



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

Hans Bonesrønning and Leiv Opstad

Abstract

Recent educational reforms aim at improving school or college quality by improving 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 effort within the same subject across two time periods is used to generate credible evidence. We find that grades are improved when students put in more effort. The estimated effects are of considerable size, although smaller than those reported by Stinebrickner and Stinebrickner (2008).

JEL classification: 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); Akerlof and Kranton (2002)). Policy makers, on their side, design educational reforms that aim to improve school or college quality by improving students' study incentives. These practices are not strongly supported by the existing empirical evidence. Sigfried and Fels (1979), in an early survey of the literature, conclude that student effort does not seem to be related to performance on standardised tests, but that class attendance does seem to be important for performance on some tests. In a famous contribution, Schuman *et al.* (1985) conclude, in a ten year study where they investigate the relationship between study effort and student performance at the University of Michigan, that there is no positive relation between study time and Grade Point Average (GPA); (see also the follow-up discussions in Hill (1991), Rau and Durand (2000), and Schuman (2001)). Glearson and Walstad (1988), and Krohn and O'Connor (2005) 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)) report positive correlations between attendance and student performance. Betts (1996) finds a positive effect of homework on high school students' maths performance.

The correlation between student performance and student effort could take any sign 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 putting in more effort than more able individuals. These endogeneity problems; which are well recognised by many of the authors cited above, are hard to deal with. In a recent contribution, Stinebrickner and Stinebrickner (2008) exploit random assignments of roommates at

Berea College in the United States. Students are portioned into groups that are identical in all respects except that the students are assigned roommates of a different observable type: some roommates have a videogame, others don't. Stinebrickner and Stinebrickner show that students who are assigned a roommate with a videogame perform poorer than students who are assigned 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 effort are very large (one standard deviation in effort transforms to 0.9 standard deviations in performance) and that the endogeneity problems are quite severe.

In the present paper, we seek to generate credible estimates of effort effects by exploiting 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 endogeneity 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 might magnify the endogeneity problems that are related to dynamic selection.

In our study, the students sort non-randomly into the groups of students that increase or respectively decrease their effort from the first to the second period. We could think of at least two reasons why students make such adjustments. One is that they respond to the information content in the result from the mid-semester test, another that they differ with respect to their time discount rates. To clarify where the dynamic selection effects might originate, we focus here on the latter case (in the empirical analyses we consider both cases). Myopic students most likely provide more effort in the second than in the first period. If the returns to second period effort increase in the level of first period efforts, myopic students will experience small returns to the second period efforts. In comparison, students working hard right from the beginning of the semester might experience much larger returns to their efforts in the second period. An instrumental variable approach might provide a solution to this problem, but has the unattractive feature of providing an average effort effect. An IV- approach then conceals that the returns to own effort differ across student subgroups in potentially important ways. Our strategy is therefore to break the sample into student subgroups that are reasonably homogenous with respect to their time discount rates.

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 on average. The effects are larger than this for students who put in a lot of effort in the first period, and smaller than this (and insignificant) for students who put in relatively more effort in the second period. These results thus differ from the much cited Shuman *et al.* 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. First, we present our data and then offer a few theoretical considerations. The econometric approach is then laid out before the results are presented and conclusions are drawn.

2. The data and some institutional details

The data for this study comes from Trondheim Business School, which is located in the middle of Norway. We highlight one institutional detail that might be important in the given context. The students are graded according to their performance at a final exam. There are no midterm exams, and the performance in class does not count. The results reported in this study are conditional upon the study incentives provided by this system.

The data 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 mandatory mid-semester test (the test was introduced for the students starting in 2006). The students participating in the Macroeconomics course are asked about their efforts twice; the first time in connection with the mid-semester test and the second time at 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, as well as 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 grade point averages (GPA) from upper secondary school do not differ much across students who participated in the survey and students who did not participate: the average GPA for all students is 53.76, while participating students have an average GPA of 53.70. A reasonable hypothesis is that students who participated in the surveys differ from non-participating students with respect to study effort. We return to this issue in the result section.

