared to input exclusion. The controllable educational inputs from Table 4.12 that are statistically significant include: counting the post-TUCE III toward the final course grade (CNT); earned doctorate in economics (PHD); intensive use of homework (DHW); intensive use of non-text reading and handouts (DRH); and the percent of the final exam that is comprehensive (COMPFNL). The hypothetical example, once again, presents the following:

> Professor B is delighted with the $1,000 award for teaching effectiveness. Professor B will apply for tenure in two years and wants to demonstrate to the department chair continued teaching excellence. Professor A is dismayed with performing below the national average. Professor A will have a post-tenure review in two years and wants to demonstrate to the department chair progress in teaching effectiveness. What practical advice can be given to professor A and professor B that will help each one improve teaching?

Professor A and professor B should incorporate controllable educational inputs into their principles of macroeconomics course design. Since both professors have earned doctorates (PHD) and both count the post-TUCE III 30 toward the final course grade (CNT), the remaining three inputs should be adopted. Both professors should intensively use homework (DHW), intensively use non-text reading and handouts (DRH), and give 100 percent comprehensive final exams. When these input changes are made in the course design, the predicted post-test score jumps to 21.00 for professor A and 14.47 for professor B. On average, professor A's students have the capability of achieving much more when the principles of economics course design is managed effectively. The RANGE regression in Table 4.8 should be used to challenge current teaching methodology rather than simply to identify below average instructors.

3. Class size can be increased beyond 100 students when controllable educational inputs are managed effectively. Table 5.4 demonstrates that when CNT, PHD, DHW, DRH and COMPFNL are incorporated into a principles of economics course design, the marginal increment to mean achievement will rise if the inputs have not been previously used. The increment to mean achievement offsets the negative influence of large class size. Gains in mean student achievement and reductions in faculty contract cost can be the benefits from enactment of this class size increase policy.

Section 2: Suggestions for Further Research.

The suggestions for further research are based on the design and results of this investigation. Points 1 through 8 suggest ways to extend the class size issues raised in this study. Points 9 through 13 contain a few suggestions that do not necessarily follow from the research, but that are relevant to the question of increasing student achievement.

1. The effect of class size on outcome measures other than the TUCE III measure of achievement should be investigated. There are multiple desired outcomes in a principles of economics course. The impact of class size on student interest, student attitude, course grade, written communication skills, oral communication skills and other outcomes should be explored. Objective and validated measurement instruments will need to be identified or developed in order to assess each type of outcome. The relative weight, rank or importance of each outcome should also be discussed.

2. The indirect and direct class size models should be tested using a definition of ability other than the pre-TUCE III 30 measure of achievement. Alternate measures of ability could include SAT or ACT scores.

3. The influence of class sizes above 132 students should be tested. The RANGE regression in Table 4.12 excludes class sizes above 132 students. In the TUCE III database sample of 2,942 observations no observations in the 133 through 193 class size range and no observations in the 195 through 231 class size range are available. The observations in class size 194 and class size 232 are dropped from the regression as they represent only two instructors at two institutions.

4. The influence of class size on upper-division economics courses should be determined.

5. Actual rather than hypothetical cost savings associated should be calculated for universities where principles of economics class size is increased beyond 30 students per class. The savings can be generated at single universities or with a university system.

6. Alternative delivery systems for principles of economics education instruction should be evaluated since the negative influence of large class size on achievement may be eliminated. Small classes are not technically efficient. Classroom size may limit the cost reduction that can be realized when introductory economics classes are constrained by the size of a classroom. Perhaps principles of economics can be taught just as effectively but more efficiently in a classroom without walls. The use of interactive cable television or interactive multimedia computer software are two possible ways to deliver economics instruction when adopting the large class size paradigm.

7. The feasibility of requiring students to pay an incremental tuition differential for desired attendance in a class size of 30 students or fewer should be assessed. One policy recommendation of this investigation is that principles of economics class sizes should be increased. However, some students may prefer a small class size. Small classes cost more

and are not technically efficient, especially when large class sizes can be shown to increase achievement. Students desiring the more expensive small class would have to pay extra for this educational service.

A problem of efficient pricing in that marginal cost may not equal marginal benefit could exist in universities that promote small class size policies for introductory economics. Motivated students attending small classes may not need the increased faculty attention associated with the small class. However, the motivated student is required to pay tuition that supports small class size policies. The link between the less motivated students who want more intensive faculty attention and those who must pay for the service is broken. The link should be restored. The efficient pricing problem may be reduced when incremental tuition differentials are established for classes of 30 or fewer students.

The user of the small class service should have to pay for the extra costs associated with offering the small class. The demand in the market for small class education already exists. Private universities characteristically have smaller classes than state universities. Some students are willing to pay higher tuition for the bundle of private university services. Consumers recognize the difference between the bundle of services at state universities versus the bundle of services at private universities.

8. The influence of class size on college principles courses other than economics should be determined. When class size increases are accompanied by the effective management of controllable educational inputs, student achievement in principles of economics can rise. Perhaps the effective management of large classes in college principles courses other than economics can also lead to achievement gains. The influence of class size on student achievement in introductory college courses like management, marketing, sociology, psychology, political science and history should not be ignored. Additional funds can be saved if the conclusions for class size influence on student

achievement in principles of college economics holds true for other introductory college courses.

9. Economics education research efforts should continue investigating the impact of alternate teaching methodologies and the impact of various teaching ancillaries on student achievement in principles of college economics. Some questions may include: Is a Socratic method preferred to a lecture method? Is teaching theory less effective than teaching the application of economics to current events found in business periodicals? How does the use of video or computer software influence achievement, especially today with the current generation of students being more visually oriented and computer literate?

10. One may want to explore the differences between private and public college education in principles of economics. This study has found statistically significant evidence that private universities positively influence achievement when compared to public universities. An investigation that identifies the differences in institutional resources available to students, teaching methods, faculty availability and other factors may be useful.

11. One may want to identify objective ways to document teaching innovations. Accreditation agencies require the objective documentation of faculty scholarly activity. One common and easy way to measure faculty productivity is to count the number of articles published in refereed academic journals. However, a peer reviewed publication by an introductory economics professor does not appear to add to student achievement in the principles course. The definition of faculty scholarly activity may need to be broadened to include teaching. With the renewed emphasis on teaching by the American Assembly of the Collegiate Schools of Business (AACSB) and other accreditation agencies, objective documentation of creative teaching efforts should be explored.

12. The backward and forward linkages to achievement in principles of college economics should be thoroughly investigated. The influence of college courses taken

prior to entering principles of economics should be assessed. The influence of the achievement level in introductory economics on other college courses, the college degree earned or career earnings should be discussed. Investigating the linkages may require economics educators to define what value is added to a college education by taking a course in principles of economics. Outcomes assessment and production function modeling of the educational process may also play a role.

13. Longitudinal databases need to be developed in order to provide more complete information, especially at the classroom level, on the educational process through time. Access to accurate information on the educational process through time will encourage the development of more precise educational production functions. Longitudinal databases should consider providing information on: student demographic data; grades earned in previous or subsequent courses; performance on specific learning activities within classes; faculty demographic data; scores on graduate school standardized exams; student employment profiles after graduation; math and English placement tests; and university exit exams. Comprehensive databases may allow for the identification of new independent variables which can be controlled in order to improve student achievement. Without the development of longitudinal databases, input-output analysis of the educational process is incomplete and theories of learning may simply be resigned to ad hoc models with unspecified ceteris paribus clauses.

