NBER WORKING PAPER SERIES

THE CAUSAL EFFECT OF STUDYING ON ACADEMIC PERFORMANCE

Todd R. Stinebrickner
Ralph Stinebrickner

Working Paper 13341
http://www.nber.org/papers/w13341

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2007

The work was made possible by generous funding from The Mellon Foundation, The Spencer Foundation, The National Science Foundation, The Social Science Humanities Research Council and support from Berea College. We are very thankful to Anne Kee, Lori Scafidi, Dianne Stinebrickner, Pam Thomas, and Albert Conley who have played invaluable roles in the collection and organization of the data from the Berea Panel Study. The authors would like to thank Dan Black, John Bound, Brian Jacob, Lance Lochner, Jeff Smith, and seminar participants at Northwestern, Maryland, Syracuse, The University of British Columbia, and NBER. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

The Causal Effect of Studying on Academic Performance
Todd R. Stinebrickner and Ralph Stinebrickner
NBER Working Paper No. 13341
August 2007
JEL No. I2,J22,J24

ABSTRACT

Despite the large amount of attention that has been paid recently to understanding the determinants of educational outcomes, knowledge of the causal effect of the most fundamental input in the education production function - students' study time and effort - has remained virtually non-existent. In this paper, we examine the causal effect of studying on grade performance using an Instrumental Variable estimator. Our approach takes advantage of a unique natural experiment and is possible because we have collected unique longitudinal data that provides detailed information about all aspects of this experiment. Important for understanding the potential impact of a wide array of education policies, the results suggest that human capital accumulation is far from predetermined at the time of college entrance.

Todd R. Stinebrickner
Department of Economics
University of Western Ontario
London, Ontario, N6A 5C2
CANADA
and NBER
trstineb@uwo.ca

Ralph Stinebrickner
Dept. of Mathematics
Berea College
Berea, KY 40404
ralph_stinebrickner@berea.edu

Section 1. Introduction

Understanding the impact of most potential education policy changes is made difficult by the reality that the large majority of variation in student outcomes is unexplained by traditionally observable individual and school characteristics. Thus, it is important that, while substantial recent attention has been paid to understanding the determinants of educational outcomes, knowledge of the causal impact of the most fundamental input in the education production function - students' own study time and effort - has remained essentially non-existent.[1]

One primary reason for the current void in our understanding is that standard data sources have not traditionally collected information about how much time students spend studying. The very small amount of existing work that has provided direct evidence about the relationship between studying and academic performance has focused on collecting measures of study-effort and has obtained estimates of the (conditional) correlation between the number of hours that a person studies and his/her academic performance. In the first of this work, Schuman et al. (1985), over the course of a ten year period, took four different measurement approaches in an explicit attempt to "produce a positive relation between amount of study and GPA" at the University of Michigan and found that none of the approaches was "very successful in yielding the hypothesized substantial association." Similar replication results at different schools by Hill (1991) and Rau and Durand (2000) produced generally similar results. [2]

The bias associated with viewing the descriptive relationships in previous work as estimates of the causal role that studying plays in the grade production process arises, in part, because students who spend more

[1]While in some cases research has uncovered evidence that schools, parents, class sizes, vouchers, and competition are related to educational outcomes, much remains unknown about why some students have better outcomes than others. Student effort is potentially important both for explaining some of the large amount of variation that remains and for thinking about why some education policies are found to have significant effects while others are not found to have significant effects.

[2]Within the economics literature, the only work that examines this relationship is Stinebrickner and Stinebrickner (2004) who estimated the descriptive relationship between a student's first semester grade performance and his/her average daily study hours using the same data as in this paper. Betts (1997) finds that the amount of homework assigned by teachers between grades seven and eleven has a quantitatively important relationship with student achievement as measured by test scores. A number of authors, including Ehrenberg and Sherman (1987), Ruhm (1997) and Stinebrickner and Stinebrickner (2003), have studied the relationship between employment during school and academic performance.

time studying may be different in unobserved ways related to, say, ability than those who spend less time studying. However, further confounding the endogeneity problem is the possibility that individuals who receive bad grade shocks or have difficult classes during a particular semester may react by changing their effort during that semester. Not only is it not possible to know the size of the bias that is present if one views the correlations found in previous papers as estimates of the causal effect, but it is also not possible to know the direction of the bias. Thus, given the central policy importance of effort and the reality that no previous work has addressed the endogeneity problem that may very well be present, it should perhaps be disconcerting that a recent review of the current evidence led Schuman (2000) to write that "for now, we can conclude that the amount of studying has some but not a great deal to do with students' achievement as measured by grades, especially GPA."

