The Role of Homework on Student Learning Outcomes: Evidence from a Field Experiment*

Andrew Grodner
East Carolina University
Economics Department
Greenville, NC 27858
grodnera@ecu.edu

Nicholas G. Rupp
East Carolina University
Economics Department
Greenville, NC 27858
ruppn@ecu.edu

PRELIMINARY DRAFT – COMMENTS ENCOURAGED

Download working paper at: http://www.ecu.edu/cs-educ/econ/wp2010.cfm

Abstract

This paper examines the impact of homework assignments on student learning outcomes using evidence from a field experiment of 400 students in an introductory Economics course. Students are randomly assigned to either a treatment group (homework required) or control group (homework is not required). A novel finding of this paper, which to our knowledge has not been previously been documented in the literature, is that homework required students have significantly higher retention rates. In addition, we show that requiring homework is strongly correlated with student learning outcomes in the classroom (more A's or B's and fewer F's). While requiring and doing homework are both important indicators of student performance, of these, doing homework has a larger impact on performance than requiring homework. Finally, students who are submitting high quality homework (i.e., homework average is better than the mean) have significantly higher test scores. In sum, submitting quality homework is the best indicator of exceptional test performance.

JEL Codes: H52, I21
Keywords: classroom experiment, student performance, homework assignment

*We acknowledge the comments from participants at the Economic Education Poster Session at the 2010 ASSA meetings. We thank Rob Lange and Kristy Merritt for assisting us with the data acquisition.

I. Introduction

The U.S. government's education spending topped $1 trillion dollars in 2010, comprising 5.7% of gross domestic product (www.nationmaster.com). Spending more on education, however, has not been shown to improve American students' academic achievement (Hoxby 2004; Lips, Watkins, and Fleming 2008)[1]. These educational shortcomings are not limited the United States. For example, the provision of official government textbooks in Kenya did not increase average test scores (Glewwe et al. 2009). Given this disconnect between educational spending and academic results, educational researchers have extensively studied ways to improve student achievement. Policymakers in the U.S. have attempted to find more effective ways to use educational dollars through the passing of the No Child Left Behind Act of 2001. This program tied federal funding to schools based on individual outcomes in education, has produced mixed results (Meier & Wood 2004). This paper explores an inexpensive way to potentially improve student performance through assigning homework. A novelty of this paper is the use of data collected from a carefully designed field experiment in the classroom. These data are rich enough to allow us to make inferences about how both the completion and the quality of the homework affect test scores.

Previous research has shown that assigning homework positively affects student achievement (Aksoy and Link 2000; Betts 1997; Grodner and Grove 2009). A recent study by Eren and Henderson (2010) finds that homework effects differ across subjects with math homework being associated with higher test scores for eighth graders while there is little to no effect on test scores from an additional hour of homework in science, history, or English. Another homework study by Eren and Henderson (2008) deals with restrictive functional form assumptions and heterogeneity in student ability which are two major obstacles confronting homework studies. The authors employ nonparametric estimation using National Education Longitude Study (NELS) data, which offers a rich set of controls for a large sample of students. After demonstrating the bias in parametric specification, they find that homework is an important determinant of student math test scores both statistically and economically. These effects, however, are not uniformly distributed across all students with the largest effects occurring for the high and low achievers.

Given that most of the above studies show that students benefit from homework assignments, one would expect that homework would be widely employed. To determine the prevalence of homework assignments in introductory classes at the university level, we surveyed (see Appendix Table 1A) every instructor who taught Economics Principles in Spring 2010 at the seven largest public universities in

---

[1] The U.S. 15 year olds average science score ranked 16th of 30 OECD countries and 23rd of 30 in math (*Washington Post*, 5 December 2007).

North Carolina. The survey reveals that a majority of university instructors (63 percent) currently assign homework when teaching introductory Economics courses.[2] The most commonly cited reason given by instructors who do not give homework assignments is that doing so "requires too much of my time." Since homework assignments are not universally employed, we intend to make the case that the marginal benefit from assigning homework in Economics Principles courses exceeds the marginal cost.

We use a field experiment to measure the effects of homework assignments on student test performance. The key component of our experimental design is that students in the same class are randomly assigned (via a coin-flip) into one of two groups: either homework is required or homework is not required. This random assignment allows for consistent estimates of homework effects because it eliminates the negative impact of omitted observable and unobservable student characteristics, which can cause selection bias. Our experimental design also allows us to address an existing shortcoming in the homework literature since previous studies have relied on either time reportedly spent by the student doing homework or homework hours assigned by the teacher, both of which are imperfect homework input measures. In comparison, this paper includes indicators for both homework requirement and completion, in addition to actual homework performance, information widely regarded as critical in measuring true impact of homework assignment on test performance (see footnote 8, p. 333 in Eren and Henderson, 2008; Kazantzis et al. 2000).

This paper on homework assignments reveals four primary results. First, students who are required to do homework have significantly higher retention rates. It is a novel finding which to our knowledge has not been previously been documented in the literature. Second, requiring homework is strongly correlated with student learning outcomes in the classroom (more A's or B's and fewer F's). Third, while requiring and doing homework are both important indicators of student performance, of these, doing homework has a larger impact on performance than requiring homework. Fourth, and finally, students who are submitting high quality homework (i.e., homework average is better than the mean) have significantly higher test scores. In sum, submitting quality homework is the best indicator of exceptional test performance.

The remainder of the paper is organized as follows. Next we turn to the experimental design, followed by a methods section, and then a discussion of the data. We then present our empirical results and close with some concluding comments.

---

[2] We surveyed 57 instructors regarding their homework policies. For instructors who did not respond to the survey we collected syllabi for their Principles courses to determine if homework is being assigned. Between syllabi and completed surveys we are able to determine homework policies for 95% of instructors surveyed (54 of 57). We found that 34 require homework, while 20 do not.

II. Experimental Design

Our experimental design meets the "field experiment" criteria suggested by Harrison and List (2004). For example, when we consider: (i*) the nature of the subjects*: introductory economics students are representative of freshman/sophomore undergraduates and hence good subjects to assess student learning outcomes; (ii) *the nature of the stakes*: given that homework comprises a portion of the course grade, there are substantial stakes in completing the assignments and thus students have sufficient motivation to take this homework experiment seriously; (iii) *the nature of the environment*: this investigation is carried out in a natural environment of student learning – the classroom; and (iv) *the nature of the commodity*: our commodity (the student's grade) is an actual good rather than a virtual/hypothetical good.

Our experiment is designed to randomly assign students to either receive exposure or non-exposure to the treatment (i.e., requiring homework) and then compare student learning outcomes of those who received the treatment with those who did not (i.e., the control group). Logistically, this random assignment was implemented via a coin-flip.[3] This completely randomized experimental design where the treatment is assigned to students independent of any subjects' observed or unobserved characteristics ensures that estimate of the average treatment effect is unbiased (List, Sadoff, and Wagner 2010).

Beginning in the Spring Semester 2008 through Fall 2009, students in four sections of a particular instructor's Principles of Microeconomics courses at East Carolina University (ECU) were subjects in a field experiment. During the initial class of the semester, students signed an informed consent (see Appendix) acknowledging that they agree to participate in this research study. While no student (to our knowledge) refused to sign the consent form, they could opt out of the study by dropping the course and registering for another section of principles (each semester typically 12-14 sections are offered). To ensure that the subject pool in our field experiment is representative of other economics courses at ECU, we track the number of students who dropped the field experiment section during the drop-add period and compare it with the departmental drop rate.

Table 1 provides an overview of the experimental design for the four experimental Principles sections including the number of subjects, drop rate, retention rate, instruction days and times, and classroom size. The drop rate of 17% for the Spring 2008 section is one-standard deviation larger than the Economics Department drop rate of 10% for other Principles sections. We suspect that students were dropping this class not to avoid participating in this homework study, but rather due to the unfavorable scheduling time

[3] Specifically, we announce in class that students with an even last digit of their student identification number would be required to do homework if the coin-flip outcome is "Heads," while students with an odd ID number are required to do homework if "Tails."

of the course (9am on MWF). In subsequent semesters, the classroom experiment was conducted during TuTh afternoon sections, which had a considerably lower drop rate (within one-standard deviation of the departmental average). Using a Mann-Whitley test to compare the drop rates of the four experimental sections (7.06%) with the Departmental average (7.16%) we find no significant differences between the two samples, hence we conclude that the attrition rate in the field experiment sections is representative of the typical Economics Principles course at East Carolina University.

