Homework: To Require or Not? Online Graded Homework and Student Achievement

Tisha L. N. Emerson and Kimberly D. Mencken[1]

Abstract: *The recent development in economics of online automated graded homework products (e.g. Aplia) has significantly reduced the cost to instructors of assigning and grading homework. Using a quasi-experimental approach, we investigate the effect that online graded homework has on student achievement (as measured by the Test of Understanding in College Economics, final course grades, and performance on the departmental final exam). Students are "randomly" placed in one of four microeconomics principles classes taught by the principal investigators in the study. All students are required to purchase the Aplia homework product and were assigned weekly homework, but final course grades for students in only two of the study sections include a homework component. The remaining two sections serve as the control group. We find that those students who are assigned required homework through an automated homework system experienced significantly higher achievement as measured by performance on a common portion of the final exam and their course letter grade than those in the control group. There was no significant effect of the treatment when achievement is measured by performance on the TUCE. (JEL A22)*

Introduction

According to Becker (1997), between a quarter and one half of introductory economics instructors assign problem sets. To the extent that we can infer by instructors' behavior, their use of problem sets indicates a belief that such assignments are useful in promoting student learning. Problem sets provide students with additional exposure to the concepts discussed in class and practice in the mathematical analysis required in economics. This reinforcement outside of class may improve student

[1] Emerson and Mencken: Associate Professor and Senior Lecturer, respectively, in the Department of Economics, Hankamer School of Business, Baylor University, One Bear Place, #98003, Waco, TX 76798-8003. Emerson is the corresponding author. Phone: 254.710.4180 Fax: 254.710.6142 Email: Tisha_Nakao@baylor.edu

achievement as indicated by performance on exams and quizzes. Evidence from other disciplines (Aksoy and Link, 2000; Cooper, 1989; Lewis and Seidman, 1994) regarding the effect of homework on student achievement would support such a belief. Increased student learning as a result of homework, however, comes at a cost – a cost that has traditionally been borne largely by the instructor. Given class size considerations and opportunity costs, many instructors find thorough grading of regular homework assignments very costly. As a result, homework may or may not be assigned and may or may not be graded. The trade-off between benefits and costs associated with homework and perhaps more specifically the extent to which the benefits and costs accrue to different parties has, in part, given rise to automated online homework programs (e.g. Aplia, DiscoverEcon, MyEconLab). With most online graded homework programs, the grading is performed and grades recorded automatically for the instructor. Such programs significantly lower the instructor's cost of assigning homework and as a result, ceteris paribus, increase the likelihood that instructors will assign and include homework as a component in determining a student's course grade.[2]

The adoption of automated homework systems continues at an increasing rate.[3] Although the automated homework systems vary somewhat in their specific features and content, they generally allow for automatic grading of and feedback (in the form of fixed comments) on homework assignments at rather low cost to instructors. We assume that the lowering of costs increases the likelihood that instructors will require homework and include it as a portion of students' course grades. Given the significant rise in the use of such systems, we investigate whether requiring homework has any

[2] In this study we differentiate between homework that is assigned but not graded (and thus not factored into a student's course grade) and homework that is assigned and students' performance on the homework is a component in the calculation of their final course grade. It is this distinction that we refer to when we use the terms "graded homework" or "required homework".

[3] According to an Aplia representative, student enrollment in Aplia courses increased between 30 and 99 percent each academic year from the 2003-2004 to the 2006-2007 academic year. Also during this time span the increase in professors using Aplia in their courses ranged from 34 to 72 percent. Aplia offers both online, automated homework services and online experiments. These figures represent adoption of Aplia services in general as we were unable to acquire homework specific adoption rates. Nonetheless, it is clear that Aplia's services are being increasingly adopted and it is likely that a considerable portion of those adopting Aplia are using the automated homework system. We contacted the publishers of both DiscoverEcon and MyEconLab but were unable to obtain adoption information on these products.

significant impact on student achievement in a microeconomic principles course. We find that those students who were assigned graded/required homework experience higher achievement than students for whom homework is not required when achievement is measured as performance on a common portion of the final exam and their course grade. We find no differential in achievement when achievement is measured by performance on the Test of Understanding in College Economics (TUCE). Thus, required homework appears to lead to better performance on course-specific measures, but not on standardized tests.

