A joint initiative of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute for Economic Research

Area Conference on

Economics of Education

03 - 04 September 2010 CESifo Conference Centre, Munich

How much is Students' College Performance Affected by Quantity of Study?

Hans Bonesrønning and Leiv Opstad

CESifo GmbH
Poschingerstr. 5
81679 Munich
Germany

Phone: +49 (0) 89 9224-1410
Fax: +49 (0) 89 9224-1409
E-mail: office@cesifo.de
Web: www.cesifo.de

How much is Students' College Performance Affected by Quantity of Study?

Hans Bonesrønning

Department of Economics, NTNU

and

Leiv Opstad

Trondheim Business School

Abstract: Recent educational reforms aim at improving school or college quality by improving the students' study incentives. However, surprisingly little is known about the effects of study on grade performance. This paper seeks to fill some of the gap by combining survey and administrative data from one Norwegian business school. A differences- in-differences approach exploiting within-student variation in efforts within the same subject across two time periods is used to generate credible evidence. We find that grades are improved when students provide more efforts. The estimated effects are of considerable size, although smaller than those reported by Stinebrickner and Stinebrickner (2008).

Keywords: higher education, education production, student effort

JEL codes: I20, I21, J24

1. Introduction

Most people would probably agree – from introspection - that study effort is a crucial determinant of academic success. Consistent with this view, economists frequently make such assumptions in theoretical models of education production (see for instance Correa and Gruver (1987), Costrell (1995), Hoxby (1999)), and policy makers design educational reforms that aim at improving school or college quality by improving the students' study incentives. These beliefs stand in stark contrast to the existing empirical evidence. Even though the empirical literature is rather thin – probably reflecting that economists do not often collect data from within the black box of education production- the small literature that exist fails to provide consistent and supportive evidence that student effort is important. Sigfried and Fels (1979), in an early survey of the literature, conclude that student effort do not seem to be related to performance on standardized test, but that class attendance seem to be important for performance on some tests. Perhaps the most famous contribution is Schuman et al (1985) who, in a ten year study, investigate the relationship between study effort and student performance at the University of Michigan, and find no strong evidence that there is a positive relation between study and Grade Point Average. (See also the following-up discussions in Hill (1991), Rau and Durand (2000), and Schuman (2001)). Other examples are Glearson and Walstad (1988), and Krohn and O'Connor (2005) who report a negative association between student performance and study time. On the other hand, a number of studies (i.e Park and Kerr (1990), Romer (1993), Marburger (2006)) have reported positive correlations between attendance and student performance. Betts (1996) is an early study finding a positive effect of homework on high school students' math performance.

The correlation between student performance and student effort could take any sign (negative, positive, no relationship) in non-experimental data. Most likely there will be a positive correlation if more able students study harder than less able students and a negative correlation if less able individuals systematically compensate for these disadvantages by providing more efforts than do more able individuals. In a recent contribution, Stinebrickner and Stinebrickner (2008) deal with the endogeneity problems by exploiting random assignments of roommates at Berea College. Students are portioned into groups that are identical in all respects except that the students are assigned roommates of different observable type: some roommates have a videogame, others don't. Stinebrickner and Stinebrickner show that students that are assigned a roommate with a videogame performs poorer than students with a roommate without such a game, and, by using the roommates' videogame as an identifying variable in an IV approach, they provide evidence that the negative effect on performance is mediated through the students' study efforts. Their IV-estimates are almost 10 times the OLS-estimates, indicating that the returns to study efforts are very large (one standard deviation in effort transforms to 0.9 standard deviations in performance) and that the endogeneity problems are quite serious.

In the present paper, we use a combination of survey and administrative data from one business school in Norway to investigate the potential importance of students' study effort and attendance on achievement. Our strategy is to exploit within-student variation in effort allocations within one subject for two subsequent periods. This fixed-student, fixed-course approach addresses the most nearby endogeneity problems by differentiating away important time-invariant unobservable factors such as student ability.

However, there are other and less obvious endogenity problems related to "dynamic selection effects" to worry about. Examples are that students may respond (non-randomly) to difficult classes or bad grade shocks by increasing their efforts. Stinebrickner and Stinebrickner (2008) provide evidence that such effects explain much of the gap between the

OLS- and IV-estimates in their study. They warn that fixed effect approaches (that focus only on problems related to unobserved student characteristics) might magnify the endogenity problems that are related to dynamic selection.

To clarify our identification strategy, and also address the concerns of Stinebrickner and Stinebrickner, it is convenient to start out from a value added education production function (EPF) augmented with student effort; which means that the single course grade is regressed against prior achievement, individual student characteristics, school inputs and the student efforts provided in that particular course. By controlling for prior achievements we most likely reduce correlations between effort and the residual which otherwise will cause the estimated effect of effort on performance to be biased. However, we cannot be sure that this approach will take us all the way: Prior achievement not only reflects student's ability, but also the unobserved efforts provided earlier in the schooling career, implying that two students with identical prior achievements might have different returns to their efforts in college. This is to say that the value added approach does not take away all types of unobserved student heterogeneity and hence, does not provide credible effects of student effort on college grades.