Homework Assignments

The informed consent document explains the two scenarios in which a student's grade could be determined: either 100% (or 90%) from four equally-weighted tests and 0% (or 10%) weight on homework assignments. All students, regardless of whether homework is required or not, had access (via Blackboard) to the multiple choice homework assignments and answer keys. Students submitted homework answers on Scantron forms (which cost ten cents each or $1.40 for the course). Homework assignments and test questions were compiled from the textbook's (Mankiw) test bank.[4] There were fourteen homework assignments which covered material from thirteen chapters during this one semester course. The lowest homework score is "dropped"; hence their homework grade is the average of their remaining thirteen homework scores. There were four multiple choice tests during the semester. The fourth test was a comprehensive final exam (yet it carried the same weight as the other tests). All reported test scores are the raw scores before a curve (if necessary).

Figure 1 provides a visual depiction of the flow of subjects in our field experiment. At the conclusion of the drop-add period, there were 421 enrolled students across the four sections. Due to student attrition during the semester, 399 students took at least three of the four tests in the course. These students are represented in Figure 1. There was an even split between those who were required (n = 201) and not required (n = 198) to submit homework. Next, for each field experiment section we track homework submission rates and hence group students into one of two categories: either they submit almost every homework assignment (i.e., completing at least 13 of 14 homeworks); or they miss multiple homework assignments (i.e., submitting fewer than 13 of 14 homeworks). Not surprisingly, completing every homework assignment is much more likely (76%) for the treatment group (152 of 201) versus the control group since just 7% (13 of 198) completed every assignment.

[4] Since these sections had 100+ students enrolled, multiple choice questions were used for both homework and tests. The test questions were similar (but not identical) to the homework questions.

Figure 2 shows considerable attrition in homework submission rates during the semester. Prior to the first exam, students appear very eager to complete homework assignments since average homework submission rates exceed 90% (for treatment group) and about 14% (for control group). After the first exam, homework submission rates begin to taper off with 87% of treatment group and about 5% of control group completing homework. This tapering continues throughout the semester culminating with less than three-fourths of the treatment students submitting every homework and less than 3% of the control group turning in every assignment prior to the final exam.

Descriptive Statistics

Descriptive statistics for our experimental data appear in Table 2. This table reveals that 419 students took Test 1 with the control group (homework not required) comprising slightly more than half of the sample (213 vs. 206). With each subsequent test, we collect fewer student observations since some students either dropped the course or gave up on the class and hence stopped taking tests. Seven percent of the class did not take the final exam. By the end of the semester, there were more homework required students remaining in the course than not required (197 vs. 194).[5]

Student retention rates are plotted on Figure 3. Measuring student retention based on the proportion of enrolled students who take the final exam, we find higher retention rates for the homework required treatment group (exceeding 95%) compared to the control group (less than 90%). More formally, a Mann-Whitney test rejects the hypothesis that the treatment and control groups have equivalent retention rates. Therefore we discover a positive externality of requiring homework appears in the form of higher student retention rates. Finding ways to improve student retention is especially important and timely since University of North Carolina (UNC) system President Erskine Bowles has proposed tying all future university funding growth for the UNC system universities to freshman student retention benchmarks. Schools that fail to reach their targeted benchmark would receive no new funds (see Appendix Table 2A).

We use two measures to track homework submission rates (*HW_Submit_100%* and *Proportion HWs completed*) which appear on Table 2. *HW_Submit_100%* is a binary variable that takes the value of 1 if the student submits every homework assignment preceding the test. For example, the first test covered material that appeared on the first four homework assignments. We find that about three-fourths of the treatment group submitted every homework assignment prior to Test 1. In comparison, just 6% of the control group submitted each of the first four homework assignments. The other homework submission variable: *Proportion HWs completed* is a simple average of homework submitted to homework assigned.

[5] The high withdraw in this course maybe partly due to the fact that in this introductory class, students are able to use a grade replacement if they obtain a poor grade.

For the first test, the *Proportion HWs completed* were considerably higher at 91% for those in the treatment group. For both the treatment and control groups our homework submission measures show a declining rate of homework submission as the semester progresses.

Finally, Table 2 reveals that the random homework assignment has done an excellent job in creating nearly identical distributions of our student ability measures for the control and treatment groups. For example, there is only a 1 point difference among our two groups for SAT verbal scores and just a 7 point spread between SAT math scores (with higher math scores for the control group). Average high school GPA's are also quite similar with just 0.04 points separating the two groups. Next, we turn to the methodology used to examine our experimental data.

III. Methods

The challenges of estimating the education production function are well known (Hanushek 1979). Previous homework studies have been handicapped with poor data on student homework and hence have been forced to use imperfect inputs of education (i.e., self-reported hours of homework completed or hours of homework assigned). Also, these studies must confront several issues including measurement of output in the education production process, multiple outcomes, specification and modeling of inputs, and specification of criteria for efficiency when evaluating academic achievement. Our model attempts to address some of these challenges.

Most previous studies which model homework as an input in the academic performance function assume that time is the major constraint for students (Betts 1997). Thus, assuming that students exert the same amount of effort per time unit when preparing for tests, we can specify our baseline model as:

$$ln(test\_score)_i = \alpha + \beta(HW\_effort)_i + \varepsilon_i \qquad (1)$$

where i indicates an individual student and β measures the effect of time spent doing homework as indicated by the amount of effort. We proxy the variable *HW_effort* by using information about whether students are required to do any homework (*HW_Req)* or whether they did every homework (*HW_Submit_100%)*. Similar to the results by Rich (2006), we expect that homework effort is associated with higher test scores, hence we should find a positive β coefficient.

Because students may spend time preparing for homeworks and not studying enough for tests, there is a possibility that β could in fact be negative (Neilson 2005). Thus, the model may predict a counterintuitive result where more time spent on the homework may actually lower test scores. Neilson (2005), however, points out that such results are an artifact of using the wrong homework measure, where homework performance may actually be more important than time spent studying. Thus, in our specification (1), for students who submitted at least one homework, we also proxy *HW_effort* with homework score average (*HW_Avg)*. Higher homework scores should be strongly correlated with intensity of studying and thus we expect an increase in homework effort to raise test scores.

Inclusion of homework performance measure *HW_Avg* also addresses a potential problem of estimating (1) when the homework inputs are measured imprecisely which may result in measurement error problems. For example, Aksoy and Link (2000) rely on student responses regarding the hours spent doing homework, which has the potential for spurious correlation with test scores since hours studying likely reflects unobserved variation in student ability and motivation. Betts (1997) uses information about teacher assigned homework which may not correctly reflect the completion rates and thus create measurement error problems. Thus, our model specification follows the recommendation of Kazantzis et al. (2000) who distinguishes between homework performance (intensity) and homework compliance (completion rate).

Further examination of the education production function reveals that the relationship between time spent completing homework and test performance is most likely non-monotonic, where homework assignments create a bigger spread between best and worst students (Neilson 2005). Such non-linearity in the relationship between homework performance and test scores may create problems in finding an appropriate linear specification, which if specified incorrectly, can result in inconsistent estimates. Eren and Henderson (2008) estimate a nonparametric model and find that parametric specification overstates the impact of schooling related inputs on test scores.

Thus, in order to control for potential non-linearity, we also include in our model an interaction term between our homework effort measures:

$$ln(test\_score)_i = \alpha + \beta(HW\_effort)_i + \gamma(HW\_effort\ interactions)_i + \varepsilon_i \tag{2}$$

where γ indicates potential non-linear effect of effort on the test scores. We expect that controlling for the interaction effects between effort measures eliminates bias in estimating (1) due to potential model misspecification.