Section 3: Conclusion.

Understanding the educational process may not be as easy as understanding how to journalize accounting entries because the educational process is not governed by a simple set of rules. Human behavior is involved in the process of education and modeling human behavior is difficult. Complex manufacturing processes are difficult to understand, but the manufacturing process can be modeled. Forecasting weather is difficult, but still attempted. Determination of interest rates, prices or wages is challenging but not impossible. The complexity of the educational process should not be an excuse to stop attempts at modeling the process.

The production function method is one way to model the educational process. Production functions are used to model input-output relationships for a business, an industry or an economic system. Production functions can also be used to describe the input-output relationships at the strategic business level unit in the university--the classroom. Gains in student achievement can be realized when influential educational inputs are identified and then managed effectively. Continuous improvement in the classroom requires a knowledge of the determinants of production function outcomes.

The determinants of student achievement in principles of college economics are numerous. The theoretical description of production function inputs, such as study effort or class size, is still being explored. Educational researchers have only begun to scratch the surface in the empirical assessment of higher education classroom activities. The multiplicity of potentially influential inputs would suggest that an educational production function be estimated with a multiple regression procedure. Multiple regression is preferred since the significance of specific educational inputs can be more easily understood.

One source of misunderstanding may be the influence of class size on student achievement in principles of college economics. In chapter one, Chester E. Finn, Jr. says that the widely held belief that smaller classes will result in higher student achievement may "... be a very expensive excuse for instructional failure." Without trying to understand the factors that contribute toward increased achievement, class size policy decisions are simply not based on empirical evidence. When class size is reduced and student achievement does not increase, instructional failure results. The failure not only reduces the achievement in principles of economics but may reduce achievement in other courses or occupations that depend on a solid understanding of economic principles. The failure is also costly in the sense that more instructors are needed to teach the increase in the number of sections due to the reduction of class size. In a world of scarcity, resources should not be inefficiently used.

The recommendations of this study are based on empirical evidence. This study has demonstrated that average student achievement can be increased by effectively managing controllable educational inputs. This study also has demonstrated that larger class sizes reduce average achievement, but by increasing class size while at the same time incorporating controllable educational inputs into a course design that previously did not use the input, mean student achievement can increase. The increase in class size may reduce the number of faculty needed to teach principles of college economics. The reduction in faculty will result in contract salary and benefit savings or lower budget expenditures when faculty members can be re-deployed. The result is that many stakeholders may benefit - students, faculty, taxpayers, courses requiring principles of economics as a pre-requisite and occupations where a knowledge of introductory economics is helpful. The bottom line is that colleges can increase both class size and mean student achievement, while simultaneously lowering cost.

Appendices

Appendix A

Glossary

Glossary

AACSB : American Assembly of the Collegiate Schools of Business

ACT : American College Test

Ancillary teaching aid : Non-textbook teaching aids like homework or handouts

EA : Explicit application question category on the TUCE III exams

GPA : Cumulative grade point average

GSI : Graduate student instructors

IA : Implicit application question category on the TUCE III exams

Intensive use : Course assignment is required and graded

NAEE : National Association of Economic Education

OLS : Ordinary least squares

RU : Recognition and understanding question category on the TUCE III exams

SAT : Scholastic Aptitude Test

SET : Student evaluation of teaching

STAR : Student/teacher achievement ratio

TUCE : Test for Understanding of College Economics

TUCE III 30 : The first 30 questions of the third version of TUCE

Appendix B

List of Variables

List of Variables

Variable Name		Description
TTOT	:	Number of Post-TUCE III 30 correct answers.
TRU	:	Number of RU Post-TUCE III 30 correct answers.
TEAIA	:	Number of EA and IA Post-TUCE III 30 correct answers.
PTOT	:	Number of Pre-TUCE III 30 correct answers.
PRU	:	Number of RU Pre-TUCE III 30 correct answers.
PEAIA	:	Number of EA and IA Pre-TUCE III 30 correct answers.
DOCI	:	Code 1 if doctoral institution and code 0 otherwise.
COMPI	:	Code 1 if comprehensive institution and code 0 otherwise.
LIBI	:	Code 1 if liberal arts institution and code 0 otherwise.
TWOI	:	Code 1 if two-year institution and code 0 otherwise.
ENG	:	Code 1 if instructor's native language is English and code 0 otherwise.
CNT	:	Code 1 if Post-TUCE III 30 counts toward final course grade and code 0 otherwise.
TCHPCT	:	Percent of time instructor spends in teaching.
RSCHPCT	:	Percent of time instructor spends in research.
PHD	:	Code 1 if instructor has earned a doctorate and code 0 otherwise.
YRSTCH	:	Number of years the instructor has been teaching.
DSGWB	:	Code 1 if study guide or workbook is used intensively and code 0 otherwise.
DHW	:	Code 1 if homework is used intensively and code 0 otherwise.

DRH : Code 1 if non-text readings and handouts are used intensively and code 0 otherwise.

COMPFNL : Percent of final exam which is comprehensive.

PCTLEC : Percent of class time instructor spends in lecture.

PCTINSR : Percent of class time instructor responds to student questions.

PCTSTUR : percent of class time students respond to instructor questions.

RESPOND : [(PCTINSR + PCTSTUR) / PCTLEC].

GEN : Code 1 if student gender is male and code 0 otherwise.

GPA : Cumulative GPA of student multiplied by 100.

HSECON : Code 1 if student has completed a high school economics course and code 0 otherwise.

JOB : Code 1 if the student is working at a job and code 0 otherwise.

HRSTDY : Hours student spends studying economics.

INTRST : Student interest in economics.

IMPT : Student assessment of the importance of economics.

INTIPT : INTRST + IMPT.

ENTHU : Enthusiasm of instructor - student assessed.

PREP : Preparation of instructor - student assessed.

STAND : Rigor of instructor's grading standards - student assessed.

OTE : Overall teaching effectiveness - student assessed.

ENTPRP : ENTHU + PREP.

PUBPRI : Code 1 if a private institution and 0 otherwise.

DMACMIC : Code 1 if principle of microeconomics and code 0 otherwise.

CLASIZE : Class size.

CLASQR : The square of class size.

DIFSQR : CLASQR - CLASIZE

LOGCLAS : The natural log of class size.

PREMAX : The pre-TUCE III 30 maximum score per course section.

PREMIN : The pre-TUCE III 30 minimum score per course section.

PRESTDV : The pre-TUCE III 30 standard deviation score per course section.

RANGE : The range of pre-TUCE III 30 scores per course section.

SKEW : The skewness of pre-TUCE III 30 scores per course section.

KURT : The kurtosis of pre-TUCE III 30 scores per course section.

PCTPRE : The percent of students taking the pre-TUCE III 30 out of the total number of students in a course section.

CLAS0 : Code 1 if class size range is 0-30 and code 0 otherwise.

CLAS1 : Code 1 if class size range is 31-40 and code 0 otherwise.

CLAS2 : Code 1 if class size range is 41-50 and code 0 otherwise.

CLAS4 : Code 1 if class size range is 51 - 75 and code 0 otherwise.

CLAS4 : Code 1 if class size range is 76 - 100 and code 0 otherwise.