These results are not surprising in light of models of homework's influence on student achievement. Much of the literature in this area is focused on elementary and secondary education. However, these models can be modified to the post-secondary level. Coulter (1979) presents a three phase temporal model of the factors that influence the homework process at the elementary level. In phase one, instructors motivate, structure, and facilitate the completion of homework assignments. In the second phase, the homework process is affected by student and parental specific characteristics and actions. The third, and final, phase involves teacher follow-up and feedback. Cooper (1989) modifies and expands Coulter's model of the homework process. Cooper's model provides a comprehensive list of the variables that may influence the impact of homework on educational outcomes. According to Cooper, factors that bear upon the effects of homework include student characteristics, assignment characteristics, classroom factors (prior to and post assignment), and home-community factors. Table 1 presents Cooper's temporal model adapted for the college student. The table delineates the factors that affect a variety of outcomes including academic outcomes. According to Cooper, homework is positively associated with student achievement.

Most empirical research estimating the value of homework on learning focuses on the relationship between homework and mathematics achievement with the

majority of student samples drawn from elementary and secondary school children.[4] These studies provide some illumination for our work given the strong connection between economics and mathematics. The preponderance of the mathematics studies demonstrate a positive relationship between the amount of homework and student achievement as measured by exam scores (for example see Aksoy and Link, 2000; Cooper, 1989; Goldstein, 1960; Austin, 1974; Lewis and Seidman, 1994). If increased student achievement is a benefit of homework, then increased teacher time spent preparing and grading homework is a cost. As such, there have also been studies investigating the level of grading effort necessary to obtain the student achievement benefits. In studies of junior high and high school students respectively, Austin and Austin (1974) and Small et al. (1967) find no difference in student exam scores between students whose homework was partially graded and those whose was fully graded. Paschal et al. (1984) finds that graded homework had a larger effect on elementary and secondary school student achievement than non-graded homework.

In addition to investigating the effect of the completeness of grading, researchers have also studied the effect of teacher comments. Austin (1976) shows that ninth and tenth grade students who received comments on their homework experienced greater achievement than those whose homework was marked simply for accuracy. Page (1958) compares the achievement of junior high and high school students in three groups, those receiving (1) no comment, (2) a fixed comment, or (3) open comments specific to their particular homework answers and finds that those with fixed comments experienced greater achievement than those with no comments and those with open comments experienced even greater achievement than those with fixed comments. To our knowledge, Schoen and Kreye (1974) provide the lone homework study of differential college student achievement in mathematics. They demonstrate that for their sample of college math students, achievement was not sensitive to whether the students received comments or not on their homework.

[4] Although most of the empirical studies focus on elementary and secondary age students, Cooper and Valentine (2001) and Cooper et al. (2006) report that the (positive) association between homework completion and student achievement grows progressively stronger as students grow older. It stands to reason that such progression is likely to continue to the post-secondary educational level.

Three additional studies are relevant to our investigation. First, Porter and Riley (1996) study the effect of instructional software on college level statistics students' exam performance. They show that students whose homework came from the instructional software performed better on the homework related portion of instructor constructed exams than students whose homework was assigned from the end of the chapter questions. Porter and Riley hypothesize that the immediate feedback afforded by the instructional software was key to the differential effect. Their findings, however, could also be a result of experiment design as students using the software may have worked more problems than those with the traditional homework as the software requires students to answer questions until a minimum threshold of correct answers is achieved.

Second, Wetzel et al. (1991) study the performance of high school students on the standard Test of Economic Literacy (TEL). They demonstrate that students spending more than two hours per night on homework (from any discipline) performed better on the TEL than those who did not pass the minimum nightly homework threshold. Third, Miller and Westmoreland (1998) report their findings from a study of college economic students whose homework was either graded fully or partially (with random questions selected). A solution set was distributed and all problems were worked in class when homework was returned. Miller and Westmoreland find that students' performance on homework was not significantly different between the two treatments.

From the existing literature, homework appears to be an important factor in student achievement, although little of the literature addresses college-age students or economics directly. Since results appear to vary somewhat by discipline and student age, a focus on college economics students is instructive. Our present study adds to the relatively limited literature on the impact of grading homework on student performance at the undergraduate level, in general, and provides some evidence of the benefits associated with automated homework systems in economics, in particular.

The remainder of this paper is organized in the following manner. Section II presents the data and empirical methodology. Section III reports our results. Section IV concludes and suggests avenues for future research.