To solve the remaining problems we exploit information from one course (Macroeconomics) that has a mandatory multiple choice mid-semester test, and we use the difference in performance at exam and the mid-semester test as the dependent variable in a regression analysis against the difference in efforts reported after and before the mid-semester test. In this case, the effort estimates are identified from students that actually change their efforts from the first to the second period of time. The estimates are credible to the extent that the individual student effort allocations across the two periods are not determined by unobserved characteristics that are correlated with student performance. We take several steps to evaluate the restrictiveness of this assumption.

We find a significant and positive relationship between study efforts and test scores. Our results indicate that one standard deviation in effort translates into 0.25 standard deviations in performance. These results thus differ from the much cited Shuman-study (1985) by finding positive effects of study effort on performance, and from Stinebrickner and Stinebrickner (2008) by finding smaller effects than they do.

The remainder of this paper proceeds as follows. In the next section we present our data. Section 3 offers a few theoretical considerations, before the econometric approach is laid out in section 4. Section 5 presents the results. In this section we also illustrate the importance of the identifying assumptions underlying the DD-approach by exploiting effort variations across subjects for individual students. Section 6 concludes.

2. The data

The data come from Trondheim Business School, and cover the cohorts of students starting in the falls of 2005, 2006 and 2007. The survey data are collected for the purpose of investigating the students' study habits. For this project, these data are mixed with administrative data for the students' prior achievements in the upper secondary school.

Among the four courses the students take in the fall semester, Macroeconomics is the only course with a mid-semester test. The students taking this course are asked about their efforts twice; the first time right before the mid-semester test and the second time in the end of November just before the exam. In the last of these surveys, the students were also asked for detailed information about the time allocations to all the four different courses they took that fall and the total time allocated to studying. Approximately 60% of all entering students have participated in the surveys. The sample contains a little more than 150 students in Macroeconomics - for which we have complete lists of data. The students that have participated in the survey have somewhat higher GPAs from the upper secondary school than those students that did not participate: The average GPA for all students is 53.24, while participating students have an average GPA of 53.45. We return to the potential consequences of missing student records from students that are unwilling to provide data in the result section.

Efforts are about both the quantity and efficiency of time. There is no consensus about the relevant dimensions of efficiency of time (which for instance might be determined by the use of adequate learning strategies, access to academic input, and so on). No attempts are made to capture such features. The students are asked the following questions: "How many Macroeconomics lectures per week have you attended?" "How many hours per week have you spent on out-of-classroom work in Macroeconomics?" The same types of questions are asked for the three other subjects. We use the average attendance rates in lectures (ATTENDANCE) and the hours per week the student spend working on their own (STUDY) to measure the variables of interest. It is well-documented that reporting error from retrospective questions might be substantial. We discuss remedies to deal with such problems in the next section.

(Table 1 about here)

According to Table 1 attendance rates are 88 percent in Macroeconomics classes, with only minor differences between the pre- and post-mid-semester periods. The students spend on average 3.2 hours per week on STUDY before the mid-semester test and 3.5 hours after the mid-semester test.

The administrative data provide information about the students' multiple test scores for Macroeconomics (both at the mid-semester test and the exam), their course grades at the Business school, their Grade Point Average from the upper secondary school, and their gender. Both the mid-semester test and the exam consist of multiple choice tests with 32 questions. The average student performance at the mid-semester test and the exam are 16.2 points and 20.6 point respectively. The Business school admits high performing students from the upper secondary school. Thus the average student has a performance level of 53.2 points (the sum of the grades (scale 6-0) for 10 subjects), which says that the average performance level is somewhat better than "B", while the average student in upper secondary school performs between "C" and "D". Thus, the Business school students have higher abilities and most likely, have provided substantial own efforts in their prior schooling career than the average upper secondary student. There is a slight majority of girls (52 percent).

Table 2 provides descriptive statistics for two of the other courses the students take the same semester. We return to these data towards the end of the paper.

(Table 2 about here)

3. The determinants of student effort : Theoretical considerations

Our main purpose is to quantify the effects of study efforts on student performance. Most likely, these effects differ across students and across subjects. To fix ideas we shortly consider the individual student's effort decision problem. This is discussed in length by several authors, such as Becker (1982), and Correa and Gruver (1987), and more recently by Krohn and

O'Connor (2005). Here we make use of a very parsimonious model where the individual student is portrayed as a rational person that cares about the grades in all subjects as well as leisure time. The grade in a given subject is assumed to be an increasing function of student ability and the effort allocated to that subject (the achievement production function), conditional upon teacher and peer quality. In the simplest set-up, the individual student determines the effort allocation across subjects and leisure time by maximizing his/her utility function subject to the achievement production functions for all subjects and a time constraint. This implies that the optimal efforts and the optimal grades will be functions of student ability, teacher and peer quality, and the characteristics of the achievement and utility functions.