CLAS5 : Code 1 if class size range is 101 - 132 and code 0 otherwise.

CLAS6 : Code 1 if class size range is 133 - 232 and code 0 otherwise.

Appendix C

Additional Tables

Table C.1 Comparison of Class Size Investigation Results.

	Class Size Study - Key to Headings in Notes					
Independent Variable	2	3	4	5	6	7
Sex						
Male	+X			+X5		+X2
Female		-X2N			-X1	
College GPA	+X5S					+X2
High School GPA				+X1		+X2
SAT						
Composite						
Math						+/-X2
Verbal						+X2
ACT	+X5S					
Pre-Test		+X2				
Pre-TUCE I	+X5S					
High School Economics	-X				+X1	
College Economics	+X					
Class - College Freshman, etc.	+X	+X				
Class - High School Freshman, etc.				-X		
Race - White				+X1		
Student Course Grade	+XS5					
Elective versus Required Course	-X				-X1	
College Credit Hours of Economics					+X	
College Math	-X					
High School English Grade					+X1	
High School Math Grade					+X1	
High School Rank						+X2
Expected Student Course Grade	+X					
Class Size		-/+X2N	X	-X1	X *	
Section Size	-X					
Small Class Dummy Variable						+/-X
Class Size Squared			X			
Class Size - Interaction Term			X			
Class Size Squared - Interaction			X			
Class Type - Course Content			X			
Major						
Math/Science	+X					
Engineering		-/+X				
Commerce/Business		-/+X				
Science		-/+X2N				
Time Spent Studying	+X5	+X2				
Pre-CTA	+X5					
Post-CTA	+XS5					
Student Interest in Course	+XS5					+/-X
Student Desire to Take Course						
Importance of Course to Student	+XS5					
Should Course be Required	-X					
Teacher Arouses Interest		+/-X				
Teacher Attitude				+X1		+X
Teacher Rating						+X

Table C.1 Continued.

	Class Size Study - Key to Headings in Notes					
Independent Variable	2	3	4	5	6	7
Course Rating						
Student Reads Regularly						
Newspaper	-X					
Magazine	-X					
News Business Section	+X					
Wall Street Journal	+X5					
Business Week	-X					
Fortune	+X					
Investment News Letter	-X					
Parent Attended College				+X		
Student Plans to Attend College				+X1		
Years Between High School & College					+X1	
United States Citizen						-X2
State Resident						+/-X
Minority						-X2
Age						+/-X2
Honors Student						+X2
Instructor Variables:						
Department Evaluation of	-XS5					
Age	+X					
Highest Degree	-X			+X1		
Credit Hours of Economics	+X					
Years Teaching	-X			-X		
Years Teaching Principles	+XS5					
Pre-TUCE I	+XS5					
Pre-CTA	+X					
Uses Audio/Visual Aids				+X		
Uses Computer Software				-X		
Uses Supplemental Texts				-X5		
Consults with NCEE				+X1		
Year Class was Taken					-X1	
Semester Class was Taken						X
Which Faculty Member Taught Class					X *	
Is this an Experimental Section	-X					
Model Utility						
Adjusted R-Squared	0.43 S0.54		ALL< =0.01	0.42		0.34 through 0.45
R-Squared		0,48 N0.27	ALL< =0.02		0.35	
F-Statistic		19.12 N8.6	MOST <1.5			

Table C.1 Continued.

NOTES:

Column Headings:
1. Levin (1967) - Did not use a linear regression procedure. Not included in Table C.1
2. Lewis and Dahl (1972)
3. Crowley and Wilton (1974)
4. Williams, et. al. (1985)
5. Lopus (1990)
6. Myatt and Waddell (1990)
7. Gramlich and Greenlee (1993)

Legend:

X - Indicates the independent variable was use in this investigation.
\+ - Variable coefficient is positive in regression.
\- - Variable coefficient is negative in regression.
/ - Indicates two regressions used. Sign preceding / is result of first regression and sign following is result of second.
* - Sign of Coefficient was not reported.
S - Step-down regression used.
1 - Variable is significant at the one percent level.
5 - Variable is significant at the five percent level.
2 - Variable has a t-ratio of two or more.
N - Post-test was given under non-exam conditions.

Comments on Column Headings:

2. Lewis and Dahl (1972)
   If S5 then variable is significant in second step-down regression.
   If 5S then variable is significant in primary and in the secondary step-down regression.
3. Crowley and Wilton (1974)
   Use of N indicates variable was non-significant when post-test was given under non-exam conditions.
6. Myatt and Waddell (1990)
   Regression reported includes students with and without high school economics.

Table C.2 Comparison of Five Achievement Models.

Variable	TTOT Coefficient	VA Coefficient	T/PTOT Coefficient	PCTVA Coefficient	GAPC Coefficient
Constant	-2.8144	-6.3623	-4.9823	-0.3142	-0.4278
	(-3.1)	(-7.6)	(-6.3)	(-2.5)	(-9.9)
DOCI	1.3448	0.6608	0.9269	0.0229	0.0384
	(3.8)	(2.0)	(2.9)	(0.4)	(2.3)
COMPI	-0.4929	-0.5256	-0.5129	-0.1089	-0.0304
	(-1.4)	(-1.6)	(-1.6)	(-1.6)	(-1.8)
LIBI	0.6418	0.6379	0.6395	0.0047	0.0394
	(1.4)	(1.5)	(1.6)	(0.1)	(1.8)
PTOT			0.6110		
			(29.6)		
SEPRSX	0.3839	0.3124	0.3402	0.0126	0.0173
	(0.9)	(0.8)	(0.9)	(0.2)	(0.9)
ENG	0.9323	0.2400	0.5093	-0.0723	0.0213
	(2.6)	(0.7)	(1.5)	(-1.0)	(1.2)
CNT	0.8827	1.1702	1.0584	0.1771	0.0596
	(4.1)	(5.9)	(5.6)	(5.0)	(5.9)
RSCHPCT	0.0104	-0.0018	0.0029	-0.0005	0.0002
	(1.8)	(-0.4)	(0.6)	(-0.7)	(0.6)
PHD	1.1098	0.9692	1.0239	0.1204	0.0523
	(4.0)	(3.8)	(4.2)	(3.1)	(4.0)
YRSTCH	-0.0048	-0.0118	-0.0091	-0.0010	-0.0006
	(-0.4)	(-1.0)	(-0.8)	(-0.5)	(-0.9)
DSGWB	0.0088	0.3257	0.2024	0.0120	0.0136
	(0.0)	(1.3)	(0.8)	(0.3)	(1.1)
DHW	0.4611	0.3358	0.3846	0.0508	0.0188
	(2.3)	(1.8)	(2.2)	(1.7)	(2.0)
DRH	0.6883	0.5723	0.6174	0.0561	0.0355
	(2.8)	(2.5)	(2.9)	(1.5)	(3.1)
COMPFNL	0.0108	0.0119	0.0114	0.0009	0.0007
	(4.7)	(5.6)	(5.7)	(2.4)	(6.0)
RESPOND	0.0499	-0.0294	0.0015	-0.0074	-0.0003
	(1.5)	(-1.0)	(0.1)	(-2.5)	(-0.2)
GEN	1.6341	0.7913	1.1191	0.0451	0.0552
	(9.7)	(5.0)	(7.4)	(1.7)	(6.9)
GPA	0.0328	0.0196	0.0248	0.0016	0.0012
	(19.7)	(14.5)	(17.3)	(6.9)	(16.9)
HSECON	-0.0790	0.0127	-0.0230	0.0043	0.0006
	(-0.5)	(0.1)	(-0.2)	(0.2)	(0.1)
JOB	0.3278	0.0059	0.1311	0.0024	0.0027
	(1.9)	(0.0)	(0.8)	(0.1)	(0.3)
HRSTDY	-0.1068	-0.0394	-0.0656	-0.0032	-0.0031
	(-2.9)	(-1.2)	(-2.0)	(-0.6)	(-1.7)

Table C.2 Continued.