The simplicity of our regression equation is possible primarily due to the careful experimental design. In general, however, lack of important controls for student ability or school quality can create biased estimates. Thus, one needs a rich set of controls for student characteristics.[6] Due to the extensive information available in our data we can address this potential heterogeneity issue by estimating the following model:

$$ln(test\_score)_i = \alpha + \beta(HW\_effort)_i + \gamma(HW\_effort\ interactions)_i + \lambda(demographics)_i + \mu(ability)_i + \varepsilon_i \quad (3)$$

where coefficients in λ and μ indicate the effects of student individual characteristics. Our expectation is that the effects of both homework effort and homework average will remain unchanged after controlling for both demographic characteristics and student ability measures, even though these additional controls will increase the overall fit of the regression. Thus, the results are expected to confirm that because students in our sample are randomly assigned to the treatment and control groups, the variables in *demographics* and *ability* are in fact orthogonal to homework effort measures and they have an independent effect on *test_score*.

Finally, estimation of (1)-(3) can be compromised by issues of self-selection. Chan et al. (1997) shows positive bias in estimating the impact of time spent in class on test scores when not correcting for course attrition. Jirjahn et al. (2009) presents evidence that sorting of students in classes taught by teachers with high ability distorts the effect of class size and student performance. Grodner and Grove (2009) find that students select teachers on the basis of who assigns homework, and show that if the researcher does not control for self-selection, then the impact/benefit of homework is overstated. Our experimental design controls for these self-selection issues by randomly assigning students to the treatment and control groups.

IV. Results

We present average test scores for both the treatment and control groups on Table 3. The first test score is featured because this test not only has the largest number of observations, but more importantly it does not suffer from selection/survival biases. Students who are performing poorly are more likely to drop the course and hence not take tests administered later in the semester. The total column from Table 3

[6] Other possibilities include using past performance to control for the effect of all previous inputs in the educational production process (Hanushek, 1979 & 2005), or using a panel data model where the unobserved student characteristics can be controlled for by fixed effects (Todd and Wolpin 2003, Rivkin et al. 2005). Note that in introductory classes, where the information about the student history is limited, both of these strategies are not feasible.

shows that homework required treatment group has a two point higher test 1 average than the homework not required control group. We further divide the groups based on homework submission rates. While requiring homework appears to slightly increase test 1 scores (by about two points), homework completion rates have a much larger effect (between four and six points) on the initial test performance.

Regardless of whether homework is required or not, students who completed every homework assignment before the first test performed substantially better than students who skipped one or more homework assignments. For example, the control group on Table 3 shows a six point higher test 1 average for students who completed every homework assignment. For the treatment group, there is a four point higher test 1 performance for those who submitted every homework assignment.

The bottom half of Table 3 presents the test average of four equally weighted tests during the semester.[7] Once again the treatment group outperforms the control group by posting a two point higher test average (see the Total column). We attribute the superior performance of the homework required students to this group's higher homework submission rates. Students which submitted 100% of the homework assignments during the semester (or missed a single homework assignment) had a five point higher test average (in the control group) or a seven point higher test average (in the treatment group). The better performance observed from submitting homework is even more noteworthy considering that we place no conditions on the quality of the homework submitted.

Figure 4 provides a visual depiction of how homework requirement and completion rates affect test performance throughout the semester. We use the term "committed" to describe students in the treatment group who submit every assignment. Figure 4 shows that the committed group consistently outperforms the other groups. We note that the "super-committed" group (students in the control group who complete every homework) registers the best test performance, however, its small sample size (n=13) prevents meaningful inferences and hence it is omitted from Figure 4. The group that consistently underperforms the others we term "uncommitted" since these students in the treatment group miss multiple assignments. Finally, students in the control group who miss multiple homework assignments are considered the "baseline" group. The performance of the baseline group is sandwiched between the committed and uncommitted groups for each of the four tests during the semester.

While the previous three figures have discussed the association between homework and test performance, next we turn to what matters most to the student: their course grade. The treatment group is more likely to earn good grades (A's or B's) and less likely to receive failing grades. More formally, a Mann-Whitley test ($z=1.98$ & $p\text{-value} < 0.05$) indicates that significantly more students in the treatment

[7] Should a student take just three of four tests, they are included in this test average calculation. Students who miss multiple tests are excluded from the test average.

group (54%) earn A’s and B’s compared to the control group (45%). Moreover, the rank-sum test (z = 3.20 & p-value < 0.01) also rejects the hypothesis that the failure rate 5% vs. 15% is equivalent between the treatment and control groups, respectively.

Finally, we present the grade distributions for the committed, uncommitted, and baseline groups (see Figure 5). Recall that the committed and uncommitted student groups share the same grading criteria: 10% weight on homework and 90% weight from four exams. A’s are four times more likely for committed (23%) vs. uncommitted (6%) students while B’s are twice as likely (46%) for the committed vs. uncommitted (24%). At the other side of the grade distribution, since Economics majors are required to earn a “C” or better in this introductory principles course, we are also interested in the proportion of Economics majors that must re-take the class due to getting a “D” or “F”. Just one in ten “committed” students must re-take the course, whereas an eye-popping four in ten of the uncommitted students are forced to retake this course should they major in Economics.

The baseline group has different grading criteria than the committed and uncommitted groups since 100% of their grade is test determined. The result is a grade distribution for the baseline group that resembles a uniform distribution. This baseline group has a large number of A’s (18%) which more closely resembles the committed group of students and a nearly equal proportion of B’s and C’s (about 25% each). Finally, three out of ten students from the baseline groups fair poorly in the course earning D’s (15%) and F’s (16%).

Regression Results – Test 1

We begin our regression analysis by examining test 1 scores. As mentioned previously, we believe that test 1 scores are perhaps the cleanest measure of the link between homework and student performance since no student attrition has occurred. The OLS regressions on Table 4 and 5 do not include controls for demographic characteristics or student ability measures since these attributes have been randomly distributed across both treatment and control groups. When such controls are included in the model (see Table 6), we find that only modest changes occur (e.g., homework measures having a slightly reduced coefficient estimate).

Turning our attention to the results from Table 4 the first model shows that students who submitted every homework assignment prior to the first test had a 5% higher test 1 score than students who missed at least one homework assignment. These findings are consistent with Table 3 which reported a 4% to 6% higher test 1 score (depending on whether homework is required or not required) for students who submitted every homework. The *HW_Required* coefficient in the second model reveals that treatment group outperforms the control group by 3.5% on the initial test. Comparing the magnitude of the coefficients for the first two models on Table 4 reveals that doing every homework (model 1) has a larger impact on student test performance than simply requiring homework (model 2).

The third model includes *HW_Average* which measures the quality of the completed homework assignment. *HW_Average* is conditional on actually submitting a homework assignment (i.e., we exclude missing homework assignments). We don't observe homework averages for more than one-third of the sample, since these students did not submit any homework assignments (and hence are excluded from the regression). We find that students with higher quality homework assignments do better on the first test. We interpret the 0.0072 HW_Average coefficient as follows: among the students that submitted at least one homework, a ten point increase in homework average (on the four homeworks prior to Test 1) will increase Test 1 scores by seven percent. In the fourth model, we include two homework measures (required and submitted) and an interaction term: *HW_Submit*HW_Required*, results in insignificant estimates for all included homework measures.

The fifth model includes *HW_Submit_100%, HW_Average*, and the interaction term of these two variables. We find that only the *HW_Average* term has a significant effect on test 1 scores. The magnitude is slightly smaller than previously reported as a ten point increase in homework average now contributes to a six percent increase for the first test. In sum, we find three different measures of homework effort (*HW_Submit_100%, HW_Required,* and *HW_Average*) all contribute to better performance on the first test. Of these three, only *HW_Average* retains its statistical significance when interaction terms are included in the estimation. This result underscores the importance of including a measure of homework quality like homework average rather simply counting the number of assignments completed.

Regression Results – Test Average

Table 5 shows results when test average is the dependent variable. The homework effort results are quite robust across the estimated models since the regressions indicate that *HW_Submit_100%* (model 6), *HW_Required* (model 7), and *HW_Average* (model 8) all have significantly higher test averages. The biggest difference between test average and test 1 is the magnitude of the coefficients: *HW_Submit_100%* is nearly three times as large for test average (comparing models 1 vs. 6), and twice as large for *HW_Required* (models 2 vs. 7) and about 30 percent larger for *HW_Average* (models 3 vs. 8). The *HW_Submit_100%* criteria is much more rigorous for test average since this requires that students submit at least 13 of 14 homework assignments; whereas, a student can submit the first four homeworks and receive the *HW_Submit_100%* designation for test 1.