From this simple model it follows that the observed efforts as well as the observed returns to studying efforts will vary across students. Without adding more structure to the model, it is hard to make more precise predictions. For instance, we cannot say whether high ability students will provide more or less effort than low ability students. If effort and ability are complementary inputs; high ability students will experience higher returns to additional efforts than do low ability students for the same level of effort. Preferences also matter: students with strong performance preferences might for instance put in much efforts and experience decreasing returns to additional efforts (these students might have to put in much additional effort to increase their performance from grade B to grade A).

The individual student will allocate different amounts of studying efforts across subjects in a given period of time because (s)he experiences that the returns to study effort – and the marginal utility from improvement in grades - differ across the subjects. One example of the former is that a low ability student might allocate a lot of effort to the most difficult courses – to be sure that she fulfills the minimum pass requirements - even though the returns to these efforts in terms of improved grades are small. One example of the latter is that a student planning an investment career might allocate much more efforts to Investment Analysis than to Macroeconomics, despite the fact that the marginal effort returns to Investment (measured as the number of correct answers) are smaller than in Macroeconomics.

Small expansions of the model are needed to discuss variations in effort allocations within a subject over time. This kind of variation might reflect the students' time discount rates. Myopic students might be unwilling to allocate much time to study in the beginning of the semester, but might be willing to increase their efforts towards the exam. The returns to these efforts might be rather small if a firm understanding of the basic material provided in the first part of the semester is required for making progress. An increase in effort from the first to the second period might also reflect a more difficult course material in the second half of the semester. Or, in the case where the course has a mid-semester test, the students might adjust their efforts dependent on whether they achieve better or worse than expected on this test.

This theoretical discussion has left out several important issues. A couple of examples should be mentioned. First, students might reallocate effort from studying to leisure if they realize increasing returns to their own leisure time (the situation exploited by Stinebrickner and Stinebrickner (2008)). Second, peers and teachers might influence the effort allocations. For instance, good peers might induce students to increase their efforts during a semester, or a good teacher might have positive effects on the returns to student effort.

4. The econometric approach

Identification

We start out from a simple value added education production function (EPF) for each of the courses:

$$A_{i,j} = \alpha_{0,j} + \alpha_{1,j}E_{i,j} + \alpha_{2,j}X_{i,j} + \alpha_{3,j}S_j + \varepsilon_{i,j} \quad (1)$$

where $A_{i,j}$ denotes achievement for student i in course j, $E_{i,j}$ denotes student effort in course j, $X_{i,j}$ is a vector of individual and family background characteristics including the student's GPA, and Sj denotes a vector of college inputs. ε is the residual.

This is a reduced form education production function augmented with student effort. Sometimes the chosen formulation is referred to as an attainment model because the lagged dependent variable is included among the independent variables. The merits of this approach are much discussed (see for instance, Allison (1990), Hanushek (1987) and Todd and Wolpin (2003)). Here this specification is chosen because it is a convenient point of departure for clarifying the assumptions that are necessary for the differences in differences approach. A stated above, student effort is determined by student preferences and other student characteristics that are unobservable to the researcher, implying that the estimated effects of effort on student achievement are biased. The estimates will be biased downwards if effort is negatively related to student ability, and only imperfect measures of student ability are included among the covariates. There will be an upward bias if high ability students reinforce their genetic advantages by providing more efforts than less able students.

Taking the difference in achievement across two subjects – or across two periods for the same subject - we arrive at:

$$A_{i,j} - A_{i,k} = \gamma_{0,jk} + (\alpha_{1,j}E_{i,j} - \alpha_{1,k}E_{i,k}) + (\alpha_{2,j} - \alpha_{2,k})X_{i,j} + (\alpha_{3,j}S_j - \alpha_{3,k}S_k) + \vartheta_{i,j} \quad (2)$$

where υ is a residual involving the difference between error terms. Assuming that the returns to student and background characteristics are equal across the two subjects/periods, the observable background characteristics cancel out ($\alpha_{2,i} = \alpha_{2,k}$), and moreover, the residual contains no elements that cause it to be correlated with the individual student's efforts. This is one of the assumptions that are required for this specification to generate credible effort estimates

An additional reason why equation (2) might generate biased effort estimates is that adequate measures of teacher quality are not available. From a theoretical point of view, the relationships between teacher quality and student effort are complicated. The characteristics of the education production function and the students' utility functions are important. For instance, teacher quality might be complementary to student effort, implying that the returns to student effort increase when teacher quality increases. But not all students might respond to higher teacher quality by providing more efforts; generally the response depends on the characteristics of their utility functions. Equation (2) cannot fully address this problem. Thus, there is a potential omitted variable problem. However, in the fixed-student, fixed-course approach the teacher characteristics are - under reasonable assumptions - differentiated away. It remains though that our results are conditional upon the teacher quality at the Trondheim Business School, and cannot be broadly generalized.

Thereafter we impose the important DD-assumption that the returns to the efforts provided in both subjects/periods are equal: $\alpha_{1,i} = \alpha_{1,k} = \gamma_1$. The differences in differences (DD) equation to be estimated is then:

$$A_{i,j} - A_{i,k} = \gamma_0 + \gamma_1 \Delta E_{i,jk} + \vartheta_{i,jk} \quad (3)$$

Equation (3) is the adequate specification for the fixed-student, fixed-course approach. In this case the variation in $\Delta E_{i,jk}$ comes from students who provide different amounts of effort in subsequent periods within one course.