Variable	TTOT Coefficient	VA Coefficient	T/PTOT Coefficient	PCTVA Coefficient	GAPC Coefficient
INTIPT	0.0526	0.0380	0.0437	0.0040	0.0021
	(9.0)	(7.2)	(8.6)	(4.7)	(7.7)
ENTPRP	-0.0105	-0.0001	-0.0041	0.0007	-0.0002
	(-1.5)	(0.0)	(-0.6)	(0.6)	(-0.5)
PUBPRI	3.0809	1.1004	1.8708	0.0218	0.0881
	(11.0)	(4.0)	(7.3)	(0.5)	(6.4)
DMACMIC	1.0585	-0.3156	0.2189	-0.1130	0.0074
	(6.0)	(-1.9)	(1.4)	(-4.1)	(0.9)
Observations.	2,942	2,942,	2,942	2,942	2,942
R-squared	0.35	0.18	0.49	0.06	0.25
Adjusted R-squared	0.34	0.18	0.48	0.06	0.25
Regression F-test [DF]	68.22	28.56	116.26	8.55	42.51
	[23; 2,918]	[23; 2,918]	[24; 2,917]	[23; 2,918]	[23; 2,918]
Prob of F-test	0.00	0.00	0.00	0.00	0.00
Breusch-Pagan	76.52	41.43	46.82	743.76	60.153
[DF]	[23]	[23]	[24]	[23]	[23]
Correctly Predicted Coefficient Signs	16	16	17	17	16
Statistically Significant Coefficients	14	10	14	8	11
Correct Sign & Statistically Significant	11	9	11	5	10

Notes: The numbers within the parentheses are the t-ratios derived from White's (1980) heteroscedasticity corrected variance-covariance matrix. The abbreviations for the five achievement models are: absolute achievement (TTOT); absolute improvement (VA); absolute improvement with the pre-test as an independent variable (T/PTOT); percentage improvement (PCTVA); and gap closing (GAPC).

Table C.3 Comparison of Direct Class Size Influence Models

Variable	INSTRUCT Coefficient	INTERACT Coefficient	FIXED Coefficient
Constant	-4.9490	-3.8680	-6.3530
	(-6.2)	(-4.8)	(-6.8)
DOCI	0.5394	0.8834	0.6500
	(1.6)	(2.6)	(1.9)
COMPI	-0.6258	-0.6600	-0.5861
	(-2.0)	(-2.1)	(-1.8)
LIBI	0.5686	0.6020	0.5811
	(1.4)	(1.5)	(1.4)
PTOT	0.6155	0.6138	0.6140
	(30.0)	(29.9)	(29.8)
SEPRSX	0.2219	0.7234	0.1867
	(0.6)	(1.9)	(0.5)
ENG	0.2639	0.2177	0.3860
	(0.8)	(0.7)	(1.2)
CNT	1.1730	1.4094	1.0689
	(6.2)	(7.2)	(5.6)
RSCHPCT	0.0000	0.0022	0.0009
	(0.0)	(0.5)	(0.2)
PHD	0.9688	1.1809	0.9743
	(4.0)	(4.9)	(4.0)
YRSTCH	-0.0211	-0.0200	-0.0151
	(-1.9)	(-1.8)	(-1.4)
DSGWB	0.3896	0.2835	0.3181
	(1.6)	(1.1)	(1.3)
DHW	0.4285	0.4741	0.4162
	(2.4)	(2.7)	(2.4)
DRH	0.7308	0.6986	0.7030
	(3.4)	(3.2)	(3.2)
COMPFNL	0.0106	0.0109	0.0108
	(5.2)	(5.4)	(5.3)
RESPOND	0.0138	0.0035	0.0108
	(0.5)	(0.1)	(0.4)
GEN	1.1178	1.1075	1.1188
	(7.4)	(7.4)	(7.4)
GPA	0.0246	0.0240	0.0248
	(17.1)	(16.9)	(17.2)
HSECON	-0.0017	-0.0221	-0.0101
	(0.0)	(-0.1)	(-0.1)
JOB	0.1730	0.2209	0.1487
	(1.1)	(1.4)	(1.0)
HRSTDY	-0.0631	-0.0672	-0.0629
	(-2.0)	(-2.1)	(-1.9)

Table C.3 Continued.

Variable	INSTRUCT Coefficient	INTERACT Coefficient	FIXED Coefficient
INTIPT	0.0452	0.0454	0.0444
	(8.9)	(8.9)	(8.7)
ENTPRP	-0.0071	-0.0069	-0.0059
	(-1.1)	(-1.1)	(-0.9)
PUBPRI	1.9776	1.9279	1.9381
	(7.7)	(7.5)	(7.5)
DMACMIC	0.1951	0.1325	0.2031
	(1.2)	(0.8)	(1.3)
CLASIZE	0.0092	-0.0237	
	(4.9)	(-3.7)	
CLASQR		0.0001	
		(5.4)	
LOGCLAS			0.4380
			(2.8)
Observations.	2,942	2,942	2,942
R-squared	0.49	0.50	0.49
Adjusted R-squared	0.48	0.49	0.49
Regression F-test	113.50	111.28	112.24
[DF]	[25; 2,916]	[26; 2,915]	[25; 2,916]
Prob of F-test	0.00	0.00	0.00
Breusch-Pagan	47.21	47.25	47.30
[DF]	[25]	[26]	[25]

Notes: The numbers within the parentheses are the t-ratios derived from White's (1980) heteroscedasticity corrected variance-covariance matrix. The abbreviations for the various direct class size models are as follows: instructional overhead (INSTRUCT); interaction (INTERACT); and fixed resource distribution (FIXED).

Table C.4 Comparison of Indirect Class Size Influence Models.