Table 5 shows that students are clearly rewarded for doing homework on a consistent basis throughout the semester. We find a nearly fifteen percent higher test average (or at least a one letter grade higher course grade) for students who submit every homework assignment. While not as large of an effect as doing homework, we find that requiring homework has nearly an eight percent increase on test average. Given that test average comprises 90% of a student's grade, an eight percent higher test average may very well translate into a one letter grade higher course grade. Among the students that submitted at least one homework during the semester, students with higher *HW_Average* also have higher test averages. Specifically, models (8) and (10) suggest that a ten point increase in *HW_Average* increases test average by about nine percent.

Table 6 provides estimates for homework effort measures when we include controls for both demographic characteristics (total credit hours, semester credit hours, race, and gender) and ability measures (SAT verbal, SAT math, and high school GPA). We find only modest changes to the results with the inclusion of these additional controls in the regressions. This result that only a few changes occur is a testament to the fact that students were randomly assigned to the treatment and control groups. The inclusion of the demographic and ability measures has slightly reduced the magnitude (but not the statistical significance) of the various homework measures. For example, conditional on submitting at least one homework assignment during the semester, a ten point increase in *HW_Average* is associated with a 6.1 percent test average increase when the additional controls are included (model 13) compared to a 9.5 percent increase in test average when these controls are excluded (model 8). Nonetheless, the story remains consistent: submitting high quality homework improves student performance.

Finally, with the inclusion of demographic characteristics and ability measures we find *HW_Submit_100%* is associated with significantly better (fourteen percent) test average (see model 14). In this model we include interaction terms for requiring and submitting homework, hence the net effect for "committed" students (who are both required and submitting every homework) is a 12.5 percent higher test average. While at first glance, the estimated coefficients for the homework variables from model (14) appear quite different from model (9), they are actually, however, very similar since the net affect of the "committed" students reveals a 14.4 percent higher test average.

V. Conclusion

This study examines the effect of homework assignments on student performance. We conduct a field experiment on 419 students in four sections of an introductory Economics course. By randomly assigning which students are required to submit homework within the same class, we are able to avoid self-selection issues that plague previous homework studies. Moreover we observe multiple measures of homework inputs (i.e., homework submission rates and homework average) which are superior to homework input measures that rely on either student reported hours spent studying or teacher assigned homework hours.

We find considerable evidence which suggests that homework is beneficial to student performance. Specifically, our treatment group of students (who are required to do homework) have significantly better test performance than the control group of students (for whom homework is not required). Both groups attended the same lecture and had access to the homework sets and answer keys. While requiring homework helped student performance by raising test average by an estimated eight percent, a much larger effect on test scores was due to students who actually completed the homework assignment. In fact, students who completed every homework (or all but one assignment) during the semester had an estimated fifteen percent higher average test scores.

We also find that homework quality matters. In our study, high homework averages are the best indicator/predictor of higher test scores. We estimate that a ten point increase in homework average is linked to a nine percent higher test average. Finally, we find that students in the treatment group are significantly more likely to receive good grades (A's and B's) and are significantly less likely to receive failing grades compared to the control group. Within the treatment group, the distinction between good grades and failing grades is magnified if we look at homework submission rates for the "committed" students (those who are required and doing homework) perform well, while the "uncommitted" students (those who are required and chose not to do homework) usually perform poorly.

What is it about assigning homework that is driving this large improvement in student performance? One possible explanation is that homework assignments create a more even distribution of student effort in the class throughout the semester. In fact, previous research has shown that an increase in studying hours or "cramming" at the end of the course may actually hurt performance (Dean, Stevens, and Stevens 2006), especially so for weaker students (Schmidt 1983). Hence students who are doing homework avoid this cramming pitfall.

A second potential explanation for the better performance of students who are submitting homework is that these students receive more frequent signals of how they are doing during the semester and hence they know whether they are mastering the course material. Students who do poorly on a particular homework assignment are now aware of which material they do not understand prior to taking the test. Or put more simply, students appear to learn from their homework mistakes.

We also discover a positive externality from assigning homework, to our knowledge that has not previously been documented in the literature, is the significantly higher student retention rates. The importance of student retention cannot be overstated in the North Carolina University System since the current system President Bowles is proposing that future increases in university funding be tied to student retention targets.

This study has shown that homework provides an effective way of increasing student performance. We hope that Economics instructors who are not currently employing homework assignments in their introductory courses might reconsider their homework policies given that this field experiment shows considerable evidence that homework improves both student learning outcomes and student retention. More generally, due to ever tightening government budgets for public education, requiring homework may offer a low cost way to improve the quality of instruction without requiring additional education expenditures.

References

Aksoy, Tevfik and C.R. Link. 2000. “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: 261-77.

Betts, Julian R. 1997. “The Role of Homework in Improving School Quality,” Unpublished Manuscript, University of California, San Diego.

Chan, Kam C; Connie Shum, and David Wright. 1997. "Class Attendance and Student Performance in Principles of Finance," *Financial Practice and Education*, 7(2): 58-65.

Dean, D., E.E. Stevens, and J.L. Stevens. 2006. “The Marginal Productivity of Short Term Versus Long Term Learning Effort in Course Performance,” working paper, University of Richmond.

Eren, Ozkan, and Daniel J. Henderson. 2008. “The Impact of Homework on Student Achievement,” *Econometrics Journal*, 11(2): 326-48.

Eren, Ozkan, and Daniel J. Henderson. 2010. "Are We Wasting Our Children's Time by Giving them More Homework?" working paper, University of Nevada, Las Vegas.

Glewwe, Paul, Michael Kremer, and Sylvie Moulin. 2009. “Many Children Left Behind? Textbooks and Test Scores in Kenya,” *American Economic Journal: Applied Economics*, 1(1): 112-35.

Grodner, Andrew and Wayne Grove. 2009. “Estimating Treatment Effects with Multiple Proxies of Academic Aptitude," working paper, East Carolina University.

Hanushek, Eric A. 1979. "Conceptual and empirical issues in the estimation of educational production functions," *Journal of Human Resources,* 14: 351-88.

Hanushek, Eric A. 2005. "Teachers, schools and academic achievement," *Econometrica,* 73: 417-58.

Harrison, Glenn W. and John A. List. 2004. “Field Experiments,” *Journal of Economic Literature*, 42(4): 1009-55.

Hoxby, Caroline M. 2004. "Productivity in Education: The Quintessential Upstream Industry", *Southern Economic Journal*, 71(2): 209-31.

Jirjahn, Uwe; Christian Pfeifer, and Georgi Tsertsvadze. 2009. "Class Size, Student Performance and Tiebout Bias," *Applied Economics Letters*, 16(10-12): 1049-52.

Kazantzis, Nikolaos, Frank P. Deane, and Kevin R. Ronan. 2000. *Clinical Psychology: Science and Practice*, 7(2): 189–202.

Lips, Dan, Shanea J. Watkins, and John Fleming. 2008. "Does Spending More on Education Improve Academic Achievement?" *Backgrounder* by The Heritage Foundation, 2179, September 8.

List, John A., Sally Sadoff, and Mathis Wagner. 2010. "So You Want to Run a Field Experiment: Now What? Some Simple Rules of Thumb for Optimal Experimental Design" NBER working paper #15701.

Meier, Deborah and George Wood, editors. 2004. *How the No Child Left Behind Act is Damaging our Children and our Schools,* Beacon Press, Boston, MA.

Neilson, William. 2005. "Homework and Performance for Time-Constrained Students," *Economics Bulletin*, 9(1): 1-6.

Rich, Steven P. 2006. "Student Performance: Does Effort Matter?", *Journal of Applied Finance*, 16(2): 120-33.

Rivkin, Stephen G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, schools and academic achievement," *Econometrica*, 73(2): 417-58.

Schmidt, Robert M. 1983. "Who Maximizes What? A Study in Student Time Allocation", *American Economic Review*, 73(2): 23-28.