Effort is about both the quantity and efficiency of time. There is no consensus on the relevant dimensions of time efficiency (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?" and "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: Descriptive Statistics for the Within-Course Analysis

	Mean	Std. Deviation	N
Gender	0.46	0.50	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
Δ ATTENDANCE	-0.11	9.20	153
Δ PERFORMANCE	4.41	5.63	153
Expected performance – actual performance/mid-semester	-2.32	4.64	153

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 upper secondary school, and their gender. Both the mid-semester test and the exam are multiple choice tests with 32 questions. The questions are not identical, but related to the curriculum covered at the time of the test. The average student performance in the mid-semester test and the exam are 16.2 points and 20.6 point respectively. The Business School admits high performing students from 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 achieves between "C" and "D". Thus, the Business School students have higher abilities and most likely, have put in more substantial effort 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 that the students take in the same semester. We use these data in the introduction of the results section, and in the appendix to the paper.

Table 2: Descriptive Statistics for the Across-Course Analysis

	N	Minimum	Maximum	Mean	Std. Deviation
Male	377	0	1	0.47	0.50
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	0.00	12.50	3.27	2.19
STUDY Investment Analysis	290	0.00	20.00	4.96	3.52
STUDY Accounting Information Systems	150	0.00	10.00	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	0.00	100.00	55.89	31.37
Δ STUDY (Macro – Finance)	288	-18.00	4.00	-1.70	3.11
Δ STUDY (Macro-Accounting Information Systems)	148	-9.00	9.00	1.92	2.21
Δ ATTENDANCE (Macro-Finance)	291	-46.67	86.19	8.18	15.99
Δ ATTENDANCE (Macro-Accounting Information Systems)	150	-13.33	88.89	32.54	27.25
Δ GRADES (Macro- Finance)	388	-5.00	4.00	0.63	2.88
Δ GRADES (Macro – Accounting Information Systems)	252	-3.00	4.00	0.34	1.28

3. The allocation of student effort: a few theoretical considerations

Our main purpose is to quantify the effects of study effort 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 Fredrick and Walberg (1980), Becker (1982), Correa and Gruver (1987), and 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 maximising 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.

It follows that the observed effort, as well as the observed returns to studying effort 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 put in more or less effort than low ability students. If effort and ability are complementary inputs; high ability students will experience higher returns to additional effort than low ability students for the same level of effort. Preferences also matter: students with strong performance preferences might put in a great deal of effort and experience decreasing returns to additional effort (these students might have to put in much more additional effort to increase their performance from grade B to grade A).

The model presented above is appropriate for discussing individual student's effort allocations across subjects in a given period of time. To portray the individual student's allocation of effort within a subject across the semester, we introduce the student's subjective time discount rate. A myopic student might be unwilling to allocate much time to study at the beginning of the semester (the first period), but might be willing to increase her/his effort nearer to the exam (the second period). All students, independent of their time discount rates, might experience that the returns to their own effort are larger in the first period compared to the second period if the course material is more difficult towards the end of the course. In addition, myopic students might experience smaller returns in the second period due to the time profile of their effort allocations. This happens if the returns to the effort put in during the second period depend on the level of effort put in during the first period. Formally, this feature might be captured by introducing the effort for two periods - together with an interaction term between the effort levels in the two periods - as arguments in the achievement production function. One potentially important implication for the empirical analysis should be noted. Using within-student difference in effort across two subsequent periods to identify the effects of study effort on performance means that estimates will depend on the fraction of myopic students to all students.

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 realise 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 or period j , $E_{i,j}$ denotes student effort in course or period j , $X_{i,j}$ is a vector of individual and family background characteristics including the student's GPA, and S_j 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 (1986) and Todd and Wolpin (2003)). Here, this specification is chosen because it is a convenient starting point for clarifying the assumptions that are necessary for the differences-in-differences approach. As 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 putting in more effort 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 ϑ is a residual involving the difference between error terms. Assuming that the returns to students 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 causes it to be correlated with individual student's effort. 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. 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 putting in more effort; 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 (which are the same in both periods) 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 generalised.

Thereafter we impose the important DD-assumption that the returns to the effort put in, 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 our favored specification for the fixed-student, fixed-course approach. In this case the variation in ΔE_{ijk} comes from students who put in different amounts of effort in subsequent periods within one course.