	HETERO	PACE	GROUP
Variable	Coefficient	Coefficient	Coefficient
Constant	-7.3332	-8.6573	-8.1328
	(-8.0)	(-9.5)	(-9.3)
DOCI	0.6043	0.5390	0.4124
	(1.8)	(1.7)	(1.2)
COMPI	-0.5154	-0.5011	-0.6012
	(-1.6)	(-1.6)	(-1.9)
LIBI	0.6124	0.4516	0.2143
	(1.5)	(1.1)	(0.5)
PTOT	0.5938	0.5405	0.5543
	(28.1)	(24.1)	(25.1)
SEPRSX	0.1449	0.1708	0.0255
	(0.4)	(0.5)	(0.1)
ENG	0.2331	0.1616	-0.0349
	(0.7)	(0.5)	(-0.1)
CNT	0.9481	1.1651	1.0977
	(4.9)	(6.2)	(5.7)
RSCHPCT	-0.0022	-0.0056	-0.0092
	(-0.5)	(-1.2)	(-1.9)
PHD	1.0427	1.1020	1.1619
	(4.3)	(4.6)	(4.8)
YRSTCH	-0.0092	-0.0141	-0.0143
	(-0.9)	(-1.3)	(-1.3)
DSGWB	0.4479	0.5452	0.6161
	(1.8)	(2.3)	(2.5)
DHW	0.2783	0.2206	0.1727
	(1.6)	(1.3)	(1.0)
DRH	0.3402	0.4416	0.4350
	(1.5)	(2.0)	(2.0)
COMPFNL	0.0141	0.0118	0.0140
	(6.7)	(5.9)	(6.9)
RESPOND	-0.0084	-0.0057	-0.0145
	(-0.3)	(-0.2)	(-0.6)
GEN	1.1076	1.1111	1.1138
	(7.4)	(7.5)	(7.5)
GPA	0.0244	0.0240	0.0240
	(16.9)	(16.6)	(16.6)
HSECON	0.0335	0.0308	0.0403
	(0.2)	(0.2)	(0.3)
JOB	0.0929	0.0507	0.0879
	(0.6)	(0.3)	(0.6)
HRSTDY	-0.0705	-0.0868	-0.0871
	(-2.2)	(-2.7)	(-2.7)

Table C.4 Continued.

Variable	HETERO Coefficient	PACE Coefficient	GROUP Coefficient
INTIPT	0.0445	0.0446	0.0445
	(8.8)	(9.0)	(8.9)
ENTPRP	-0.0017	-0.0011	-0.0015
	(-0.3)	(-0.2)	(-0.2)
PUBPRI	1.6770	1.0176	1.1337
	(6.2)	(3.6)	(4.1)
DMACMIC	-0.0615	-0.3902	-0.4195
	(-0.4)	(-2.3)	(-2.4)
RANGE	0.0985		
	(2.5)		
PRESTDV	0.4873		
	(2.5)		
PREMIN			0.1741
			(2.9)
DEVFAVG		0.0493	
		(0.7)	
PREAVG		0.5009	
		(7.6)	
SKEW		0.4112	
		(2.1)	
PREMAX			0.2655
			(6.9)
KURT			-0.1525
			-(2.0)
Observations.	2,942	2,942	2,942
R-squared	0.49	0.50	0.50
Adjusted R-squared	0.49	0.50	0.50
Regression F-test [DF]	109.32	108.54	108.56
	[26; 2,915]	[27; 2,914]	[27; 2,914]
Prob of F-test	0.00	0.00	0.00
Breusch-Pagan	52.60	57.20	51.90
[DF]	[26]	[27]	[27]

Notes: The numbers within the parentheses are the t-ratios derived from White's (1980) heteroscedasticity corrected variance-covariance matrix. The abbreviations for the various indirect class size models are as follows: heterogeneity (HETERO); instructional pacing (PACE); and ability grouping (GROUP). PREMIN is not included in the empirical model since it is perfectly correlated with DEVFAVG and PREAVG.

Table C.5 Direct and Indirect Class Size Variable Regressions

Variable	CLASIZE Coefficient	CLASQR Coefficient	LOGCLAS Coefficient	RANGE Coefficient	PRESTDV Coefficient	PREMIN Coefficient
Constant	-4.9490	-4.5739	-6.3530	-6.5624	-7.0993	-5.5214
	(-6.2)	(-5.8)	(-6.8)	(-7.7)	(-7.8)	(-6.9)
DOCI	0.5394	0.5729	0.6500	0.5768	0.7778	0.9892
	(1.6)	(1.8)	(1.9)	(1.8)	(2.4)	(3.1)
COMPI	-0.6258	-0.6618	-0.5861	-0.5235	-0.5058	-0.5403
	(-2.0)	(-2.1)	(-1.8)	(-1.6)	(-1.6)	(-1.7)
LIBI	0.5686	0.5647	0.5811	0.5892	0.6492	0.3860
	(1.4)	(1.4)	(1.4)	(1.5)	(1.6)	(1.0)
PTOT	0.6155	0.6159	0.6140	0.6040	0.5907	0.5849
	(30.0)	(30.1)	(29.8)	(29.2)	(28.0)	(27.1)
SEPRSX	0.2219	0.3671	0.1867	0.0304	0.3537	0.4467
	(0.6)	(1.0)	(0.5)	(0.1)	(1.0)	(1.2)
ENG	0.2639	0.1953	0.3860	0.2396	0.3499	0.4428
	(0.8)	(0.6)	(1.2)	(0.7)	(1.1)	(1.3)
CNT	1.1730	1.2781	1.0689	0.9204	1.0269	1.2031
	(6.2)	(6.7)	(5.6)	(4.8)	(5.5)	(6.3)
RSCHPCT	0.0000	0.0001	0.0009	-0.0014	-0.0007	-0.0005
	(0.0)	(0.0)	(0.2)	(-0.3)	(-0.2)	(-0.1)
PHD	0.9688	1.0291	0.9743	1.0079	1.0712	1.1387
	(4.0)	(4.3)	(4.0)	(4.2)	(4.4)	(4.6)
YRSTCH	-0.0211	-0.0233	-0.0151	-0.0082	-0.0103	-0.0118
	(-1.9)	(-2.2)	(-1.4)	(-0.8)	(-0.9)	(-1.1)
DSGWB	0.3896	0.3938	0.3181	0.4583	0.3271	0.1778
	(1.6)	(1.6)	(1.3)	(1.8)	(1.3)	(0.7)
DHW	0.4285	0.4535	0.4162	0.3477	0.2523	0.2899
	(2.4)	(2.6)	(2.4)	(2.0)	(1.4)	(1.7)
DRH	0.7308	0.7443	0.7030	0.3983	0.4026	0.7106
	(3.4)	(3.4)	(3.2)	(1.8)	(1.8)	(3.3)
COMPFNL	0.0106	0.0105	0.0108	0.0129	0.0141	0.0114
	(5.2)	(5.2)	(5.3)	(6.4)	(6.7)	(5.7)
RESPOND	0.0138	0.0130	0.0108	-0.0002	-0.0126	-0.0073
	(0.5)	(0.5)	(0.4)	(0.0)	(-0.5)	(-0.3)
GEN	1.1178	1.1141	1.1188	1.1148	1.1050	1.1260
	(7.4)	(7.4)	(7.4)	(7.4)	(7.3)	(7.5)
GPA	0.0246	0.0243	0.0248	0.0245	0.0244	0.0246
	(17.1)	(17.0)	(17.2)	(17.0)	(17.0)	(17.2)
HSECON	-0.0017	-0.0040	-0.0101	0.0249	0.0174	-0.0371
	(0.0)	(0.0)	(-0.1)	(0.2)	(0.1)	(-0.2)
JOB	0.1730	0.1983	0.1487	0.1188	0.0824	0.1281
	(1.1)	(1.3)	(1.0)	(0.8)	(0.5)	(0.8)
HRSTDY	-0.0631	-0.0640	-0.0629	-0.0669	-0.0722	-0.0765
	(-2.0)	(-2.0)	(-1.9)	(-2.1)	(-2.2)	(-2.3)

Table C.5 Continued.