Data and Empirical Methodology

Study participants were enrolled in one of four sections of microeconomic principles at Baylor University during the fall 2005 semester.[5] Individual classes were relatively small and ranged in enrollment from 32 to 40 students. All four sections required students to enroll in an automated online homework system from which they were assigned weekly homework. Two of the four sections (the treatment group consisting of 76 students) included a homework component constituting 12-18% of students' final course grade. The remaining two sections were also assigned homework (the control group consisting of 72 students), but their performance on the homework was not factored into their final course grade. The authors taught two sections each – one treatment and one control – and assigned the same homework to each of their two sections.

Our experimental design provided all students with access to the same homework assignments, their solution sets, and feedback on student performance. Aside from the difference in grading policy for the treatment and control groups, the pedagogical approach, course material coverage, and textbook were uniform across and within the two groups. In addition to the homework assignments, all study participants were also assigned three midterm exams and a comprehensive final exam. Exams consisted of a mixture of multiple choice, short answer, and numerical word problem questions. The multiple choice section of the final exam consisted of 20 questions and was common across all sections.

[5] Optimally students would be randomly assigned across sections. Such assignment, however, was not possible. At the time of student enrollment students did not know whether they had selected a section with or without graded homework assignments.

A. Model of Student Learning

To motivate our empirical work, we use an educational production function approach that is standard in the literature (see, e.g., Siegfried and Fels, 1979). In this approach, the following reduced-form model is specified:

> Student learning = f(aptitude; educational background; other student-specific characteristics; required/graded homework; observed and unobserved instructor-level effects).

Our student learning measure takes on three forms: (1) a "gap-closing" measure defined as the difference in post- and pre-course TUCE scores expressed as a percentage of the maximum possible point improvement available based on the student's pre-course TUCE score, (2) departmental final exam score, and (3) course grade. In addition to measures of aptitude (e.g., students' GPA or SAT scores), educational background (e.g., a student's major and whether a student has taken high school economics) and other student-specific characteristics (e.g., gender, age, ethnicity), we include dummy variables for instructor and for the treatment group ("homework graded" or "treatment section") that captures the differential effect, if any, on student learning associated with the treatment (required/graded homework).

B. Measures of Student Learning

Becker et al. (1991) call for multiple measures of student outcomes as indicators of efficacy of teaching approaches. We study the relationship between graded homework and three general outcome measures. These measures are the Test of Understanding in College Economics (gap-closing measure), final course grade, and performance on the 20 common multiple choice questions from the final exam. By employing multiple outcome measures, we increase the likelihood that we will be able to reliably measure the efficacy of requiring homework.

The TUCE is a standardized test of economic knowledge aimed at the principles level student. Both microeconomic and macroeconomic versions of the TUCE are available. In this study, the microeconomics version of the third edition of the TUCE was administered to all students in the sample on the first and last days of class.[6] This research design allows us to measure differences in learning, or value-added, across students. To provide incentive to exert effort on the pre-course TUCE, students were informed that their performance on the pre-course TUCE would impact their final course grade, but were not told explicitly how their score would be included in this calculation.[7] To induce effort on the post-course TUCE, students were informed that their course grades would be based, in part, on their improvement over their pre-course TUCE score.[8]

In addition to the TUCE, we also use students' performance on the 20 common multiple choice questions on the final exam and their final course grades. A departmental final exam that consisted of 20 common multiple choice questions was administered to all students taking principles of microeconomics in the fall of 2005. The multiple choice questions were written collectively by all instructors teaching microeconomic principles that semester including the two instructors participating in this study. Thus, the exam included questions contributed by the study instructors, but also questions from other non-participating instructors. The aim of the departmental portion of the exam is to measure a minimum competence level with regard to

[6] The third edition of the TUCE has a total of 33 questions for each version. Instructors are given the option of having students complete either the first 30 questions or all 33 questions. Students in this study were instructed to answer all 33 questions to the best of their ability.

[7] The pre-course TUCE was designed to be a surprise exam; i.e., students were to have had no knowledge of the exam before coming to class because such knowledge could have affected attendance and participation in the study. Because some sections began on Monday, and others on Tuesday, there is some possibility that some students taking the exam on Tuesday knew of the exam beforehand. Thus, models (described later) that explicitly account for possible selection bias include as a control whether the class met first on Monday or Tuesday. In general, this day-of-the-week effect is not present in either the selection issue or other empirical findings presented herein.