An important question is thus whether the within-student, within-course, between-periods variation in effort is reasonable exogenous to student performance. Since the students select themselves into the student subgroups that increase their effort in the second period, decrease their effort in the second period, or keep their efforts unchanged, the question is whether this sorting is based on observable and unobservable characteristics that are correlated with student performance. We provide three pieces of empirical evidence to shed light on this issue. First, we investigate whether the change in efforts is related to observable student characteristics. If there are no systematic relationships between efforts and observable student characteristics, it seems less likely that changes in efforts across time periods are driven by unobservable student characteristics. Second, we investigate whether the average student increases her/his effort from the first to the second period. One hypothesis is that most students do that (because the exam is more high stake than the mid-semester test and/or because the students are myopic), and hence some of the effort effects are identified from this kind of behavior. Third, we investigate whether the effort response is correlated with the difference between expected grade in the macroeconomics course and the test performance. This exercise is motivated by the hypothesis that students decide their initial efforts based on an expected course grade, and adjust their efforts if the test results show that this expectation is unlikely to be fulfilled.

The assumption that the returns to the efforts provided in the two periods are equal might be too restrictive. For instance, the Macroeconomics course might be more difficult towards the end. To evaluate the restrictiveness of this assumption we estimate

$$A_{i,j} - A_{i,k} = \gamma_0 + \alpha_{i,k}(E_{i,j} - E_{i,k}) + (\alpha_{1,i} - \alpha_{1,k})E_{i,j} + \mu_{i,jk} \quad (4)$$

Equation (4) is a reformulation of equation (3) with $\alpha_{2,i} = \alpha_{2,k}$, $\alpha_{3,i} = \alpha_{3,k}$, and $\alpha_{1,i} \neq \alpha_{1,k}$.

Measurement error

The variables of interest – the students' efforts - are likely reported with error; implying that the estimates of effort on college grades have a downward bias. The most adequate way to reduce measurement errors of this type would be to collect time-use information at more than one time (see Stinebrickner and Stinebrickner (2003, 2004)). No such additional information has been available. Note however that if students systematically underreport their efforts (across courses or across periods of time), the DD-approach effectively does away with this bias. Some additional steps are nonetheless taken to minimize the measurement problems. The most important is that we examine the robustness of our results by excluding outliers.

5. Results

OLS estimates for all courses

To provide a point of reference, OLS-estimations of the value added education production function, as portrayed in equation (1), are estimated. In addition to Macroeconomics, the equation is also estimated for the two courses Investment Analysis and Accounting Information Systems, and in all cases we use the college grades as our outcome variable. All courses at the Business School are graded on the same A-F scale, where A is the best grade and F is failure. In this paper the grades are converted to a 5-0 scale, where 5 is equivalent to A, and 0 is equivalent to F. The results reported in Table 3 show that there is a significant association between attendance and performance in a majority of the courses.

STUDY is positively associated with performance in all courses, but the estimates and their precision vary much across the subjects. One standard deviation in STUDY transforms into 0.13, 0.09 and 0.01 standard deviations in the grades in Accounting Information Systems, Macroeconomics, and Investment Analysis respectively. In none of the cases are the STUDY estimates significantly different from zero. There are several critical issues that are not raised in these initial analyses, for instance related to the grading practices across subjects, which limit the value of this exercise. Nonetheless, this exercise demonstrates that the effort effects might vary across subjects. We should keep this in mind when we now turn to Macroeconomics for the DD-analysis.

(Table 3 about here)

Exploiting within-student. within-course, between- periods variation

Table 4 reports the results from the DD-analysis for the Macroeconomics course. Note that in this case we use the difference in the number of correct answers at the two multiple tests as our outcome variable. This measure is superior to a grade-based measure that typically suppresses much information and is subject to much more subjective judgments. It is evident that an increase in study efforts leads to better test performance. The effect; which is significant at the 1 percent level, is of some considerable size: using the results from the most parsimonious specification, we find that one standard deviation (2.45 hours) increase in study efforts per week increases test performance with 0.26 standard deviations in the December test. This estimate, which is almost three times the OLS-estimate reported above, is not much affected when controls for different cohorts and for student background characteristics are added as controls – indicating that there is no serious sorting into effort changes.

Our OLS estimates are of approximately the same size as those reported by Stinebrickner and Stinebrickner (2008). As in their study, the estimates increase substantially when we address the endogeneity problems. However, while their IV-estimates are nearly 10 times their OLS-estimates, our DD-results are a little more than 3 times the OLS-estimates. There are several potential explanations for this difference. One is that the Berea grading practices might reward student effort, another that we have been unable to do away with all the endogeneity problems. Below we perform some exercises that potentially might shed light on the latter issue by examining the likelihood that the assumptions underlying the DD-approach are fulfilled.