	CLASIZE	CLASQR	LOGCLAS	RANGE	PRESTDV	PREMIN
Variable	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
INTIPT	0.0452	0.0456	0.0444	0.0447	0.0440	0.0431
	(8.9)	(9.0)	(8.7)	(8.8)	(8.7)	(8.5)
ENTPRP	-0.0071	-0.0077	-0.0059	-0.0035	-0.0008	-0.0035
	(-1.1)	(-1.2)	(-0.9)	(-0.5)	(-0.1)	(-0.5)
PUBPRI	1.9776	1.9837	1.9381	1.8428	1.5878	1.4479
	(7.7)	(7.7)	(7.5)	(7.2)	(6.0)	(5.3)
DMACMIC	0.1951	0.1687	0.2031	0.0013	-0.0027	0.0571
	(1.2)	(1.1)	(1.3)	(0.0)	(0.0)	(0.4)
CLASIZE	0.0092					
	(4.9)					
CLASQR		0.0000				
		(6.4)				
LOGCLAS			0.4380			
			(2.8)			
RANGE				0.1498		
				(4.7)		
PRESTDV					0.7864	
					(4.8)	
PREMIN						0.2866
						(4.8)
DEVFAVG						
PREAVG						
SKEW						
PREMAX						
KURT						
Observations.	2,942	2,942	2,942	2,942	2,942	2,942
R-squared	0.49	0.50	0.49	0.49	0.49	0.49
Adjusted R-squared	0.48	0.49	0.49	0.49	0.49	0.49
Regression F-test [DF]	113.50 [25; 2,916]	114.69 [25; 2,916]	112.24 [25; 2,916]	113.30 [25; 2,916]	113.22 [25; 2,916]	113.49 [25; 2,916]
Prob of F-test	0.00	0.00	0.00	0.00	0.00	0.00

Table C.5 Continued.

Variable	CLASIZE Coefficient	CLASQR Coefficient	LOGCLAS Coefficient	RANGE Coefficient	PRESTDV Coefficient	PREMIN Coefficient
Breusch-Pagan [DF]	47.21 [25]	49.99 [25]	47.30 [25]	52.26 [25]	50.95 [25]	40.57 [25]

Table C.5 Continued.

	DEVFAVG	PREAVG	SKEW	PREMAX	KURT
Variable	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
Constant	-6.0128	-8.1840	-5.0062	-8.0947	-5.4182
	(-7.0)	(-9.3)	(-6.3)	(-9.3)	(-6.7)
DOCI	0.7220	0.5825	0.9276	0.3961	0.9174
	(2.2)	(1.8)	(2.9)	(1.2)	(2.9)
COMPI	-0.4931	-0.5165	-0.5116	-0.5545	-0.4877
	(-1.5)	(-1.6)	(-1.6)	(-1.8)	(-1.5)
LIBI	0.7698	0.4900	0.6317	0.3340	0.5913
	(1.9)	(1.2)	(1.6)	(0.8)	(1.5)
PTOT	0.5993	0.5402	0.6115	0.5765	0.6149
	(28.7)	(24.0)	(29.5)	(27.3)	(29.6)
SEPRSX	0.1734	0.1561	0.3466	-0.0846	0.2753
	(0.5)	(0.4)	(0.9)	(-0.2)	(0.7)
ENG	0.4491	0.2618	0.4994	0.0003	0.4765
	(1.4)	(0.8)	(1.5)	(0.0)	(1.4)
CNT	0.9941	1.1661	1.0598	0.9542	1.0018
	(5.3)	(6.3)	(5.6)	(5.0)	(5.2)
RSCHPCT	0.0024	-0.0043	0.0028	-0.0073	0.0021
	(0.5)	(-0.9)	(0.6)	(-1.5)	(0.4)
PHD	0.9733	1.1102	1.0241	1.0974	1.0328
	(4.0)	(4.6)	(4.2)	(4.6)	(4.2)
YRSTCH	-0.0096	-0.0150	-0.0089	-0.0100	-0.0067
	(-0.9)	(-1.4)	(-0.8)	(-0.9)	(-0.6)
DSGWB	0.3700	0.5294	0.1994	0.6087	0.2044
	(1.5)	(2.2)	(0.8)	(2.5)	(0.8)
DHW	0.3845	0.2213	0.3847	0.2402	0.3857
	(2.2)	(1.3)	(2.2)	(1.4)	(2.2)
DRH	0.4844	0.4848	0.6160	0.3326	0.5598
	(2.2)	(2.2)	(2.8)	(1.6)	(2.6)
COMPFNL	0.0115	0.0115	0.0115	0.0139	0.0115
	(5.7)	(5.8)	(5.7)	(6.9)	(5.7)
RESPOND	0.0064	-0.0027	0.0010	-0.0089	0.0012
	(0.2)	(-0.1)	(0.0)	(-0.3)	(0.0)
GEN	1.1059	1.1019	1.1206	1.1179	1.1266
	(7.3)	(7.4)	(7.4)	(7.5)	(7.5)
GPA	0.0246	0.0241	0.0248	0.0242	0.0248
	(17.1)	(16.8)	(17.2)	(16.7)	(17.3)
HSECON	0.0054	0.0153	-0.0219	0.0448	-0.0230
	(0.0)	(0.1)	(-0.1)	(0.3)	(-0.2)
JOB	0.0929	0.0418	0.1333	0.1080	0.1465
	(0.6)	(0.3)	(0.9)	(0.7)	(0.9)
HRSTDY	-0.0678	-0.0892	-0.0652	-0.0773	-0.0633
	(-2.1)	(-2.7)	(-2.0)	(-2.4)	(-1.9)

Table C.5 Continued.

Variable	DEVFAVG Coefficient	PREAVG Coefficient	SKEW Coefficient	PREMAX Coefficient	KURT Coefficient
INTIPT	0.0445	0.0444	0.0437	0.0448	0.0437
	(8.8)	(8.9)	(8.6)	(8.9)	(8.6)
ENTPRP	-0.0035	-0.0016	-0.0041	-0.0025	-0.0045
	(-0.6)	(-0.3)	(-0.6)	(-0.4)	(-0.7)
PUBPRI	1.8342	1.0613	1.8679	1.4547	1.9064
	(7.1)	(3.8)	(7.3)	(5.7)	(7.4)
DMACMIC	0.0854	-0.3545	0.2191	-0.2861	0.2231
	(0.5)	(-2.1)	(1.4)	(-1.7)	(1.4)
CLASIZE					
CLASQR					
LOGCLAS					
RANGE					
PRESTDV					
PREMIN					
DEVFAVG	0.2241				
	(3.5)				
PREAVG		0.4941			
		(8.3)			
SKEW			0.0487		
			(0.3)		
PREMAX				0.2504	
				(7.8)	
KURT					0.1394
					(2.1)
Observations.	2,942	2,942	2,942	2,942	2,942
R-squared	0.49	0.50	0.49	0.50	0.49
Adjusted R-squared	0.49	0.50	0.48	0.49	0.49
Regression F-test [DF]	112.49 [25; 2.916]	116.97 25; 2,916]	111.58 [25; 2,916]	116.19 [25; 2,916]	111.89 [25; 2,916]
Prob of F-test	0.00	0.00	0.00	0.00	0.00

Table C.5 Continued

Variable	DEVFAVG Coefficient	PREAVG Coefficient	SKEW Coefficient	PREMAX Coefficient	KURT Coefficient
Breusch-Pagan	50.19	48.54	49.30	55.23	48.33
[DF]	[25]	[25]	[25]	[25]	[25]

Notes: The numbers within the parentheses are the t-ratios derived from White's (1980) heteroscedasticity corrected variance-covariance matrix. Each regression includes the determinants of achievement identified in section two of chapter three. The label over each column identifies the sole class size influence variable which is added to the determinants of achievement regression using the absolute improvement model.