[8] During the pre-course TUCE assessment, students were not made aware of this grading method to prevent strategic behavior that could have led to a downward bias of the pre-course TUCE scores. Additionally, care was taken to ensure that the change in TUCE scores entered into grade calculation in a similar fashion across all sections.

principles level microeconomic concepts. The questions are thus written at a level where by an "average" (i.e. C-level) student could answer them correctly.

Final course grades in all sections depended upon student performance on three midterm exams and a comprehensive final exam. For students in the treatment group, their final course grades also included their homework performance (with a weight of 12-18 percent). Final course grades for students in the control group did not include homework performance. In all sections, final course grades were assigned using a standard grading scale (A for 90 and above, B for 80-89, etc.). Since homework for students in the treatment group was weighted between 12 and 18 percent – a non-trivial level – it is conceivable that homework performance may have directly affected the calculation of students' final course grades. In this sample, however, final course grades in the treatment group were the same with and without homework in the calculation. The grading scale at the study university includes A, B+, B, C+, C, D, and F.[9]

Summary statistics for each of our three outcome measures are presented in Table 2. The unconditional mean pre-course and post-course TUCE scores do not statistically differ across the two study groups. Similarly the unconditional mean scores on the 20 common multiple choice questions from the final do not differ significantly between the control and treatment groups. Finally, the average grade assigned to each group was a C.[10]

C. Measuring Other Inputs in the Educational Production Function

It is essential to control for individual student characteristics. Student achievement depends upon many factors in addition to the pedagogical approaches that are the focus of our study. In order to accurately measure the impact of required/graded homework, our analysis must control for student specific characteristics including gender, age, aptitude, effort, and socio-economic status. We

[9] Grades are recorded in the dataset as points on a 4-point scale with "+'s" adding 0.5 points to the letter's base value.

[10] While an average grade of C may seem low, it is not unusual for average grades in this course to fall in the C range. Principles of microeconomics is one of the required courses for admission to the business school and often found to be challenging by students.

collect a portion of this data through pre- and post-course student surveys including gender, ethnicity, age, employment, high school economics background, and motivations for enrolling in the course. Additional data, however, are more accurately collected directly from student records. This includes student grade point average and SAT (or ACT) scores. Becker and Powers (2001) argue against using student provided data for aptitude measures due to their unreliability. Further, Maxwell and Lopus (1994) demonstrate that student self-reporting of GPA and SAT scores may suffer from systematic reporting bias. Such nonrandom reporting would produce biased estimates of the relationship between student achievement and educational inputs. Summary statistics for student characteristics are presented in Table 2. On all measured dimensions the control and treatment groups were similar, expect for their mean SAT math score where the control (non-graded homework) group had a significantly higher mean score than did the treatment group.

D. Notes on Estimation Methods

Censoring and selection are each potential issues in our data. For example, each of our dependent variables are potentially subject to censoring problems. The TUCE scores can potentially range between $-\infty$ and 1 for the gap closing measure. In our data these values actually range between -2 and 0.83. Similarly, final exam scores can theoretically range from 0 to 20 and in our data range from 6 to 20. Course grades (on a 4 point scale) range from 0 to 4, both theoretically and in our data. In light of the potential censoring issue, we estimated all of our models using a Tobit estimation procedure. Our results were robust to a variety of specifications.

Selection bias also posed a potential issue in our data. That is, it is possible that the same variables that influence a student's performance with regard to our outcome measures also influence whether or not the student continues in the course such that we are able to construct one or more of our outcome measures for the student. If outcome measures are missing in a systematic manner, failure to control for sample selection will result in biased estimates. In such cases, Becker and Powers (2001)

recommend the use of a standard Heckman selection correction to control for any potential nonrandom attrition from the sample. In our sample of 148 students, between 3 and 15 values of the dependent variable are missing (15 TUCE related measures, 3 letter grades, and 12 final exam multiple choice scores). To control for any potential selection bias, we employ the Heckman approach and, not surprisingly, find statistically significant selection issues in our data. Importantly, however, estimates of the coefficients on the predictors of student learning, particularly our treatment-group dummy variable, are materially unchanged when controlling for selection.

We find no qualitative (or significant quantitative) differences in our estimates of the effects of participation in the treatment group when we specifically account for censoring and selection issues. Thus, in the results that we present below, we report only our OLS estimates.

Results

A. Qualitative Outcomes

Although we investigate the relationship between graded homework and three (relatively) objective measures of student achievement, these outcomes are still but a few of a variety of outcomes of interest when evaluating different pedagogical approaches (Becker, 1997, 2001). Before we present results of the TUCE, final exam performance, and course grade, we first report a set of additional outcomes for our control and treatment groups in Table 3.