Todd P.E. and Wolpin K.I. 2003. "On the specification and estimation of the production function for cognitive achievement," *Economic Journal*, 113: F3-F33.

Appendix Table 1A: Homework Survey (conducted via e-mail)

My name is Nicholas Rupp and together with my colleague Andrew Grodner, we are examining the link between homework assignment and classroom performance. We are conducting a survey of the six largest public Universities in North Carolina to determine the prevalence of homework in introductory Economics classes. You could help us tremendously by answering the following questions:

1. This semester in your Principles classes at NC State do you assign homework?

2. If you answered "yes" to question #1, then who administers these homework assignments? (you can select more than one answer)

   A. Aplia
   B. Textbook's web site (e.g., myeconlab.com)
   C. You do it yourself
   D. Internet homework site other than Aplia
   E. Other (please specify)

3. If you answered "no" to question #1, then why aren't you assigning homework? (again, you can select multiple answers)

   A. Requires too much of my time
   B. Provides little benefit to students
   C. Students will simply copy each other's answers
   D. Homework administered sites are too expensive
   E. Other (please specify)

4. When you are teaching non-principles classes, do you usually assign homework?

   A. Yes
   B. No

All of survey responses will remain confidential. Finally, if you would like a copy of the working paper (once it's competed), please let me know and I would be happy to send it to you. Thank you for your time in completing this short survey. Good luck teaching this semester.

Best regards,

Nicholas G. Rupp

Appendix Table 2A: Minutes of the September 2009 East Carolina University Board of Trustees meeting

CHAIRMAN'S REPORT Chairman Hannah Gage announced the members of the 2010 O. Max Gardner Award Committee. They were Mr. Mills, chair; Mr. Blackburn; Dr. Buffaloe ; Dr. McCain; and Mr. Mitchell. Chairman Gage stated that the previous day's policy discussions were some of the best. Important initiatives were identified separate from the budget cuts which demonstrated that the University was moving in a new direction. **She thought that the idea of tying enrollment growth rates to graduation and retention rates was a significant change – one that will better serve the students and the State.** Once a strategy was adopted, the next step would be to find a way to reward campuses for accomplishing this task.

A task force of the ECU Chancellor's was put together to study the issues.

CONSENT TO PARTICIPATE IN PROJECT

Title: Estimating Treatment Effects on Academic Success in the Classroom

UMCIRB#: _07-0753___________________

Principle Investigator:__Nicholas G. Rupp, Ph.D Economics___________________________

You are being asked to participate in a research study. If you choose not to participate in this research study, please drop this Econ 2113 course and choose another Econ 2113 section from the sections that are available this semester.

Before you agree, the investigator (Dr. Rupp) will tell you what your participation involves. This study examines the treatment effects on academic success in the classroom. Students will be randomly assigned to a treatment. All students will be given the homework assignments. The two treatments are:

1. "homework required"
2. "homework is not required"

If you are in the "homework required" treatment, then a small portion of your grade (10%) will be determined based on your homework performance. If you are in the "homework is not required" treatment, then no part of your grade will be based on the homework assignments. You may still turn in your homework. It will be graded and returned, however, no credit will be given for doing the homework for the "homework is not required" students.

A coin-flip will determine whether you are assigned to the "homework required" or "homework is not required" group.

All personal information of participants, including you, will remain confidential and will only be used for statistical analysis. We plan to conduct this study in three Econ 2113 classes involving approximately 750 ECU undergraduate students. If you agree to participate, you must be given a signed copy of this document and a written summary of the research plan.

This Consent Document also grants the researchers the ability to obtain your ECU GPA, high school GPA, SAT score, and current hours enrolled in. All information from this study will be coded without names, and all individual homework and test performance will remain confidential.

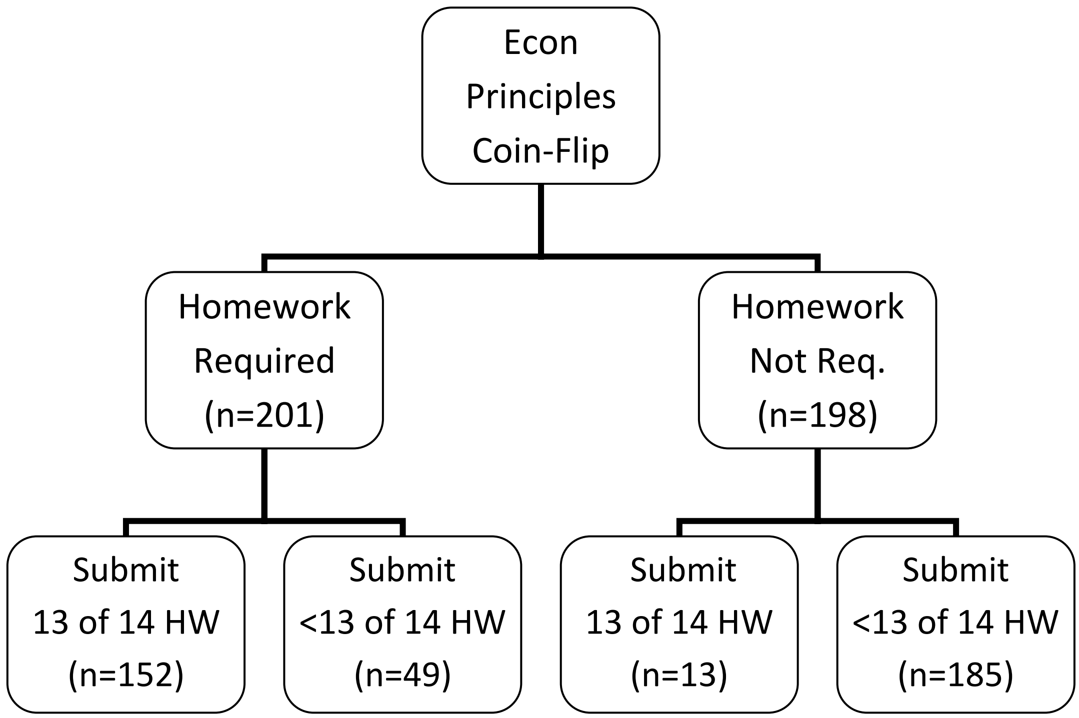
You may contact __Dr. Nicholas G. Rupp at _328-6821______ any time you have questions about the research. You may contact the University and Medical Center Institutional Review Board at (252) 744-2914 if you have questions about your rights in this study or what to do if you are injured.

Signing this document means that the research study, including the above information, has been described to you orally, and that you voluntarily agree to participate.