(Table 4 about here)

The specifications reported in the three first columns of Table 4 apply the assumption that the returns to own study efforts are equal across the two periods of time, that is, before and after the mid-semester test. This seems like a reasonable assumption, but it might nevertheless be argued that the Macroeconomics course is tougher towards the end than in the beginning – or that students who have put in little effort in the first period do not stand on firm ground; which potentially implies that the returns to their own efforts are smaller in the second period. To evaluate this claim we have included the level of study in the first period among the explanatory variables (i.e. we have estimated equation (4)). There is some weak evidence that the returns to the study efforts are larger in the first period (0.588) compared to the second period returns (0.588-0.194 = 0.354), but the estimated coefficient for the *level* variable STUDY of -0.194 is of poor precision and not significantly different from zero. Thus, the hypothesis that the returns are similar for the two periods is not formally rejected.

Another crucial identifying assumption is that that the changes in efforts between periods are not due to unobservable student characteristics that are correlated with student performance. As mentioned above, several steps are taken to evaluate the restrictiveness of this assumption. First, note that neither gender nor the GPA from the upper secondary school is significantly associated with the change in test performance (see Table 4, column (3)). Although the estimated ΔSTUDY-coefficient decreases with nine percent when these controls are included, the estimate is still significant at the 1 percent level. Thus, there are no strong indirect indications that the changes in efforts from the first to the second period are associated with observable student characteristics.

More direct evidence is provided from an equation with ΔSTUDY as the dependent variable and student characteristics as independent variables. As can be seen from Table 5, neither gender nor GPA is significantly associated with the change in study efforts. These findings -especially the lack of correlation between GPA and effort- might be taken as indications that the effort choices across the two periods are not driven by troublesome unobservable student characteristics.

(Table 5 about here)

Nonetheless, it remains that some students for some reasons choose to provide more efforts in the second period. One hypothesis is that this variation reflects the students' responses to the information content in the mid-semester test result (see Krohn and O'Connor (2006)). More specifically, students who perform poorer than expected at the mid-semester test might decide to increase their effort in the subsequent period. Since the students in the first questionnaire are asked about their expected performance at the mid-semester test, we use this answer to generate an additional explanatory variable which is the difference between expected and actual performance at the mid-semester exam. This new explanatory variable is positively correlated with ΔSTUDY, implying that students who perform poorer than expected at the first test increase their efforts in the next period. This finding is consistent with Krohn and O'Connor (2006). The estimated coefficient is rather small and not significantly different from zero, probably reflecting that the reported expectation is most likely to be a very noisy variable. It is not clear what we econometrically should make out of this finding.

A somewhat different hypothesis is that the most ambitious students work hard right from the beginning of the semester, and thus, do not change their efforts much from the first to the second period, while the less ambitious students increase their efforts towards the exam. If this is the case, much of the estimated effects are driven by the latter subgroup, and we might expect that our effort estimates are biased downwards. We have made an attempt to evaluate this hypothesis by breaking the sample into two; one subsample consisting of students that increase their efforts in the second period, and one subsample consisting of students that decrease their efforts in the second period. This exercise, which clearly can be criticized for separating students according to an endogenous variable, reveals that the estimated effort returns are driven by students who decrease their effort in the second period. The estimated effort coefficient for this subsample is 0.776, while the comparable estimate for the full sample is 0.686. The effort estimate for the other subsample is thus much smaller (0.350) and statistically insignificant. These findings, which are not reported in the tables, are consistent with the hypothesis that the students who work relatively hard from the beginning of the semester have higher returns to their efforts in the last part of the semester. We cannot tell whether this is due to selection (for instance, these students are less myopic than other students), or due to the fact that the returns to own efforts in the second period depends on the investments made in the first period.

About 40 percent of the students did not provide data, either because they were unwilling to do so or because they were absent from class on the day the data were collected.

Neither of these two events is likely random. For instance, we might conjecture that at least some of the absent students provide less effort than the average student. Our hypothesis is that a student's failure to complete a questionnaire provides direct information on the student's performance, and hence, indirect information about their study efforts. (see Becker and Powers (2001)). To evaluate this conjecture we have estimated the value added equation (equation (1)) with all students included and with a dichotomous variable that is 1 if the student participated in the survey and 0 otherwise among the independent variables. (not completed).

Finally, a couple of additional results should be noted. The last column of Table 4 reports the results when students in the two first and the two last deciles are excluded. There is no evidence that the returns to own study efforts are larger for students in the middle of the distribution than for students in the tails. And, there is some weak evidence that attending the lectures on a more regularly basis leads to improved student performance. The effect, which is rather small, is significant at the 7 percent level.

Exploiting within-student, across-subjects variation

Here we shortly present and evaluate an alternative approach that exploits effort variations across two different courses. We can choose from a menu of three courses; Macroeconomics, Investment Analysis and Accounting Information Systems. Table 2 provides the essential descriptive statistics for the two "new" courses. Attendance is 80 percent in Investment Analysis classes, and 56 percent in Accounting Information Systems classes. The variation in attendance is substantial across the subjects. Also the between-student variation in attendance within each of the courses is of some considerable size, as indicated by the standard deviations of 12-30 percent. STUDY varies from slightly more than

5 hours per week in Investment Analysis, to 3.2 hours in Macroeconomics, and about 1.5 hours in Accounting Information Systems. (The differences in STUDY across the courses partially reflect that the courses are of different size: Investment Analysis gives 10 credits, Macroeconomics gives 7.5 credits, and Accounting Information Systems 5 credits.)