One concern with this strategy is that the Macroeconomics course might be more difficult in the second period, implying that the assumption that the returns to the effort put in during the two periods are equal is not fulfilled. To evaluate the restrictiveness of this assumption we estimate:

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

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

Another remaining concern is whether the within-student, within-course, between-periods variation in effort is reasonably exogenous to student performance, that is, the concern is whether the explanatory variable in equation (3) is uncorrelated with the residual. We provide several pieces of empirical evidence to shed light on this issue. First, we investigate whether the change in effort is related to the available observable student characteristics. Second, we seek to evaluate whether the estimates are conditional on the students' subjective time discount rates. There are several potential approaches. Inspired by Schmidt (1983) and Borg, Mason and Shapiro (1989) who emphasise that average effects of study effort might be misleading, we seek to evaluate the importance of the unobserved time discount rates by providing separate effort estimates for subgroups of myopic and non-myopic students. Regrettably, these subgroups cannot be identified with great precision, but we try to approximate the two subgroups in two exploratory exercises. First, we separate the students according to their choice of an additional elective course in the fall semester. At the start of the semester some students choose a "difficult" course, others choose an "easy" course, thus revealing some of their effort preferences. Second, we separate students according to their study time allocations across the two period of the semester, one subgroup consisting of students who increase their efforts in the second relatively to the first period, and one subgroup that decrease their efforts in the second period.

Measurement error

The variable of interest – the students' effort - is 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 point in the study (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 minimise 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

The value added education production function, as portrayed in equation (1), is estimated by OLS. This exercise has two purposes. One is to provide a point of reference for the DD-estimations. The other is to provide some indications as to whether the returns to study effort vary across different subjects. Equation (1) is therefore estimated for the two courses Investment Analysis and Accounting Information Systems in addition to Macroeconomics, and in all cases we use college grades as our outcome variable (because no other information is available for the two former courses). 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.

Serious objections, for instance related to the use of grades that reflect the teachers' grading practices, can be raised against these analyses. Since the main body of the analyses in this paper uses student scores on multiple tests as outcome variables, we do not address these concerns. This exercise is included mainly to draw attention to the fact that the effort estimates are conditional on the subjects investigated. This should be kept in mind when we now turn to Macroeconomics for the DD-analyses.

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

	Macroeconomics	Investment Analysis	Accounting Information Systems
Male	0.212 (0.131)	0.256 (0.189)	-0.050 (0.194)
GPA	0.066 (0.016)	0.076 (0.022)	0.057 (0.026)
Attendance	0.023 (0.006)	0.015 (0.006)	0.010 (0.003)
STUDY	0.048 (0.032)	0.008 (0.017)	0.101 (0.056)
R^2_{adj}	0.131	0.058	0.110
N	264	258	138

Note: Standard errors in parentheses

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

Table 4 reports the results from the DD-analysis for the Macroeconomics course. These are our main results. As already noted we use the difference in the number of correct answers in the two multiple choice 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 effort leads to better test performance. The effect, which is significant at the 1 percent level, is of considerable size: using the results from the most parsimonious specification, we find that one standard deviation (2.45 hours) increase in study effort 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 – indicating that there is no serious sorting on observables into effort changes.

Our OLS estimates reported in Table 3 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 we have been unable to do away with all the endogeneity problems. Below we perform some exercises that might potentially shed light on this issue. We start by examining the likelihood that the assumptions underlying the DD-approach are fulfilled.

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%]
Δ ATTENDANCE	0.081 (0.045)	0.080 (0.045)	0.083 (0.047)	0.080 (0.049)	0.117 (0.057)
Δ STUDY	0.490 (0.184)	0.479 (0.187)	0.447 (0.194)	0.588 (0.251)	0.491 (0.288)
2008-cohort		-0.283 (0.809)	-0.165 (0.842)	0.018 (0.871)	0.247 (1.090)
Gender			0.062 (0.829)	0.006 (0.834)	1.438 (1.075)
GPA			0.116 (0.112)	0.115 (0.112)	-0.095 (0.369)
STUDY				-0.194 (0.224)	-0.132 (0.255)
ATTENDANCE				-0.001 (0.038)	0.043 (0.046)
Constant	4.113 (0.402)	4.262 (0.587)	-2.188 (6.003)	-1.464 (6.898)	4.726 (19.613)
R^2_{adj}	0.048	0.043	0.032	0.025	0.032
N	183	183	176	176	104

The specifications reported in the three first columns of Table 4 apply the assumption that the returns to own study effort 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 of the course compared to 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 effort 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 study effort 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 the changes in effort between periods are not due to 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 upper secondary school is significantly associated with the change in test performance (see Table 4, column (3)). Although the estimated Δ STUDY-coefficient decreases by 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 effort from the first to the second period are associated with observable student characteristics.

More direct evidence is provided from an equation with Δ STUDY 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 effort, which is as expected from the results reported above.