Table C.6. Number of Observations per Class Size

Class Size	Observations	Class Size	Observations
10	4	50	19
13	5	51	9
14	2	52	25
15	5	53	26
16	20	54	28
18	7	55	24
20	14	56	28
21	20	57	2
23	11	58	32
24	7	59	14
25	18	60	26
27	39	61	36
28	19	62	69
29	61	64	29
30	75	69	27
31	10	72	10
32	100	73	44
33	83	85	41
34	40	91	93
35	38	93	83
36	16	96	163
37	122	97	152
38	114	98	32
39	56	100	38
40	48	101	35
41	105	106	57
42	23	108	45
43	45	109	29
44	49	113	70
46	45	122	56
47	75	132	51
48	49	194	105
49	47	232	72

Table C.7 Comparison of National Percentile Rank for Students Completing Post-TUCE III 30

Correct Answers	Macro Post-test	Micro Post-test
30	99	99
29	99	99
28	99	98
27	98	97
26	97	96
25	96	95
24	94	92
23	92	89
22	90	85
21	87	80
20	84	76
19	80	72
18	75	67
17	70	61
16	65	55
15	58	50
14	52	44
13	45	38
12	38	31
11	30	25
10	22	19
9	15	14
8	9	9
7	5	5
6	2	3
5	1	1
4	1	1
3	1	1
2	1	1
1	-	-
Mean Correct	14.31	15.36

Source: TUCE III Examiners Manual (1991); pages 17 and 21.

Bibliography

Bibliography

Access for Windows. 1992. Redmond, WA: Microsoft Corporation.

Adams, John L. and William Becker. "Course Withdrawals: A Probit Model and Policy Recommendations." Research in Higher Education. (31:6 1990): 519-538.

Allen, Robert F. "A Note on the Economic Meaning of Correlations between Teaching and Research Outputs." Journal of Economic Education. (Spring 1980): 49-54.

Allison, E. 1982. "Education Production Function for an Introductory Economics Course." In Rendigs Fels and John Siegfried (eds.), Research on Teaching College Economics: Selected Readings. New York: Joint Council on Economic Education, 171-194.

Becker, William, Robert Highsmith, Peter Kennedy and William Walstad. "An Agenda for Research on Economic Education in Colleges and Universities. Journal of Economic Education. (Summer 1991): 241-250. Again in American Economic Review Papers and Proceedings. (May 1991): 26-31.

________. "Behavior and Productivity Implications of Institutional and Project Funding of Research: Comment." American Agricultural Economics Association. (August 1982): 595-598.

________. "The University Professor as a Utility Maximizer and producer of Learning, Research, and Income." Journal of Human Resources. (Winter 1975): 107-115.

________. "Professorial Behavior Given a Stochastic Reward Structure." American Economic Review. (December 1979): 1,010-1,017.

________. "Economic Education Research: Part II, New Directions in Theoretical Model Building." Journal of Economic Education. (Spring 1983A): 4-10.

________. "Economic Education Research: Part III, Statistical Estimation Methods." Journal of Economic Education. (Summer 1983B): 4-15.

_______ and William B. Walstad. "Data Loss From Pretest To Posttest As A Sample Selection Problem." Review of Economics and Statistics. (February 1990): 184-188.

Belsley, David, Edwin Kuh and Roy Welsch. 1980. Regression Diagnostics, Identifying Influential Data and Sources of Collinearity. New York: Wiley.

Blinder, Alan S. "Research in Economic Education and the Teaching of Economics." Journal of Economic Education. (Summer 1991): 251- 254.

Bloom, Benjamin S., ed. 1956. Taxonomy of Educational Objectives: The Classification of Educational Goals. New York: David McKay Company, Inc.

Breusch T.S. and A.R. Pagan. "A Simple Test for Heteroscedasticity and Random Coefficient Variation." Econometrica. (September 1979): 1,287 - 1,294

Buckles, Stephen and Vera Freeman. "A Longitudinal Analysis of a Developmental Economics Education Program. Journal of Economic Education. (Winter 1984): 5-10.

Cameron, A. Colin and Pravin K. Trivedi. "Regression Based Tests for Overdispersion in the Poisson Model." Journal of Econometrics. (December 1990): 347-364.

Card, David and Alan B. Krueger. "Does School Quality Matter? Returns to Educaiton and the Characteristics of Public Schools in the United States. Journal of Political Economy. (February 1992): 1-40.

Clauretie, Terrence M. and E. William Johnson. "Factors Affecting Student Performance in Principles of Economics." Journal of Economic Education (Spring 1975): 132-134.

Coleman, James, Ernest Campbell, Carol Hobson, James McPartland, Alexander Mood, Frederic Weinfeld and Robert York. 1966. Equality of Educational Opportunity. Washington, D.C.: U.S. Government Printing Office.

Crowley, Ronald W. and David A. Wilton. "A Preliminary Report on the Efficiency of Introductory Economics Courses. Journal of Economic Education. (Spring 1974): 103-108.

Davisson, William I. and Frank J. Bonello. 1976. Computer-assisted Instruction in EconomicEducation: A Case Study. South Bend, IN: University of Notre Dame Press.

DBase III Plus. 1986. Version 1 - IBM/MS DOS. Culver City, CA: Ashton-Tate.

DeCanio, Stephen J. "Student Evaluations of Teaching - A Multinomial Logit Approach." Journal of Economic Education. (Summer 1986): 165-176.

Dossugi, Samuel. 1992. Determinants of Student Course Withdrawals. Doctor of Arts dissertation, Middle Tennessee State University.

Ewell, Peter T. 1992. "Longitudinal Student Databases: A Critical Tool for Managing Quality in Higher Education." In Quality Quest in the Academic Process, eds. John W. Harris and J. Mark Baggett. Birmingham, AL: Samford University.

Examiner's Manual: Test of Understanding in College Economics. Phillip Saunders and Joint Council on Economic Education. 1991. New York: Joint Council on Economic Education.

Fels, Rendig. "A New Test Of Understanding In College Economics." American Economic Review Papers and Proceedings. (May 1967): 660-666.

Folger, John and Carolyn Breda. "Evidence From Project STAR About Class Size and Student Achievement." Peabody Journal of Education. (Fall 1989): 17-33.

Gramlich, Edward M. and Glen A. Greenlee. "Measuring Teaching Performance. Journal of Economic Education. (Winter 1993): 3-13.

Greene, William H. 1990. Econometric Analysis. New York: Macmillan Publishing Company.

______________. 1992. LIMDEP: User's Manual and Reference Guide. Bellport, NY: Econometric Software, Inc.

Haley, Mary Lewis Purnell. 1991. Effects of Computer Assisted Instruction in Macroeconomics Education: An Experimental Course Design. Doctor of Arts dissertation., Middle Tennessee State University.

Hansen, W. Lee. "What Knowledge is Most Worth Knowing For Economics Majors?" American Economic Review. (May 1986): 149-152.

____________ and Allen C. Kelley. "Political Economy of Course Evaluations." Journal of Economic Education. (Fall 1973): 10-21.