No multiple choice test results are available for the DD-analyses that exploit between-course effort variation. In this case we therefore make use of the students' grades. The average performance in the three courses varies from 3.12 in Macroeconomics to 2.42 in Investment Analysis, with Accounting Information Systems in a middle position with an average of 2.95.

Within the available menu of courses, Macroeconomics and Accounting Information Systems stand out as the most similar courses: they have approximately the same average returns to GPA (see Table 2), and the marginal returns to GPA do not differ much (see Table 3).The DD-approach also imposes the requirement that the returns to own efforts are equal across the two subjects. To provide some indications whether this requirement is likely to be fulfilled, we have estimated "effort equations" for the three courses under consideration; that is, the efforts provided in the course are regressed against GPA and gender. The results from these exercises are reported in Table 6.

(Table 6 about here)

There appear to be no statistically significant effects of GPA on study effort; which is similar to the findings reported by Stinebrickner and Stinebrickner (2008). However, for Macroeconomics and Accounting Information Systems the positive signs of the GPA estimates in the STUDY equation indicate reinforcing responses. Note also that attendance is not systematically related to the background characteristics for any of the courses.

The left hand side of Table 7 reports the results from estimating equation (3) with Macroeconomics and Accounting Information Systems as the two subjects. By controlling for cohorts (column (2)), the ΔSTUDY estimate is 0.068, and not significant at conventional levels. One standard deviation in the ΔSTUDY variable transforms into 0.1 standard deviations in ΔGRADE. This effect is between ½ and ⅓ of the effect estimated from within-course variation in effort.

(Table 7 about here)

The results from estimating the less restrictive equation (4) are reported in column (3). The ΔSTUDY estimate is then 0.126 with p-value 0.05. For the current specification, this is the estimate for the effort returns in Accounting Information Systems. The sign of the STUDY coefficient is negative, indicating that the returns to STUDY might be smaller in Macroeconomics than in Accounting Information Systems (0.126- 0.084 = 0. 042). Notice however that the coefficient for STUDY is not significantly different from zero.

The effects of studying time do not differ much between column (3) and the OLS estimations reported earlier. The column (3)-estimate of 0.126 is approximately equal to the OLS estimate for Accounting, and taking the insignificant estimate for the level variable STUDY in Macroeconomics into account the estimated STUDY-effect for this course is close to the OLS-estimate of 0.048. In one interpretation the DD-estimations reported in columns (1) and (2) thus provide an imprecise average of the returns to efforts for the two subjects. The estimated average effort effect on performance is much smaller than the within-course estimate for Macroeconomics, indicating that this approach is unable to do away with unobserved student characteristics that bias the estimated coefficient downwards. The two courses are simply "not equal enough".

Comparison with the Stinebrickner and Stinebrickner study

Our most credible estimates - one standard deviation in study efforts transforms into 0.25 standard deviations in performance in Macroeconomics - is a little less than ⅓ of the Stinebrickner and Stinebrickner's (2008) effect of STUDY on GPA. Indirectly we have pointed to a number of potential explanations for this difference. First, some of our evidence indicates that the effort returns differ across subjects. The Stinebrickner and Stinebrickner study might have included subjects with higher returns to study efforts than the Macroeconomics course investigated here. Second, student quality might differ between the universities/colleges under study: when we break our sample according to the GPA from the upper secondary school it is evident that the returns to studying efforts are higher for students in the middle of the GPA distribution than for students located in the tails. Appendix Table 1 provides evidence from exploiting between-course variation in efforts, that students between the 2nd and 5th deciles of the GPA distribution experience marginal returns to STUDY of the same size as that reported by Stinebrickner and Stinebrickner (2008). Finally, the DD-approach might not have done away with all the unobserved student characteristics that bias the estimates downwards. More research, using IV- and DD- approaches, is required to pin down the exact size of the study effort effects.

6. **Conclusion**

In many countries, policy makers are looking for reforms that will improve the academic performance of the (higher) education institutions. Incentive-based reforms with the aim of increasing students' efforts are among the most popular proposals. Unfortunately, we do not know much about how such reforms will work. One part of these discussions is about what kind of incentives that are likely to be the most effective ones. At a more fundamental level, no incentives-oriented policies in higher education are likely to succeed if college outcomes are driven by background factors that are determined before students arrive at college. There is little empirical evidence that can shed light on these issues.

In this paper we seek to fill some of this gap. Using data for three subsequent cohorts of students in one Norwegian business school we show that student effort is an important determinant of test scores/college grades. The average returns to the study effects seem to be quite large in the second year course Macroeconomics. One standard deviation in study time per week transforms into approximately 0.25 standard deviations in performance. These results might indicate that incentives directed towards students' study efforts will improve their performance.