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

	All students
Gender	0.005 (0.335)
GPA	0.013 (0.047)
Expected performance-Actual performance	0.052 (0.036)
Constant	-0.262 (2.562)
R^2_{adj}	0.00
N	152

Nonetheless, it remains that some students for various reasons change their level of effort across the periods. One hypothesis is that this variation reflects students' responses to the information content in the mid-semester test result (see Krohn and O'Connor (2005)). More specifically, students who perform poorer than expected in 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 in the mid-semester test, we use this answer to generate an additional explanatory variable, which is the difference between expected and actual performance in the mid-semester exam. This new explanatory variable is positively correlated with Δ STUDY, implying that students who perform poorer than expected at the first test increase their efforts in the next period. This finding is consistent with the findings reported by Krohn and O'Connor (2005). The estimated coefficient is rather small and not significantly different from zero, probably reflecting that the reported expectation is a noisy variable. It is not clear what we econometrically should make of this finding.

A different hypothesis is that the most ambitious/less myopic students work hard right from the beginning of the semester, and thus, do not change their effort much from the first to the second period, while the less ambitious students increase their effort nearer to the exam. The estimated

average effect conceals that the effort returns differ across these student subgroups in potentially important ways. To deal with this worry, we first break the sample into two; one subsample consisting of students who increase their effort in the second period, and one subsample consisting of students who decrease their effort in the second period. The estimated effort coefficient for students who decrease their effort in the second period is significant at the 1 percent level and equal to 0.776. The effort estimate for the other subsample is 0.350 and statistically insignificant.

The exercise above can be criticised for separating the students according to an endogenous variable. A more attractive approach is to use the revealed effort preferences that originate from the students' choices of elective courses. At the start of the fall semester, the students choose between Accounting Information Systems and Organizational Psychology, of which the former has the reputation of being the more challenging course. Consistent with this, the students also report much higher efforts for the former. When estimating the returns to study effort in Macroeconomics separately for the subgroups of students that are enrolled in the Accounting Information Systems and the Organizational Psychology courses we find estimates of 0.538 (0.200) and 0.276 (0.510) respectively. The findings from these two exercises (which neither is reported in the tables) are consistent with the hypotheses that students who work relatively hard from the beginning of the semester or are characterised by strong work ethics experience higher returns to their effort than students who put in less effort right from the beginning of the semester. Schmidt (1983) finds that studying for the final examination may have a negative marginal product for weak students. This finding is more extreme than those reported above, but has the same flavour to it.

Finally, there is a sample selection issue to be analysed. A little less than 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 in November when the last round of survey 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; (see Becker and Powers (2001)). To evaluate this hypothesis we have exploited that the first survey was "almost mandatory": it was provided together with the mandatory mid-semester test, implying that almost all students have reported the effort they put in during the first part of the semester. Table 6 reports summary statistics separately for students who participated and students who did not participate in the second survey. While the two subgroups of students have approximately the same average GPA, the participating students perform better than the non-participating students in the mid-semester test, and consistent with this, they put in more effort prior to the test. Since non-participating students put in relatively little effort prior to the mid-semester test, it seems likely that their returns to additional effort in the second period are below the average for the participating group of students, implying that our average STUDY-estimate is somewhat biased upwards.

Table 6: Participants and non-participants in the second survey

	All	Participants	Non-participants
GPA	53.76	53.70	53.88
Study	3.12	3.28	2.86
Test result	15.17	16.06	13.85
Expected test result	13.42	13.97	12.49
Number of students	307	195	112

Finally, more insights into the importance of student effort can potentially be gained by exploiting the variations in effort allocations across different subjects/courses. We have performed these exercises, but, it turns out that this approach does not fulfill the identifying assumptions to any reasonable degree. Since it might be useful to see where the problems originate, we present the results from one of these exercises in the Appendix.

Comparison with the Stinebrickner and Stinebrickner study

Our most credible estimates - one standard deviation in study effort transforms into 0.25 standard deviations in performance in Macroeconomics - is a little less than $\frac{1}{3}$ 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 effort 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 upper secondary school, it is evident that the returns to studying effort 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 effort, 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). Third, our results indicate that it matters how the effort is allocated across the semester. Students putting in relatively more effort from the beginning of the semester seem to experience higher returns than students who put in relatively little effort from the beginning of the semester. A larger fraction of the students in the Stinebrickner and Stinebrickner study might have put in more effort early in the semester. This is not an unlikely conjecture. As noted in section 2, the grading in Trondheim Business School is entirely based on performance in the final exam, indicating that the incentives for putting in study effort right from the start of the semester are weaker than at Berea. 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 academic performance of (higher) education institutions. Incentive-based reforms with the aim of increasing students' effort 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 the kind of incentives that are likely to be the most effective. At a more fundamental level, no incentive-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. The existing empirical evidence provides no consistent evidence.