Homework is assigned, in part, to provide students with practice in applying concepts presented in class. The additional practice and review is the benefit associated with homework. Students, however, face non-trivial opportunity costs and must decide how to ration their scarce time and energy. One might expect that homework that is required/graded may be more likely to be completed than non-graded homework. Students' self-reported time devoted to class preparation is a mean of 4.15 hours per week for the control group and 4.95 hours for the treatment group. The difference, however, was not statistically significant. Also not significantly different, the mean

levels of student agreement with the statements "assignments contributed to my understanding of the course" and "I learned a great deal from this course" were 3.93 and 3.81 for the treatment group and 3.67 and 3.63 for the control group (on a 5-point scale with 5 representing strong agreement to 1 signaling strong disagreement).

B. Regression Analysis of Student Learning

We estimate the impact of required/graded homework on our measures of student learning for our usable sample of 145 students.[11] Tables 4, 5, and 6 report estimates for our gap-closing (TUCE), final exam, and course grade measures respectively. As we discuss in Section II, all estimates are obtained using OLS, but these estimates are robust to corrections for sample selection and censoring.[12]

Table 4 presents estimates on the TUCE measure. We present three specifications for the gap-closing measure. The first specification is a simple difference-in-means estimation. Students in the control group gained 22.1 percent of the available points on the post-TUCE as measured by the gap-closing measure while students in the treatment group gained 18 percent of the available points. The difference, however, is not statistically significant. Subsequent specifications add controls for student-specific characteristics including aptitude, gender, age, ethnicity, educational background, and time constraints as well as an instructor control. Inclusion of these additional controls improves the fit of the model, but the coefficient on graded homework is not significant for any of the specifications. Some of the student characteristics, however, are significant predictors of the gap-closing measure. For example, student's GPA is positively correlated with TUCE performance. Interestingly and somewhat counter-intuitively, standardized absences are also positively correlated with TUCE performance. Additionally, the coefficient on age is negative and marginally significant. The remaining

[11] The usable sample differed for the various outcome measures. The total usable sample was 133 for the TUCE measures, 136 for the common questions on the final exam, and 145 for the final course letter grade.

[12] Note that the estimates reported use GPA as a measure of aptitude instead of SAT scores. While the results are not qualitatively different between the two, the sample size is somewhat smaller using SAT scores as our aptitude measure due to missing data. Thus, we opt to use GPA in order to retain as large a sample as possible.

student specific characteristics (gender, ethnicity, whether the student is retaking the course, the number of student hours completed, the number of current semester hours the student is carrying, and the number of weekly hours devoted to employment) are not significant predictors of the student's improvement on the TUCE.

Table 5 presents the estimates for student performance on the 20 common multiple choice questions on the final exam. The first specification is again a simple difference-in-means estimation where no significant difference in performance is detected between the control and treatment groups. Subsequent specifications, (3) and (4) in particular, control for a number of student-specific characteristics. Once student-specific differences are controlled for, the treatment group's performance on the 20 common multiple choice final exam questions is an average of 0.72 questions higher than the control group. For these specifications, student GPA is also a significant, positive predictor of student performance. Coefficients on the remaining controls are not statistically significant.

Results for the final outcome measure, course grade, are presented in Table 6. As with the other outcome measures, a simple difference-in-means estimation (specification (1)) reveals no significant difference between the performance of the treatment and control groups. Subsequent specifications that control for student-specific characteristics, however, do indicate a significant difference in the performance between the two groups. For specifications (3) and (4), students in the treatment group achieved higher course grades by 0.32 and 0.36 grade points, respectively. Students with a higher initial aptitude (measured both by their pre-course TUCE score and their GPA) also achieved higher course grades; as did students that were retaking the course and those that had completed more semester hours at the start of the course.

IV. Conclusion

In our study of 145 microeconomics principles students, we investigate whether students who are assigned required/graded homework through an automated homework system experienced any differential in their achievement as compared to

students who have access to the same homework but for whom homework is not required (i.e. homework completion and/or performance is not used in course grade calculation). We find that those students who are assigned required homework experience higher achievement than students for whom homework is not required where achievement is measured as performance on a common portion of the final exam and their course grade. No differential in achievement was estimated in performance on the Test of Understanding in College Economics (TUCE). Thus, required homework appears to lead to better performance on course-specific achievement measures, but not on standardized tests. Additional outcomes including student evaluations of the course and the self-reported time devoted to class preparation do not differ significantly between the two groups.