Can we believe in the results? The present study solves the most obvious endogeneity problems - related to unobservable student characteristics - by utilizing within-student variations in effort allocations within the same subject for different time periods. Several pieces of evidence are provided showing that the effort variation used for identification is not "suspect". In addition, the assumption of equal returns to own efforts in two periods for the same course is likely to be fulfilled by good approximation. However, it remains that the effort estimates vary widely across different studies. To some extent this is due to different methodological approaches, but some of the evidence provided above indicates that the returns to own efforts depend on the courses investigated and also, the individual student's allocation of efforts across the semester. More research is needed to settle this issue.

Table A1 Heterogeneous effects across the GPA distribution. Across-subject variation using Macroeconomics and Accounting Information Systems

	[20%, 80%]	[20%, 50%]	[50%, 80%]
ΔATTENDANCE	.002	-.008	.004
	(.005)	(.009)	(.006)
ΔSTUDY	.188	.338	.028
	(.085)	(.131)	(.132)
ATTENDANCE macro	-.016	-.013	-.020
	(.011)	(.017)	(.016)
STUDY macro	-.068	-.166	.019
	(.085)	(.164)	(.104)
R^2_{adj}	.042	.009	.00
N	83	47	36

Note: Cohort dummies included as control variables are not reported in the table.
Standard errors in parentheses.

References

Allison, P. D. (1990) Change scores as dependent variables in regression analysis. *Sociological Methodology*, 20: 93-114

Becker, W.E and J. Powers (2001) Student performance, attrition, and class size given missing student data. *Economics of Education Review*, 20: 377-388

Becker, W. E. (1982) The educational process and student achievement given uncertainty in measurement. *American Economic Review*, 72(1): 229-236

Betts, J. (1996) The role of homework in improving school quality. Discussion paper 96-16, University of California at San Diego

Correa, H., and G. W. Gruver. (1987) Teacher-student interaction: a game-theoretic approach to the economic theory of education. *Mathematical Social Science*, 13:.19-47

Costrell, R.M. (1995) A simple model of educational standards. *American Economic Review*, 84(4): 956-971

Fredrick, W.C., and H.J. Walberg (1980) Learning as a function of time. *Journal of Economic Education* (March/April): 183-84

Glearson, J.P., and W.B. Walstad (1988) An empirical test of the inventory model of student study time. *Journal of Economic Education* (Fall): 315-321

Hill, L. Jr. (1991) Effort and reward in college: a replication of some puzzling findings. In Replication Research in the Social Science, by James W. Neuliep (ed), Newbury park, CA: Sage, 139-56

Hoxby, C. M. (1999) The productivity of schools and other local public goods producers. *Journal of Public Economics*, 74: 1-30

Krohn, G. A. and C. M. O'Connor (2005) Student effort and performance over the semester. *Journal of Economic Education*, 36(1): 3-28

Marburger, D.R. (2006) Does mandatory attendance improve student performance. *Journal of Economic Education* 37(2): 148-155

Park, K.H., and P.M. Kerr (1990) Absenteeism and undergraduate exam performance. *Journal of Economic Education* (spring): 101-111

Rau, W., and A. Durand. (2000) The academic ethic and college grades: does hard work help students to 'make the grade'? *Sociology of Education,* 73: 19-38

Schuman, H., E. Walsh, C. Olson, and B. Etheridge. (1985) Effort and reward: the assumption that college grades are affected by the quantity of study. *Social Forces*, 63: 945-66

Schuman. H. (2001) Comment: Students' efforts and reward in college settings. *Sociology of Education*, 74(1): 73-74

Siegfried. J., and R. Fels (1979) Research and teaching of college economics: A survey. *Journal of Economic Literature*, 17(3), 923-969

Stinebrickner, R., and T. Stinebrickner. (2003) Working during school and academic performance. *Journal of Labor Economics*, 21(2): 473- 491

Stinebrickner, T., and R. Stinebrickner. (2004) Time-use and college outcomes. *Journal of Econometrics*, 121 (1-2): 243-269

Stinebrickner, T., and R. Stinebrickner. (2008) The causal effect of studying on academic performance. *The B.E. Journal of Economic Analysis & Policy*, 8(1), Article 14

Weinberg, B. A., B.M. Fleisher and M. Hashimoto (2009) Evaluating teaching in higher education. *Journal of Economic Education*, 40(3): 227-261

Table 1 Descriptive Statistics for the Within-Course Analysis

	Mean	Std. Deviation	N
Gender	.46	.500	153
GPA	53.78	3.54	153
Test October - #correct answers	16.17	4.94	153
Exam December – #correct answers	20.58	4.65	153
STUDY (per week) after mid-semester	3.54	2.45	153
STUDY (per week) before mid-semester	3.19	1.89	153
ATTENDANCE (percentage) after mid-semester	88.21	11.92	153
ATTENDANCE (percentage) before mid-semester	88.32	12.81	153
ΔATTENDANCE	-.11	9.20	153
ΔPERFORMANCE	4.41	5.63	153
Expected performance – actual performance /mid-semester	-2.32	4.64	153