Name of Participant____________________________

Signature of Participant__________________________ Date_______________

Figure 1: Experimental Design



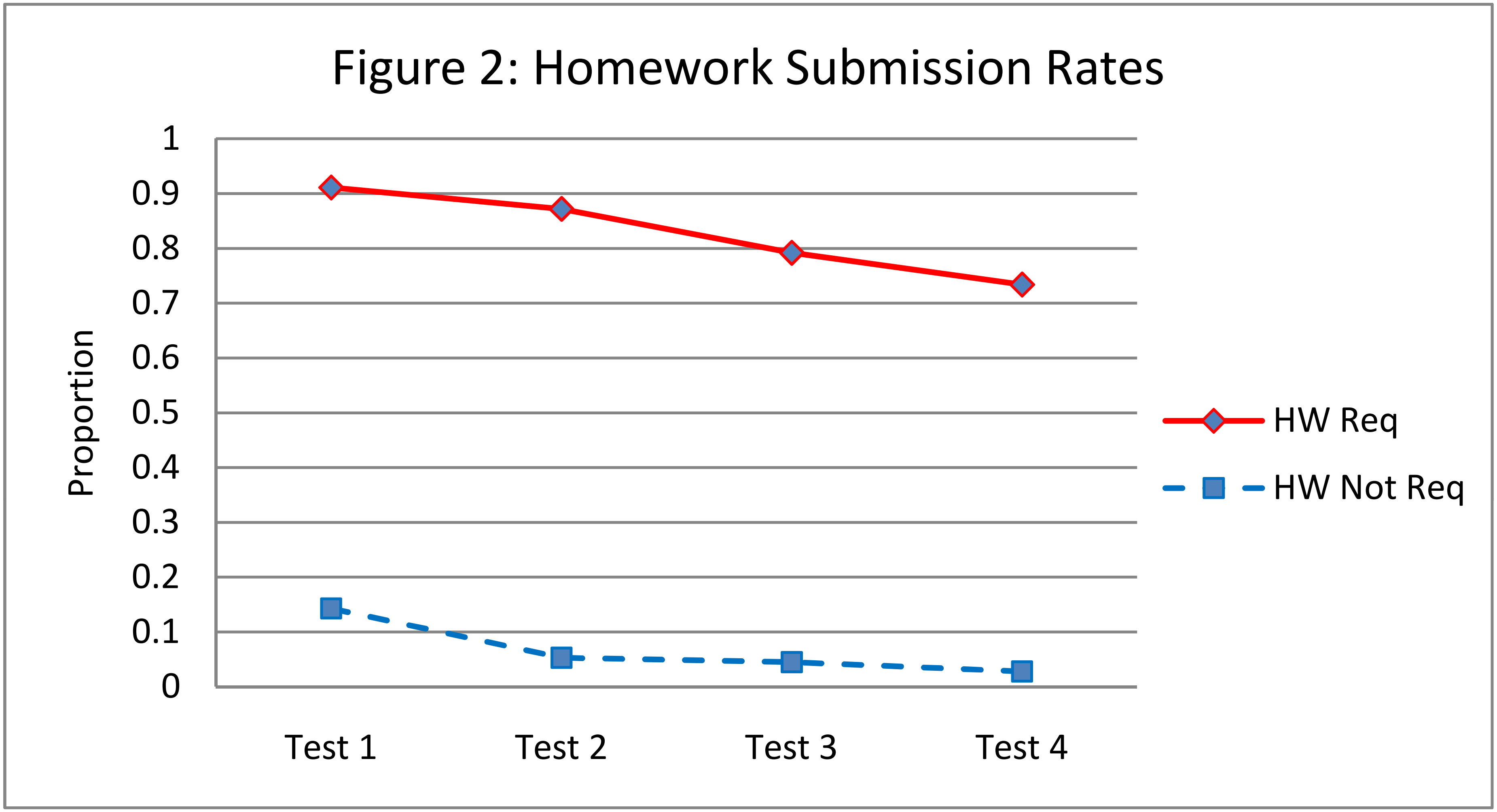


Figure 2: Homework Submission Rates
Proportion
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Test 1
Test 2
Test 3
Test 4
HW Req
HW Not Req

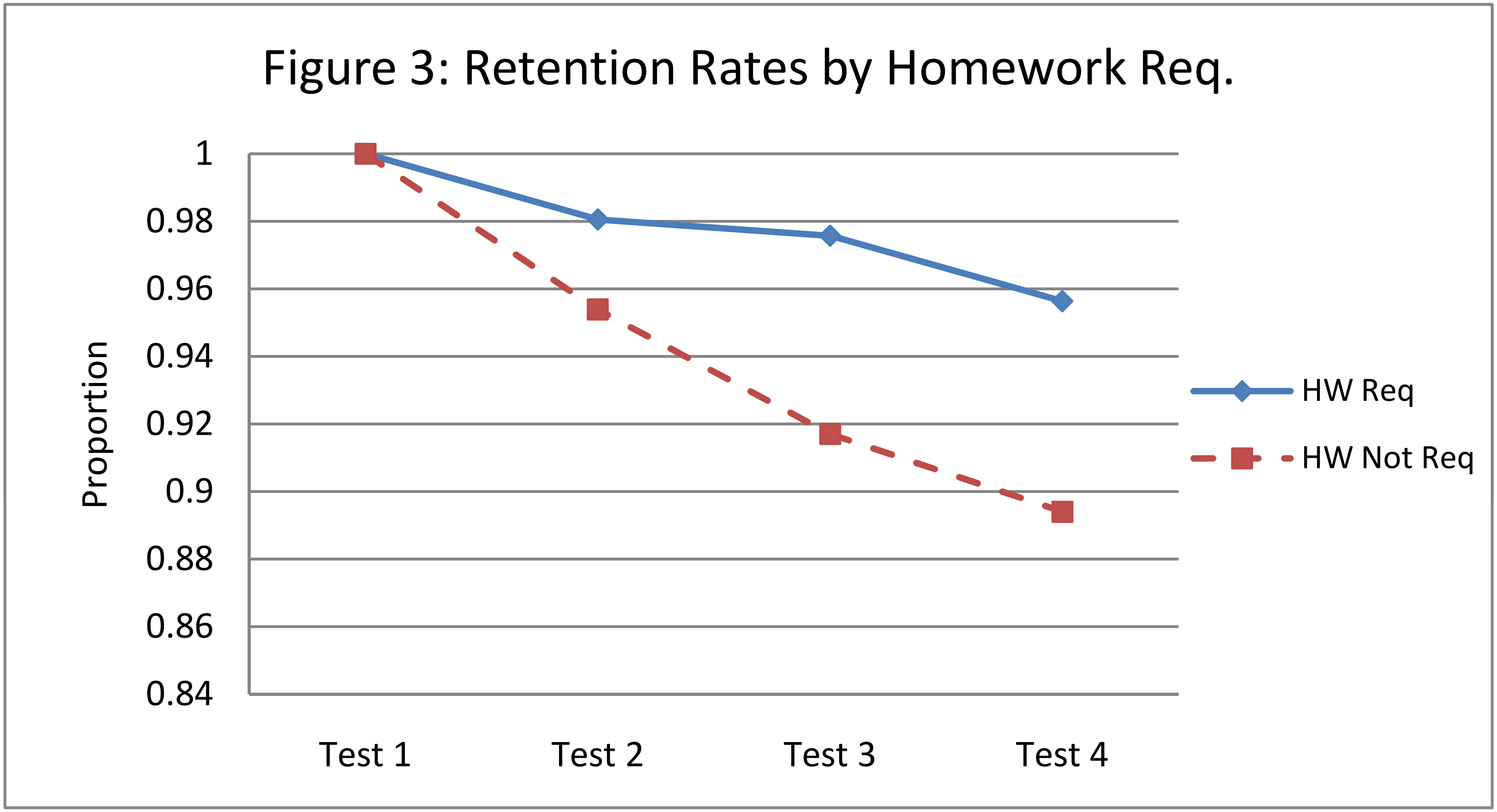


Figure 3: Retention Rates by Homework Req.
Proportion
1
0.98
0.96
0.94
0.92
0.9
0.88
0.86
0.84
Test 1
Test 2
Test 3
Test 4
HW Req
HW Not Req