The results of this study will help develop a greater understanding of the efficacy of requiring homework of college level economics students. Homework and problem sets are a long standing pedagogical approach in undergraduate economics instruction, but relatively little more than anecdotal evidence exits regarding the benefits to students of such homework. Recent advances in technology have led to the development of a number of automated homework grading systems. These systems are used increasingly by instructors, likely due in part to the fact that they significantly reduce the cost to instructors of employing homework as a pedagogical tool. Our results suggest that there may be a benefit to students from being required to complete homework assignments and thus justify their use by instructors.

While our results suggest that students for whom homework is required and graded will achieve significantly higher achievement as measured by course specific measures, we must consider some caveats. First, the magnitude of the effects appears to be small. Students with required homework achieved an average of 0.72 questions more correct out of 20 multiple choice questions on the common portion of the final exam and one-third of a letter grade higher on their final course grade. These are relatively small improvements. Second, the homework system itself is not likely the factor driving differences in student achievement and grades. Instead, the homework

system is only the means through which students are spending additional time studying. The driving factor, we suggest, is the grading of the homework and its inclusion in final course grades. The feature of the homework system that facilitates such inclusion is the automated grading as opposed to any special aspect of the questions, their format, or form of delivery. Third, it has been suggested that the homework questions available through online automated systems tend to focus on formal model construction and analysis and much less on developing our students' abilities to "think like an economist." To the extent that this critique is true, the time spent on online graded assignments comes at a considerable cost (e.g. reading and writing assignments that may further students' ability to think like an economist). Of course, if the choice is no homework or online graded homework, then the cost would be lower. Even in this event, however, there is a cost to the student of less time spent on other courses and the monetary cost of registering for the system.

This study is a first step in expanding our understanding of the effects of required homework. Considering the widespread use of homework and problem sets, future research in the area of required/graded homework in undergraduate economics is necessary. In particular, it would be instructive to investigate the effect of required/graded homework in other courses (both introductory level and upper level courses), other university environments, and with other class sizes.

References

Aksoy, Tevfik and Charles R. Link. "A panel analysis of student mathematics achievement in the US in the 1990s: does increasing the amount of time in learning activities affect math achievement?" Economics of Education Review, 19, 2000, pp. 261-77.

Austin, Joe Dan. "Homework research in mathematics 1900-1974. paper presented at the 1974 Annual Georgia Mathematics Education conference, Rock Eagle, Georgia.

Austin, Joe Dan. "Do Comments on Mathematics Homework Affect Student Achievement?" School Science and Mathematics, 76(2), 1976, pp. 159-64.

Austin, Joe Dan and Kathleen Austin. "Homework Grading Procedures in Junior High Mathematics Classes" School Science and Mathematics, 74(4), 1974, pp. 269-72.

Becker, William E. "Teaching Economics to Undergraduates." Journal of Economic Literature, 35, September 1997, pp. 1347-73.

Becker, W., Highsmith, R., Kennedy, P., and Walstad, W. "An Agenda for Research on Economic Education in Colleges and Universities." American Economic Review Papers and Proceedings, 81(2), May 1991, pp. 26-31.

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

Cooper, H. Homework, White Plains, NY: Langman, 1989.

Cooper, Harris, Kristina Jackson, Barbara Nye, and James Linday. "A Model of Homework's Influence on the Performance Evaluations of Elementary School Students." Journal of Experimental Education, 69(2), 2001, pp. 181-99.

Cooper, Harris, Jorgianne Robinson, and Erika Patall. "Does Homework Improve Academic Achievement? A Synthesis of Research, 1987-2003." Review of Educational Research, 76(1), Spring 2006, pp. 1-62.

Cooper, Harris, and Jeffrey C. Valentine. "Using Research to Answer Practical Questions About Homework." Educational Psychologist, 36(3), 2001, pp. 143-53.

Coulter, F. "Homework: A Neglected Area of Research." British Educational Research Journal, 5, 1979, pp. 21-33.

Goldstein, A. "Does homework help? A review of research." Elementary School Journal, 1, 1960, pp. 212-24.

Lewis, Kenneth A. and Laurence S. Seidman. "Math-Time Capital Matters: A Cross-Country Analysis." Economics of Education Review, 13(3), 1994, pp. 215-26.