Table 2 Descriptive Statistics for the Across-Course Analysis

	N	Minimum	Maximum	Mean	Std. Deviation
Male	377	0	1	.47	.500
GPA	375	43.50	69.60	53.07	4.14
Grade Macroeconomics	416	0	5	3.12	1.15
Grade Investment Analysis	395	0	5	2.42	3.08
Grade Accounting Information Systems	255	0	5	2.95	1.25
STUDY Macroeconomics	299	.00	12.50	3.27	2.19
STUDY Investment Analysis	290	.0	20.0	4.96	3.52
STUDY Accounting Information Systems	150	.0	10.0	1.55	1.71
ATTENDANCE Macro	302	20.00	100.00	87.74	12.34
ATTENDANCE Finance	291	6.67	100.00	79.59	16.13
ATTENDANCE Accounting Information Systems	150	.00	100.00	55.89	31.37
ΔSTUDY (Macro – Finance)	288	-18.00	4.00	-1.70	3.11
ΔSTUDY (Macro-Accounting Information Systems)	148	-9.00	9.00	1.92	2.21
ΔATTENDANCE (Macro-Finance)	291	-46.67	86.19	8.18	15.99
ΔATTENDANCE (Macro-Accounting Information Systems)	150	-13.33	88.89	32.54	27.25
ΔGRADES (Macro- Finance)	388	-5.00	4.00	.63	2.88
ΔGRADES (Macro – Accounting Information Systems)	252	-3.00	4.00	.34	1.28

Table 3 OLS estimation results for the value added education production function

	Macroeconomics	Investment Analysis	Accounting Information Systems
Male	.212	.256	-.050
	(.131)	(.189)	(.194)
GPA	.066	.076	.057
	(.016)	(.022)	(.026)
Attendance	0.023	.015	.010
	(.006)	(.006)	(.003)
STUDY	.048	.008	.101
	(.032)	(.017)	(.056)
R^2_{adj}	.131	.058	.110
N	264	258	138

Note: Standard errors in parentheses

Table 4 Differences-in-differences estimation using Macroeconomics after and before the mid-semester test

	All students	All students	All students	All students	Students in [20%, 80%]
ΔATTENDANCE	.081	.080	.083	.080	.117
	(.045)	(.045)	(.047)	(.049)	(.057)
ΔSTUDY	.490	.479	.447	.588	.491
	(.184)	(.187)	(.194)	(.251)	(.288)
2008-cohort		-.283	-.165	.018	.247
		(.809)	(.842)	(.871)	(1.090)
Gender			.062	.006	1.438
			(.829)	(.834)	(1.075)
GPA			.116	.115	-.095
			(.112)	(.112)	(.369)
STUDY				-.194	-.132
				(.224)	(.255)
ATTENDANCE				-.001	.043
				(.038)	(.046)
Constant	4.113	4.262	-2.188	-1.464	4.726
	(.402)	(.587)	(6.003)	(6.898)	(19.613)
R^2_{adj}	0.048	0.043	0.032	0.025	0.032
N	183	183	176	176	104

Table 5 The Determinants of the change in study efforts across two periods. Macroeconomics

	All students
Gender	.005
	(.335)
GPA	.013
	(.047)
Expected performance-Actual performance	.052
	(.036)
Constant	-.262
	(2.562)
R^2_{adj}	.00
N	152

Table 6 OLS estimations of efforts against student background characteristics. All courses

	Macroeconomics		Investment Analysis		Accounting Information Systems	
	Attendance	STUDY	Attendance	STUDY	Attendance	STUDY
Male	-0.904	-0.099	2.28	-0.019	-8.68	0.088
	(1.51)	(0.267)	(1.95)	(0.435)	(5.26)	(0.29)
GPA	0.04	0.042	-0.03	-0.014	-0.71	0.062
	(0.18)	(0.031)	(0.23)	(0.054)	(0.71)	(0.039)
R^2_{adj}	0.00	0.01	0.01	0.03	0.01	0.01
N	266	266	264	264	142	142

Note: Standard errors in parentheses

Table 7 Differences-in-differences estimation using Macroeconomics and Accounting Information Systems

	All students	All students	All students	All students	Students in [20%, 80%]
ΔATTENDANCE	**.005**	**.005**	**.005**	**.004**	**.003**
	(.004)	**(.004)**	**(.004)**	**(.004)**	**(.005)**
ΔSTUDY	**.078**	**.067**	**.126**	**.126**	**.159**
	(.045)	**(.045)**	**(.064)**	**(.066)**	**(.080)**
2008 cohort		.142	.439	.325	-.251
		(.346)	(.343)	(.392)	(.496)
2007 cohort		.431	.122	.059	.036
		(.342)	(.349)	(.384)	(.516)
ATTENDANCE macro			-.002	-.003	-.014
			(.008)	(.008)	(.011)
STUDY macro			-.077	-.084	-.036
			(.063)	(.065)	(.081)
Gender				.205	-.144
				(.212)	(.260)
GPA				-.001	-.046
				(.031)	(.072)
R^2_{adj}	0.021	0.026	0.024	0.008	0.024
N	143	143	143	143	83

Note: Standard errors in parentheses