Ideal for learning about the importance of studying would be a random experiment in which two groups of students that are identical in all respects at the beginning of school are forced to study different amounts during school, but continue to behave identically in all other ways (class attendance, sleeping, drinking, study efficiency, paid employment etc.) that could influence the outcome of interest. In this paper we examine the effect of studying on college grade performance by using an Instrumental Variable (IV) approach that takes advantage of a real-world situation which we find closely resembles this ideal experiment.

The analysis in this paper is possible because we designed a sequence of surveys with the specific goal of documenting all aspects of this natural experiment and personally administered these surveys to a sample of students over the course of their freshman year in college. The survey data play a crucial role in all aspects of our work. First, specific questions in the data allow us to construct the instrument that we use to divide students into two groups that are identical at the time of college entrance: students who have a randomly assigned roommate who brought a video game to school at the beginning of the year and students who have a randomly assigned roommate who did not bring a video game to school at the beginning of the year.[3] Second, time-use diaries that were collected at multiple times during the year allow us to document that

[3]To be more precise, the two groups are identical in the sense that they are drawn from the same population distribution of student characteristics.

the assignment of a roommate with a video game causes students in the former group to study significantly less per day, on average, than students in the latter group. Finally, because we designed our own longitudinal survey with a well-defined issue in mind, we are able to directly examine the possible theoretical reasons that our instrumental variable might not be valid even in the presence of random assignment. Specifically, information from the time diaries and additional survey questions allow us to obtain information about all of the college behaviors other than study-effort (class attendance, sleeping, drinking, study efficiency, paid employment etc.) that we could imagine influencing grade performance directly. We find no evidence that being assigned a roommate with a video game influences these other behaviors.

Thus, the evidence in the survey data, when combined with the random assignment feature, suggests that it is reasonable to believe that the two groups of students are very similar in all dimensions other than study-effort that influence grade performance. In this case, we can learn about the causal effect of studying by comparing average grade outcomes between the two groups. Linking our survey data to administrative data, we find that grades are significantly lower, on average, for the group that studies less, on average, and we estimate that studying has an important effect on grade performance.

While our estimate of the effect of studying on academic performance is statistically significant at approximately .05 when we instrument using only the video game variable, our sample is relatively small and the estimator is not particularly precise. To address this issue we take advantage of the fact that, for the large majority of students in our sample, we have access to two other potential instruments: how much a student's randomly assigned roommate studied in high school and how much this roommate expects (at the time of college entrance) to study in college. The motivation for exploring the usefulness of these potential instruments comes from Stinebrickner and Stinebrickner (2006) who found that roommates interact very little on specific academic matters and that peer effects between roommates are most likely to arise through students influencing the time-use of each other. We find that these instruments are strong predictors of study-effort and find no evidence that these instruments influence other behaviors that could influence grades directly. Adding these instruments to our IV specification increases the precision of our estimator considerably. It is also worth noting that this paper can make a substantial contribution even without pinning down the size of the effect exactly; while previous work has found with certainty that the effect of studying is small, even when viewed cautiously our results indicate that the effect of studying is likely to very important.

It is worth stressing that this paper presents perhaps the only opportunity in the foreseeable future to learn about the causal role that effort plays in the production of human capital. One obvious reason for this is that standard data sources do not contain information about how much students study. However, perhaps more importantly, our results strongly suggest that it is not possible to provide convincing evidence about the causal effect of studying unless one has dealt with the potential endogeneity issue in a credible fashion, and, as a result, provide a likely explanation for the lack of a positive finding in the previous work described earlier. Specifically, we find that our IV estimate is much larger than the Ordinary Least Squares (OLS) estimate. We find no evidence that study-effort varies with our observable measure of ability - a college entrance exam score. However, we design a test which takes advantage of two semesters of data and (under an assumption that a transitory component of the grade process is independent across semesters) shows that the difference between the IV and OLS estimates can be entirely explained by a "dynamic selection" effect in which students increase effort when bad luck or other negative grade shocks occur. Thus, not only does this test provide some compelling evidence for the difference we find between the IV and OLS estimates, but it also provides a cautionary alarm about the use of certain types of estimators (e.g., fixed effects) that might be tempting to employ in the absence of the type of experiment utilized in this paper but are not necessarily appealing on theoretical grounds.