Maxwell, Nan L. and Lopus, Jane S. "The Lake Wobegon Effect in Student Self-Reported Data." American Economic Review Papers and Proceedings, 84(2), May 1994, pp. 201-5.

Miller, Ellen and Geraldine Westmoreland. "Student Responses to Selective Grading in College Economics Courses." Journal of Economic Education, 29(3), 1998, pp. 195-201.

Page, E. B. "Teachers comments and student performance: A seventy-four classroom experiment in school motivation." The Journal of Educational Psychology, 49, 1958, pp. 173-81.

Paschal, Rosanne A., Weinstein, Thomas, and Walberg, Herbert J. "The Effects of Homework on Learning: A Quantitative Synthesis." Journal of Education Research, 78(2), 1984, pp. 97-104.

Porter, Tod S. and Teresa M. Riley. "The Effectiveness of Computer Exercises in Introductory Statistics." Journal of Economic Education, Fall 1996, pp. 291-99.

Schoen, Harold L. and Kreye, Bettibel C. "Five forms of written feedback to homework in a mathematics course for elementary teachers." Journal for Research in Mathematics Education, 5, 1974, pp. 140-5.

Siegfried, John J. and Fels, Rendigs. "Research on Teaching College Economics: A Survey." Journal of Economic Literature, 17, September 1979, pp. 923-69.

Small, D. E., Halton, B. D., and E. J. Davis. "A Study of Two Methods of Checking Homework in a High School Geometry Class." The Mathematics Teacher, 60, 1967, pp. 149-52.

Wetzel, James N., O'Toole, Dennis M., and Edward L. Millner. "A Qualitative Response Model of Student Performance on a Standardized Test." Atlantic Economic Journal, 19(3), 1991, pp. 18-25.

TABLE 1. –Temporal Model of Factors Influencing the Effects of Homework

Exogenous Factors	Assignment characteristics	Initial Classroom Factors	Home-Community Factors	Classroom follow-up
Student characteristics	Amount	Provision of materials	Competitors for student time	Feedback
Ability	Purpose	Facilitators	Home environment	Written comments
Motivation	Skill area used	Suggested approaches	Space	Grading
Study habits	Degree of individualization	Links to curriculum	Light	Incentives
Grade level	Degree of student choice	Other rationales	Quiet	Testing of related content
	Completion deadlines		Materials	Use in class discussion
	Social context		Others' involvement	
			Classmates	
			Tutors	

*Adapted from Cooper et al. (2001).

TABLE 2. – Descriptive Statistics (Graded vs. Non-Graded Sections)

Variable	Non-Graded Mean (SD)	Graded Mean (SD)
Pre-course TUCE	10.65	10.19
	(3.54)	(3.43)
Post-course TUCE	15.41	14.49
	(5.78)	(4.19)
Final Exam Multiple Choice	15.41	15.19
	(2.85)	(2.97)
Course Grade	2.16	2.09
	(1.29)	(1.31)
GPA	3.02	2.90
	(0.59)	(0.61)
SAT-Math	609.70	586.31†
	(65.66)	(63.21)
SAT-Verbal	576.06	562.31
	(69.46)	(67.24)
Male	0.57	0.70
	(0.50)	(0.46)
Nonwhite	0.13	0.12
	(0.33)	(0.33)
Age	19.74	19.66
	(1.00)	(0.87)
Previously taken micro principles	0.02	0.03
	(0.13)	(0.17)
Work hours per week	5.15	6.73
	(8.44)	(9.79)
Semester hours completed	69.46	65.22
	(21.63)	(21.01)
Current semester hours	15.07	14.86
	(1.49)	(1.51)
Standardized number of absences	5.42	5.29
	(5.39)	(4.61)
High school course in economics	0.21	0.13
	(0.41)	(0.34)
Number of observations	76	79

† Graded and non-graded means are statistically different at the 5% (two-tailed) significance level or better.

TABLE 3. – Various Outcomes (Graded vs. Non-Graded Sections)

Outcomes	Non-Graded Mean (SD)	Graded Mean (SD)
Student evaluations: (strongly agree = 5; agree = 4; neither agree nor disagree = 3; slightly disagree = 2; strongly disagree = 1)		
Instructor appeared interested in the subject material	4.64 (0.82)	4.54 (0.75)
Instructor stimulated my interest in this subject	3.47 (0.98)	3.47 (1.16)
Instructor stimulated my thinking	3.67 (1.07)	3.80 (1.11)
Assignments contributed to my understanding of the course content	3.56 (1.09)	3.93 (1.01)
I learned a great deal from this course	3.63 (1.09)	3.81 (1.04)
Student agreed that academic background prepared him/her for the class	0.41 (0.50)	1.36 (0.48)
Hours per week that the student spent on preparation for class	4.15 (1.85)	4.95 (2.79)
Number of observations	55	58

† The graded and non-graded means are not statistically different at the 5% (two-tailed) significance level or better for any of the above.