In this paper we seek to fill some of the gap. Using data for two subsequent cohorts of students in a 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 Macroeconomics course. 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 effort will improve their performance.

Can we believe the results? The present study solves the most obvious endogeneity problems - related to unobservable student characteristics - by utilising within-student variations in effort allocations within the same subject for different time periods. There is a remaining problem related to

unobservable characteristics of the students, notably their subjective time discount rates. Some exploratory analyses indicate that myopic students experience less returns to their efforts than students with lower time discount rates. The estimated average effect of study effort thus is conditional on the share of myopic students in the investigated sample. More research is needed to settle this issue.

References

- Akerlof, G. A. and Kranton, R. E. (2002). 'Identity and schooling: some lessons for the economics of education', *Journal of Economic Literature*, Vol. 40, pp. 1167–1201.
- Allison, P. D. (1990). 'Change scores as dependent variables in regression analysis', *Sociological Methodology*, Vol. 20, pp. 93–114.
- Becker, W.E and Powers, J. (2001). 'Student performance, attrition, and class size given missing student data', *Economics of Education Review*, Vol. 20, pp. 377–388.
- Becker, W. E. (1982). 'The educational process and student achievement given uncertainty in measurement', *American Economic Review*, Vol. 72(1), pp. 229–236.
- Betts, J. (1996). 'The role of homework in improving school quality', Discussion Paper 96–16, University of California at San Diego.
- Borg, M. O., Mason, P. M. and Shapiro, S. L. (1989). 'The case of effort variables in student performance', *Journal of Economic Education* Vol. 20(3), pp. 308–314.
- Correa, H., and Gruver, G. W. (1987). 'Teacher-student interaction: a game-theoretic approach to the economic theory of education', *Mathematical Social Science*, Vol. 13, pp. 19–47.
- Costrell, R. M. (1995). 'A simple model of educational standards', *American Economic Review*, Vol. 84(4), pp. 956–971 .
- Fredrick, W. C. and Walberg, H. J. (1980). 'Learning as a function of time', *Journal of Educational Research*, Vol. 73, pp. 183–194.
- Gleason, J. P. and Walstad, W. B. (1988). 'An empirical test of the inventory model of student study time', *Journal of Economic Education*, (Fall), pp. 315–321.
- Hanushek, E. A. (1986). 'The economics of schooling: production and efficiency in public schools', *Journal of Economic Literature*, Vol. 24, pp. 1141–1177.
- Hill, L. Jr. (1991). 'Effort and reward in college: a replication of some puzzling findings', in *Replication Research in the Social Science*, by J. W. Neuliep (ed.), Newbury Park, CA: Sage, pp. 139–156.
- Krohn, G. A. and O'Connor, C. M. (2005). 'Student effort and performance over the semester', *Journal of Economic Education*, Vol. 36(1), pp. 3–28.
- Marburger, D. R. (2006). 'Does mandatory attendance improve student performance', *Journal of Economic Education*, Vol. 37(2), pp. 148–155.
- Park, K. H. and Kerr, P. M. (1990). 'Absenteeism and undergraduate exam performance', *Journal of Economic Education*, (Spring), pp. 101–111.