While the majority of this paper involves establishing an IV estimate for the effect of an additional hour of studying on academic performance, it is important to stress that the reduced form relationship between whether a student's randomly assigned roommate brings a video game at the beginning of the year (our instrument) and the student's academic performance is informative in and of itself. This is the case because this relationship immediately establishes that non-drastic policies can have substantial effects on grade performance, and, even without requiring that one fully establish the exogeneity condition that is necessary for the IV estimator to be valid, provides evidence that the amount and/or quality of a person's studying has an important causal effect on college grade performance.

In Section 2 we describe the survey project which takes place at Berea College. In Section 3 we describe the equation of interest and provide OLS results. Section 4 contains the IV results for the specification which includes only the video game instrument and also contains causal evidence from the reduced form specification discussed in the previous paragraph. In Section 5, we explore the usefulness of the other instruments which characterize the past effort and expected effort of a student's roommate and describe the results when these instruments are added to our IV specification. In Section 6 we examine the reasons for the difference between the OLS and IV estimates. In Section 7 we discuss the importance of this work for policymakers, including the fact that it provides perhaps the first direct evidence about an underlying avenue through which peer effects operate.

Section 2. A general overview of the Berea Panel Study

Located in central Kentucky where the "bluegrass meets the foothills of the Appalachian mountains," Berea College is a liberal arts college which operates under a mission of providing educational opportunities to students of "great promise but limited economic resources." The survey data used in this paper are part of the Berea Panel Study (BPS) that Todd Stinebrickner and Ralph Stinebrickner (hereafter referred to as S&S) started with the explicit objective of collecting the type of detailed information that is necessary to provide a comprehensive view of the decision-making process of students from low income families. The BPS involved surveying two cohorts of students approximately twelve times each year while they were in school with baseline surveys being administered to the students in the first BPS cohort prior to their freshman year in the fall of 2000 and to students in the second BPS cohort prior to their freshman year in the fall of 2001.[4]

Of direct relevance for the analysis in this paper, a sequence of time-use surveys were administered

[4]In addition to collecting detailed background information about students and their families, the baseline surveys were designed to take advantage of recent advances in survey methodology in order to collect information about students' preferences and expectations towards uncertain future events and outcomes (e.g., academic performance, labor market outcomes, non-pecuniary benefits of school, marriage and children) that could influence decisions. Substantial follow-up surveys that are administered at the beginning and end of each subsequent semester have been designed to document the experiences of students and provide information about how various factors that might influence decisions change over time.

at multiple times during each academic year. Also of relevance, the baseline and follow-up surveys collected substantial information about friends, roommates, and other information related to studying and grade performance. Student identifiers allow the survey data to be merged with Berea College's administrative data.

Section 3. The equation of interest and OLS results

Our equation of interest is

(1) $GPA_i = \alpha_0 STUDY^*_i + \alpha_1 X_i + u_i$.

The dependent variable is the first semester grade point average (GPA) of student i in his/her freshman year. $STUDY^*_i$ is the average number of hours that a person studies per day over all of the days in the first semester. X_i contains a constant, a MALE indicator variable, an indicator of whether the student is BLACK, an indicator of whether a student's health is excellent at the time of entrance (HEALTH_EXC), an indicator of whether a student's health is poor or fair at the time of entrance (HEALTH_BAD), a student's score on the American College Test (ACT), and a set of seven college major indicators $MAJOR_1,...,MAJOR_7$ where $MAJOR_i$ is equal to one if the student believes at the time of entrance that he is more likely to end up with $MAJOR_i$ than any other major.[5] u_i represents unobserved individual determinants of the grade performance of person i. It contains, for example, information about other behaviors such as class attendance that influence grade performance, unobserved measures of ability, the difficulty of a student's classes, and whether the person has good or bad "luck" in a particular semester. With respect to the latter ("luck") we have in mind, for example, whether a student gets sick at an inopportune time during the semester or finds that he/she has a bad match with his/her professors in the semester.