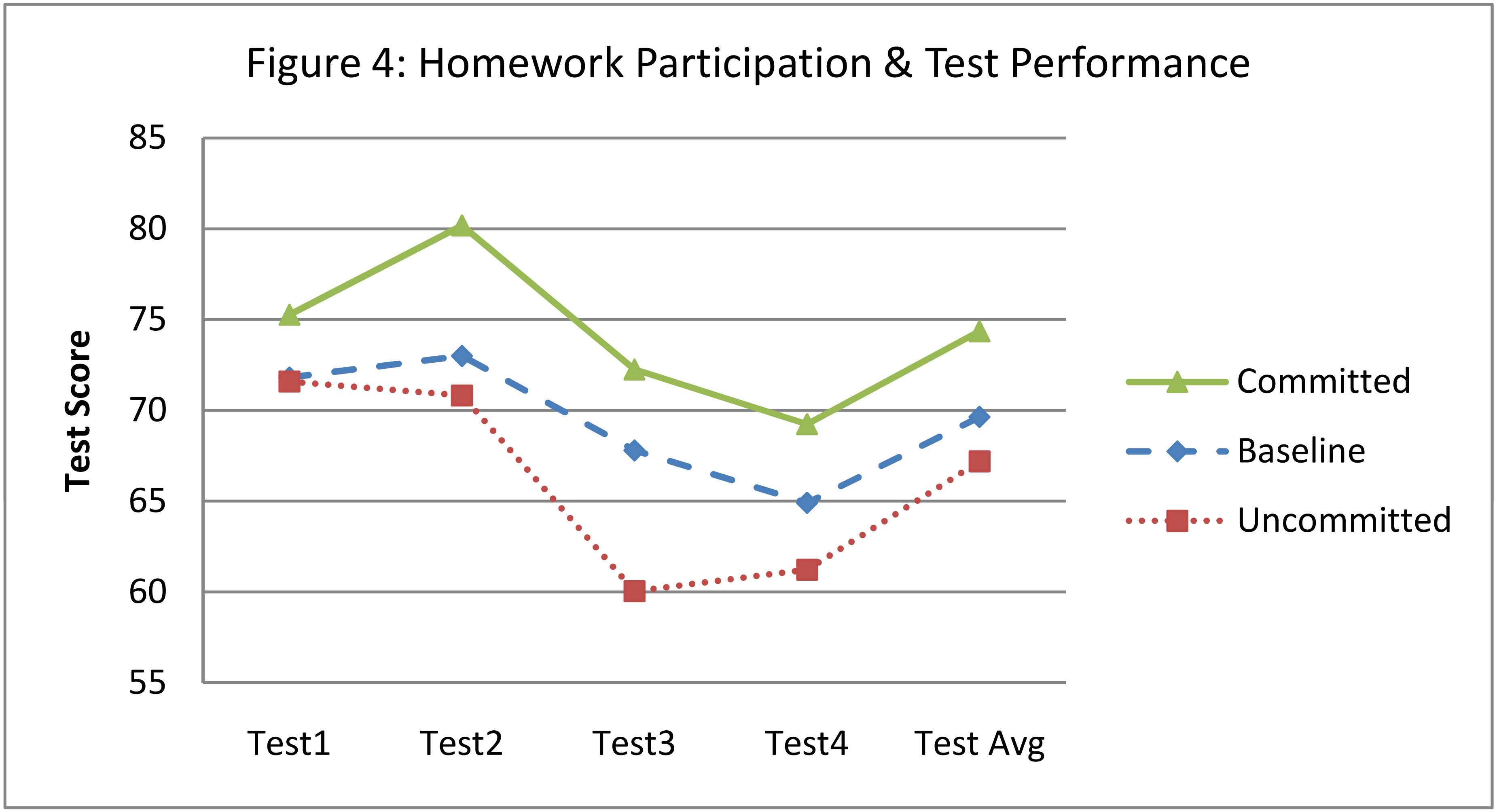


Figure 4: Homework Participation & Test Performance
85
80
75
70
65
60
55
Test Score
Test1
Test2
Test3
Test4
Test Avg
Committed
Baseline
Uncommitted

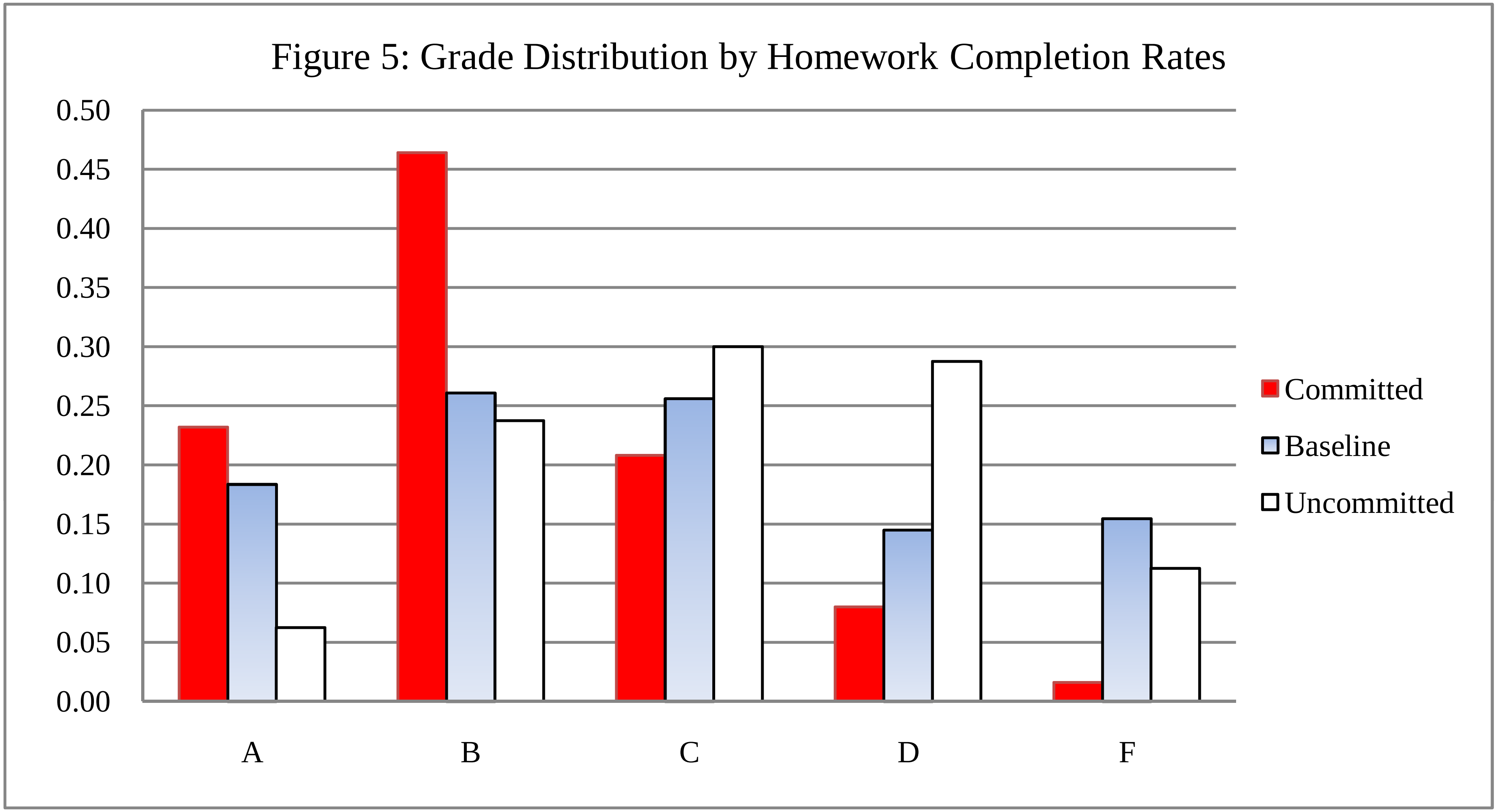


Figure 5: Grade Distribution by Homework Completion Rates
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
A
B
C
D
F
Committed
Baseline
Uncommitted

Table 1: Overview of Experimental Design for Classroom Experiment - Teaching Economics Principles Students

Semester	Students	Drop rate[1]	Econ Dept. Drop rate (std dev)	Homework Required?	Students HW req (took final)	Retention Rate if HW req	Students HW not req (took final)	Retention Rate if HW not req[2]	Instruction Days/Time	Classroom Seats
Spr '08	98	17.32% +	9.79% (5.5%)	Coin-flip	43 (42)	97.7%	55 (50)	90.9%	MWF 9am	110
Fall '08	108	6.90%	4.83% (2.1%)	Coin-flip	45 (43)	95.6%	63 (56)	88.9%	TuTh 3.30pm	110
Spr '09	105	2.68%	6.02% (4.0%)	Coin-flip	57 (54)	94.7%	50 (47)	94.0%	TuTh 2pm	110
Fall '09	110	0.00% -	8.01% (6.70%)	Coin-flip	61 (58)	95.1%	49 (41)	83.7%	TuTh 2pm	110
Total/Avg.	421	7.06%	7.16%		206 (197)	95.6%	217 (194)	89.4%		

[+]Drop rate is one-standard deviation larger ([-] indicates smaller) than Economics Department average.

[1] We employ a Mann-Whitney test to determine if the distribution of drop rates differs between the classroom experiment courses and departmental principles courses. The Z = 0.197 (p-value 0.84) indicates that we cannot reject the hypothesis of equivalent drop rates.