- Rau, W. and Durand, A. (2000). 'The academic ethic and college grades: does hard work help students to "make the grade"?' , *Sociology of Education*, Vol. 73, pp. 19–38.
- Romer, D. (1993). 'Do students go to class? Should they?' , *Journal of Economic Perspectives*, Vol. 7, pp. 167–174.
- Schmidt, R. M. (1983). 'Who maximizes what? A study in student time allocation' , *American Economic Review* Vol. 73(2), pp. 23–28.
- Schuman, H., Walsh, E., Olson, C. and Etheridge, B. (1985). 'Effort and reward: the assumption that college grades are affected by the quantity of study' , *Social Forces*, Vol. 63, pp. 945–966.
- Schuman, H. (2001). 'Comment: Students' efforts and reward in college settings' , *Sociology of Education*, Vol. 74(1), pp. 73–74.
- Siegfried, J. and Fels, R. (1979). 'Research and teaching of college economics: A survey' , *Journal of Economic Literature*, Vol. 17(3), pp. 923–969.
- Stinebrickner, R. and Stinebrickner, T. (2003). 'Working during school and academic performance' , *Journal of Labor Economics*, Vol. 21(2), pp. 473–491.
- Stinebrickner, T. and Stinebrickner, R. (2004). 'Time-use and college outcomes' , *Journal of Econometrics*, Vol. 121 (1–2), pp. 243–269.
- Stinebrickner, T. and Stinebrickner, R. (2008). 'The causal effect of studying on academic performance' , *The B.E. Journal of Economic Analysis & Policy*, Vol. 8(1), Article 14.
- Todd, P. E. and Wolpin, K.I. (2003). 'On the specification and estimation of the production function for cognitive achievement' , *Economic Journal*, Vol. 113: F3–33.

Appendix 1: Exploiting within-student, across-subjects variation

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

	[20%, 80%]	[20%, 50%]	[50%, 80%]
Δ ATTENDANCE	0.002 (0.005)	-0.008 (0.009)	0.004 (0.006)
Δ STUDY	0.188 (0.085)	0.338 (0.131)	0.028 (0.132)
ATTENDANCE	-0.016 (0.011)	-0.013 (0.017)	-0.020 (0.016)
STUDY	-0.068 (0.085)	-0.166 (0.164)	0.019 (0.104)
R^2_{adj}	0.042	0.009	0.000
N	83	47	36

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

More insight into the importance of student effort can potentially be gained by exploiting the variations in effort allocations across different subjects/courses. Here we shortly present the results from such an exercise. 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 range 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 effort 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 effort put into the course is regressed against GPA and gender. The results from these exercises are reported in Table A2.

Table A2: OLS estimations of efforts against student background characteristics for all courses

	Macroeconomics		Investment Analysis		Accounting Information Systems	
	Attendance	STUDY	Attendance	STUDY	Attendance	STUDY
Male	-0.904 (1.51)	-0.099 (0.267)	2.28 (1.95)	-0.019 (0.435)	-8.68 (5.26)	0.088 (0.290)
GPA	0.04 (0.18)	0.042 (0.031)	-0.03 (0.23)	-0.014 (0.054)	-0.71 (0.71)	0.062 (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

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 A3 reports the results from estimating equation (3) with Macroeconomics and Accounting Information Systems as the two subjects. By controlling for cohorts (column (2)), the Δ STUDY estimate is 0.068, and not significant at conventional levels. One standard deviation in the Δ STUDY variable transforms into 0.1 standard deviations in Δ GRADE. This effect is between $\frac{1}{2}$ and $\frac{1}{3}$ of the effect estimated from within-course variation in effort.

Table A3: Differences-in-differences estimation using Macroeconomics and Accounting Information Systems

	All students	All students	All students	All students	Students in [20%, 80%]
Δ ATTENDANCE	0.005 (0.004)	0.005 (0.004)	0.005 (0.004)	0.004 (0.004)	0.003 (0.005)
Δ STUDY	0.078 (0.045)	0.067 (0.045)	0.126 (0.064)	0.126 (0.066)	0.159 (0.080)
2008 cohort		0.142 (0.346)	0.439 (0.343)	0.325 (0.392)	-0.251 (0.496)
2007 cohort		0.431 (0.342)	0.122 (0.349)	0.059 (0.384)	0.036 (0.516)
ATTENDANCE macro			-0.002 (0.008)	-0.003 (0.008)	-0.014 (0.011)
STUDY macro			-0.077 (0.063)	-0.084 (0.065)	-0.036 (0.081)
Gender				0.205 (0.212)	-0.144 (0.260)
GPA				-0.001 (0.031)	-0.046 (0.072)
R^2_{adj}	0.021	0.026	0.024	0.008	0.024
N	143	143	143	143	83

Note: Standard errors in parentheses

The results from estimating the less restrictive equation (4) are reported in column (3). The Δ STUDY 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 effort 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.

Contact details

Hans Bonesrønning
Department of Economics,
NTNU, 7491 Trondheim,
Norway
Email: hansbo@svt.ntnu.no

Leiv Opstad
Trondheim Business School,
7004 Trondheim,
Norway
Email: leiv.opstad@hist.no