TABLE 4. – Gap-Closing Measure (Post-course minus Pre-course TUCE/33 minus Pre-course TUCE)

	Specifications		
Independent variables	(1)	(2)	(3)
Graded section	-0.042	-0.019	-0.008
	(0.024)	(0.048)	(0.047)
GPA		0.198**	0.222*
		(0.055)	(0.070)
Male		0.022	0.024
		(0.026)	(0.032)
Age		-0.005	-0.030*
		(0.014)	(0.012)
Non-white		-0.025	-0.020
		(0.046)	(0.040)
Retaking course			-0.104
			(0.068)
High school course in economics			-0.047
			(0.059)
Semester hours completed			0.001
			(0.001)
Current semester hours			-0.008
			(0.013)
Work hours per week			-0.002
			(0.001)
Number of standardized absences			0.011*
			(0.004)
Professor A		0.023	0.014
		(0.045)	(0.064)
Intercept	0.221***	-0.311	0.128
	(0.022)	(0.198)	(0.424)
Observations	133	117	113
Adjusted *R*-squared	0.00	0.21	0.21

* significant at 10%; ** significant at 5%; *** significant at 1%

NOTES: Robust standard errors are in parentheses and are adjusted for within-section correlation of errors.

TABLE 5. – Final Exam Multiple Choice

	Specifications			
Independent variables	(1)	(2)	(3)	(4)
Graded section	-0.212	-0.142	0.724**	0.715**
	(0.630)	(0.583)	(0.170)	(0.171)
Pre-course TUCE		0.049	-0.016	-0.004
		(0.078)	(0.069)	(0.087)
GPA			3.560**	3.734**
			(0.694)	(0.908)
Male			0.754	0.509
			(0.327)	(0.322)
Age			0.098	0.404
			(0.332)	(0.239)
Non-white			0.447	0.508
			(0.371)	(0.544)
Retaking course				0.575
				(0.512)
High school course in economics				-0.317
				(0.574)
Semester hours completed				0.003
				(0.002)
Current semester hours				-0.010
				(0.039)
Work hours per week				-0.018
				(0.015)
Number of standardized absences				0.078
				(0.142)
Professor A			0.020	0.153
			(0.355)	(0.309)
Intercept	15.406***	14.862***	1.724	-4.861
	(0.189)	(0.709)	(8.713)	(7.274)
Observations	136	131	115	111
Adjusted *R*-squared	0.01	0.01	0.38	0.37

* significant at 10%; ** significant at 5%; *** significant at 1%

NOTES: Robust standard errors are in parentheses and are adjusted for within-section correlation of errors.

TABLE 6. – Course Grade

	Specifications			
Independent variables	(1)	(2)	(3)	(4)
Graded section	-0.061	-0.038	0.323***	0.361***
	(0.454)	(0.408)	(0.039)	(0.056)
Pre-course TUCE		0.094**	0.040*	0.044*
		(0.025)	(0.016)	(0.015)
GPA			1.760***	1.691***
			(0.076)	(0.085)
Male			-0.176	-0.077
			(0.127)	(0.186)
Age			0.002	-0.142
			(0.099)	(0.133)
Non-white			-0.036	-0.045
			(0.173)	(0.169)
Retaking course				0.503**
				(0.158)
High school course in economics				-0.054
				(0.211)
Semester hours completed				0.011**
				(0.002)
Current semester hours				-0.040
				(0.038)
Work hours per week				-0.007
				(0.007)
Number of standardized absences				0.001
				(0.023)
Professor A			-0.040	0.007
			(0.042)	(0.082)
Intercept	2.159***	1.180**	-3.641	-0.839
	(0.192)	(0.349)	(2.055)	(2.619)
Observations	145	140	121	117
Adjusted *R*-squared	0.01	0.05	0.72	0.72

* significant at 10%; ** significant at 5%; *** significant at 1%

NOTES: Robust standard errors are in parentheses and are adjusted for within-section correlation of errors.