Hanushek, Eric A. "Conceptual And Empirical Issues In The Estimation of Educational Production Functions." Journal of Human Resources. (Summer 1979): 351-388.

________. "The Economics of Schooling: Production and Efficiency in Public Schools." Journal of Economic Literature. (September 1986): 1,141-1,177.

________. "Testing Economic Knowledge. Journal of Economic Education. (Summer 1991): 273-275.

________. "The Impact of Differential Expenditures on School Performance." Educational Researcher. (May 1989): 45-51.

Heckman, James J. "The Common Structure of Statistical Models of Truncation, Sample Selection, Limited Dependent Variables and A Simple Estimator for Such Models." Annals of Economic and Social Measurement. (Fall 1976): 475-492.

Hilgard, Ernest R. and Gordon H. Bower. 1975. Theories of Learning. Englewood Cliffs, NJ: Prentice-Hall.

Kelley, Allen C. "The Student as a Utility Maximizer." Journal of Economic Education. (Spring 1975): 82-92.

Lewis, Darrell R. and Tor Dahl. "The Test of Understanding in College Economics and its construct Validity." Journal of Economic Education. (Spring 1971): 155-166.

________________. 1972. "Critical Thinking Skills in the Principles Course: An Experiment. In Research Papers in Economic Education, ed. Arthur L. Welsh. New York: Joint Council on Economic Education.

Levin, Harry M. 1967. Differences in Outcomes Between Large and Small Classes in Western Civilization and Economics. Doctor of Education dissertation. Rutgers University.

Lima, Anthony K. "An Economic Model of Teaching Effectiveness." American Economic Review. (December 1981): 1,056-1,059.

Lopus, Jane S. "Do Additional Expenditures Increase Achievement in the High School Economics Class?" Journal of Economic Education. (Summer 1990): 277-286.

Manahan, Jerry. "An Educational Production Function for Principles of Economics." Journal of Economic Education. (Spring 1983): 11- 16.

Martin, David A. Implications of Learning Theory for Economic Education. Kansas City, MO: Joint Council on Economic Education Annual Meeting. October 1982.

McConnell, Campbell R. and Kim Sosin. "Some Determinants of Student Attitudes Toward Large Classes. Journal of Economic Education. (Summer 1984): 181-190.

McKeachie, Wilbert J. 1986. Teaching Tips: A Guidebook for the Beginning College Teacher. Eight Edition. Lexington, MA: D.C. Heath and Company.

McKenzie Richard B. and Robert J. Staaf. 1974. An Economic Theory of Learning: Student Sovereignty and Academic Freedom. Blacksburg, VA: University Publications.

Melfi, C. A. and Donald M. Waldman. 1987. "Limited and Discrete Dependent Variable Models: Use, Applications, and Comparisons." In Econometric Modeling in Economic Education Research, eds. William E. Becker and William B. Walstad. Boston: Kluwer-Nijhoff Publishing.

Milkman, Martin and Sarah Tinkler. Writing Intensive 'Principles' Classes and Economic Learning. New Orleans: 60th Annual Conference of the Southern Economic Association. November 1990.

Mitchell, Doulgas E., Sara Ann Beach and Gary Badarak. " Modeling the Relationship Between Achievement and Class Size: A Re-Analysis of the Tennessee Project Star Data." Peabody Journal of Education. (Fall 1989): 34-74.

Mirus, Rolf. "Some Implications of Student Evaluation of Teachers." Journal of Economic Education. (Fall 1973): 35-37.

Mulligan James G. "A Classroom Production Function." Economic Inquiry. (April 1984): 218-226.

Myatt, Anthony and Charles Waddell. "An Approach to Testing the Effectiveness of the Teaching and Learning of Economics in High School. Journal of Economic Education. (Summer 1990): 355-363.

Needham, Douglas. "The Economics of Reducing Faculty Teaching Loads: Comment." Journal of Political Economy. (February 1975): 219-223.

Pindyck, Robert S. and Daniel L. Rubinfeld. 1991. Econometric Models and Economic Forecasts. New York: McGraw-Hill, Inc.

Preece, Peter F.W. "Class Size and Learning: A Theoretical Model." Journal of Educational Research. (July-August): 377-379.

Prince, Raymond, Paul H. Kipps and Howard M. Wilhelm, with James N. Wetzel. "Scholastic Effort: An Empirical Test of Student Choice Models. Journal of Economic Education. (Summer 1981): 15-25.

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

Salemi, Michael K. and George E. Tauchen. 1987. "Simultaneous Nonlinear Learning Models." In Econometric Modeling in Economic Education Research, eds. William E. Becker and William B. Walstad. Boston: Kluwer-Nijhoff Publishing.

Siegfried, John J. and William B. Walstad. 1990. "Research On Teaching College Economics." In The Principles Of Economics Course: A Handbook For Instructors. Phillip Saunders and William B. Walstad, Ed. New York: McGraw-Hill Publishing Company.

__________ and Rendigs Fels. "Research on Teaching College Economics: A Survey." Journal of Economic Literature. (September 1979): 923-969.

__________, Robin L. Bartlett, W. Lee Hansen, Allen C. Kelley, Donald N. McCloskey and Thomas H. Tietenberg. "The Status and Prospects of the Economics Major. Journal of Economic Education. (Summer 1991): 197-224.

__________, William B. Walstad. 1990. "Research on Teaching college Economics. In The Principles Of Economics Course: A Handbook For Instructors. Phillip Saunders and William B. Walstad Ed. New York: McGraw-Hill Publishing Company.

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

TUCE III Data File Codes. 7-16-91. New York: National Council on Economic Education.

Van Metre, Dale. "A Learning Theory for Economics Instructional Development." Journal of Economic Education. (Spring 1976): 95-103.

Walstad, William B. 1987. "Statistical Methods in Economic Education Research." In Econometric Modeling in Economic Education Research, eds. William E. Becker and William B. Walstad. Boston: Kluwer-Nijhoff Publishing.

__________. 1987. "Measurement Instruments." In Econometric Modeling in Economic Education Research, eds. William E. Becker and William B. Walstad. Boston: Kluwer-Nijhoff Publishing.

__________. 1987. "Applying Two-Stage Least Squares." Econometric Modeling in Economic Education Research, eds. William E. Becker and William B. Walstad. Boston: Kluwer-Nijhoff Publishing.

_________ and John C. Soper. "A Model of Economics Learning in the High Schools." Journal of Economic Education. (Winter 1982): 40-54.

Wetzstein Michael E. and Josef M Broder. "The Economics of Effective Teaching." Journal of Economic Education. (Winter 1985): 52-59.

White, H. "A Heteroskedasticity-Consistent Covariance matrix Estimator and a Direct Test for Heteroskedasticity." Econometrica. (May 1980): 817-838.

Whitney, Simon N. "Measuring the Success of the Elementary Course." American Economic Review. (March 1960): 159-169.

________. "Tests of the Success of The Principles Course." American Economic Review Papers and Proceedings. (May 1965): 566-571.

Williams, David D. and Paul F. Cook, Bill Quinn and Randal P. Jensen. "University Class Size: Is Smaller Better?" Research in Higher Education. (23:3 1985): 307-318.

Yates, Judith. "Research in Economic Education: Are Our Horizons Too Narrow?" Journal of Economic Education. (Fall 1978): 12-17.