[2] We also use a Mann-Whitney test to determine if the retention rates (i.e., do enrolled students take the final exam?) are equivalent for those whom homework is required vs. not required. The Z = -2.419 (p-value = 0.0156) easily rejects the hypothesis that retention rates are equal for homework required vs. homework not required students.

Table 2: Descriptive Statistics of Field Experiment for Principles Students

Treatment:		Coin-flip	
Homework Requirement:	HW Req.	Not Req.	Total
Observations Test 1	206	213	419
Observations Test 2	200	204	404
Observations Test 3	201	197	398
Observations Test 4	197	194	391
Test Performance			
Test 1	74.3	72.2	73.2
(Std. Dev.)	(12.7)	(14.0)	(13.4)
Test 2	77.5	73.2	75.3
(Std. Dev.)	(12.7)	(16.7)	(15.0)
Test 3	68.0	68.0	68.0
(Std. Dev.)	(15.7)	(18.8)	(17.3)
Test 4	66.6	64.9	65.8
(Std. Dev.)	(13.2)	(15.0)	(14.1)
Test Average	71.8	70.3	71.1
(Std. Dev.)	(10.7)	(13.3)	(12.1)
Failure Rate	5.3%	14.7%	10.2%
Homework Submission			
HW_Submit_100% - Test 1	74.4%	6.0%	39.4%
HW_Submit_100% - Test 2	68.6%	2.3%	34.7%
HW_Submit_100% - Test 3	63.3%	2.8%	32.3%
HW_Submit_100% - Test 4	64.7%	2.8%	33.0%
Proportion HWs Completed			
for Test1	91.1%	14.3%	51.8%
for Test2	87.2%	5.3%	45.3%
for Test3	79.2%	4.5%	41.0%
for Test4	73.4%	2.8%	37.3%
Homework Average (if submit)	75.6	80.7	76.8
(Std. Dev.)	(10.2)	(12.1)	(10.9)
Demographic Characteristics			
Male (Pct)	31.4%	40.1%	35.8%
White (Pct)	55.6%	62.7%	59.2%
Total Credit Hours	39.2	31.0	35.0
Semester Hours	14.3	14.1	14.2
Ability Measures			
SAT Verbal	498.4	499.5	498.9
(Std. Dev.)	(68.8)	(69.3)	(69.0)
SAT Math	522.8	530.5	526.8
(Std. Dev.)	(70.6)	(71.4)	(71.0)
High School GPA	3.11	3.07	3.09
(Std. Dev.)	(0.40)	(0.42)	(0.41)

Table 3: Test scores for students by homework requirement and submission rates

Session	Homework submitted:		
	<100%	100%[1]	Total
Test 1			
Homework not required	71.81% (n=200)	77.69% (n=13)	72.17% (n=213)
Homework required	71.59% (n=52)	75.26% (n=154)	74.33% (n=206)
Test average[2]			
Homework not required	69.30% (n=185)	74.67% (n=13)	69.65% (n=198)
Homework required	66.51% (n=49)	73.30% (n=152)	71.64% (n=201)

[1] 100% of homework submitted for Test 1 includes students who completed each of the first four homeworks. Since students were allowed to drop their "lowest" homework score, 100% of homeworks submitted for Test Average includes students who completed at least 13 of the 14 homework assignments.

[2] Test Average reflects the average across four tests. Should a student take only three of the four tests, the average reflects the performance on the three tests taken. Students who miss multiple tests are excluded from the Test Average above.

Table 4 : Effect of submitting homework, requiring homework, & average homework performance on Test 1 scores: OLS regressions

	(1) ln(Test 1)	(2) ln(Test 1)	(3) ln(Test 1)	(4) ln(Test 1)	(5) ln(Test 1)
HW_Submit_100%	0.0523 **	-	-	0.0785	-0.1536
	(0.0191)			(0.0603)	(0.1628)
HW_Required	-	0.0346 ^	-	0.0046	-
		(0.0190)		(0.0267)	
HW_Average[1]	-	-	0.0072 **	-	0.0059 **
			(0.0010)		(0.0015)
HW_Submit*HW_Req	-	-	-	-0.0319	-
				(0.0659)	
HW_Submit*HW_Avg	-	-	-	-	0.0023
					(0.0020)
constant	4.2548 **	4.2586 **	3.7259 **	4.2538 **	3.8070 **
	(0.0125)	(0.0143)	(0.0013)	(0.0147)	(0.1248)
Demographic characteristics?	No	No	No	No	No
Ability measures?	No	No	No	No	No
n	419	419	262	419	262
R-squared	0.017	0.008	0.173	0.018	0.183

Notes: Robust standard errors appear in parentheses. **, *, and ^ denote statistical significance at the 1%, 5%, and 10% level, respectively. HW_Submit_100% equals 1 if the student submits all four homeworks prior to Test 1 and submits at least 13 of the 14 homeworks for Test Average.
[1]Homework Average is conditional on submitting homework (i.e., missing homeworks are excluded). The sample size is reduced by the 157 students who fail to submit any homework before test 1.

Table 5: Effect of submitting homework, requiring homework, & average homework performance on average test scores: OLS regressions

	(6) ln(Test Average)	(7) ln(Test Average)	(8) ln(Test Average)	(9) ln(Test Average)	(10) ln(Test Average)
HW_Submit_100%	0.1478 **	-	-	0.0871	0.1369
	(0.0268)			(0.0608)	(0.1998)
HW_Required	-	0.0775 *	-	-0.0217	-
		(0.0337)		(0.0493)	
HW_Average[1]	-	-	0.0095 **	-	0.0092 **
			(0.0015)		(0.0020)
HW_Submit*HW_Req	-	-	-	0.0789	-
				(0.0736)	
HW_Submit*HW_Avg	-	-	-	-	-0.0003
					(0.0025)
constant	4.1487 **	4.1574 **	3.5040 **	4.1549 **	3.4660 **
	(0.0238)	(0.0285)	(0.1189)	(0.0293)	(0.1684)
Demographic characteristics?	No	No	No	No	No
Ability measures?	No	No	No	No	No
n	409	409	260	409	260
R-squared	0.041	0.013	0.141	0.042	0.251

Notes: Robust standard errors appear in parentheses. **, *, and ^ denote statistical significance at the 1%, 5%, and 10% level, respectively. HW_Submit_100% equals 1 if the student submits all four homeworks prior to Test 1 and submits at least 13 of the 14 homeworks for Test Average.

[1]Homework Average is conditional on submitting homework (i.e., missing homeworks are excluded). The sample size is reduced by the 149 students who fail to submit any of the 14 homeworks.

Table 6: Effect of submitting homework, requiring homework, & average homework performance on average test scores: OLS regressions controlling for demographic characteristics & ability measures

	(11) ln(Test Average)	(12) ln(Test Average)	(13) ln(Test Average)	(14) ln(Test Average)	(15) ln(Test Average)
HW_Submit_100%	0.1254 **	-	-	0.1375 **	0.1375
	(0.0252)			(0.0485)	(0.1919)
HW_Required	-	0.0716 *	-	0.0011	-
		(0.0324)		(0.0503)	
HW_Average[1]	-	-	0.0061 **	-	0.0060 **
			(0.0018)		(0.0022)
HW_Submit*HW_Req	-	-	-	-0.0133	-
				(0.0666)	
HW_Submit*HW_Avg	-	-	-	-	-0.0004
					(0.0024)
constant	2.8451 **	2.8005 **	2.7240 **	2.8435 **	2.7409 **
	(0.1949)	(0.2019)	(0.2146)	(0.2035)	(0.2373)
Demographic characteristics?	Yes	Yes	Yes	Yes	Yes
Ability measures?	Yes	Yes	Yes	Yes	Yes
n	355	355	224	355	224
R-squared	0.215	0.197	0.272	0.215	0.303

Notes: Robust standard errors appear in parentheses. **, *, and ^ denote statistical significance at the 1%, 5%, and 10% level, respectively. HW_Submit_100% equals 1 if the student submits all four homeworks prior to Test 1 and submits at least 13 of the 14 homeworks for Test Average.

[1]Homework Average is conditional on submitting homework (i.e., missing homeworks are excluded). The sample size is reduced by the 131 students who fail to submit any of the 14 homeworks.