Two problems are potentially present in the estimation of equation (1). First, while our data are unique in that they contain detailed information about student study-effort, an errors-in-variables problem is present because $STUDY^*_i$ is not fully observed in the data. What is observed is $STUDY_i$, a noisy proxy for $STUDY^*_i$ which is created by averaging the number of hours that a person studies per day over the subset of

[5] To be more precise, the MAJOR variables represent groups of majors that are described in Table 1.

days during the semester that his/her study-effort is observed. During the first semester, daily study-effort was collected on four different weekdays using the twenty-four hour time diaries that are shown at the end of Appendix A. Response rates were relatively high on these surveys; the median person in our sample described below answered all four surveys and the average number of responses was 3.11. Second, STUDY* is potentially correlated with the unobservable u because decisions about how much to study in a particular semester may depend on, for example, a student's unobserved ability or may depend on the difficulty of a student's classes or information that the student receives about his/her luck in that semester.

The presence of these errors-in variables and endogeneity problems imply that the Ordinary Least Squares estimator of equation (1), obtained by replacing $STUDY^*_i$ with $STUDY_i$, may be biased with the direction of the bias unknown. The OLS estimates are shown in the first column of Table 4. The estimated effect of studying is small, with an extra one hour of daily study-time increasing first semester GPA by only .038, and not statistically significant at significance levels less than .13. Thus, our OLS results are similar in spirit to the previous literature that was discussed in the introduction.

Section 4. Results using the video game instrument

Section 4.A. Intuition underlying identification strategy

Instrumental variable estimation represents a desirable way to deal with the two issues above. In this section, we describe the results obtained when we instrument for STUDY* in equation (1) with a variable, which we refer to generically as TREATMENT, that indicates whether a student's randomly assigned freshman roommate brought any type of video game with him/her at the beginning of the school year. The intuition behind the IV approach in this section is as follows. In Section 4.B we use the TREATMENT variable to divide our sample into two groups - those who have randomly assigned roommates who brought video games and those who have randomly assigned roommates who did not bring video games. In Section 4.C we show that the presence of a video game causes students in the former group to study less, on average, than students in the latter group. In Section 4.D we use the random assignment of roommates along with additional, unique information from the BPS to argue that it is very plausible to believe that students in the two groups are very similar in all other (non-study) dimensions that influence grade performance. The IV estimator in Section 4.E is based on the fact that, if this is the case, then differences in average grade performance between the groups can be attributed to differences in average study-effort between the groups.

Section 4.B. Dividing the sample using the TREATMENT variable

The survey question which asked whether a student's roommate brought a video game(s) to school appeared for the first time in our surveys in the fall of 2001. As a result, we focus on the BPS cohort that entered Berea as freshmen in 2001. As mentioned earlier and discussed later, the validity of our instrument takes advantage of the fact that students at Berea who do not request roommates are unconditionally randomly assigned roommates.[6] Slightly more than one-third of students at Berea either live off campus or request a roommate. The sample used in this paper contains information about 210 students who live on campus and were randomly assigned roommates. The TREATMENT entry near the end of Table 1 shows that 53% of males and 24% of females in our sample have roommates that brought some sort of video game(s) to school.

It is worth noting that our sample size is small given the decrease in precision (relative to OLS) that can be expected to accompany the IV estimator. As a concession to the small sample size, we combine males and females when we apply the IV estimator. We present information in the following sections that this is reasonable.

Section 4.C. Does the Instrument Influence Study Decisions?

The descriptive statistics in the first row of Table 1 show that, for both males and females, study-effort differs in a quantitatively important manner between students in the sample whose roommates bring video games to school and students in the sample whose roommates do not bring video games to school. Specifically, the sample average of STUDY is .667 lower (2.924 vs. 3.591) for males who receive the video

[6]Unlike students at most schools, freshmen at Berea are not asked to complete a housing preference questionnaire. Approximately two weeks before the start of school (and after all members of the freshman class are determined) pairs of roommates were drawn in a purely random fashion (for this cohort using a random number generator on the campus administrative computing system) from the pool of all freshmen who need roommates. S&S (2006) provide a set of empirical checks which find no evidence of a relationship between a student's observable characteristics and those of his/her roommate.

game treatment than for males who do not receive the treatment. The sample average of STUDY is .467 lower (3.226 vs. 3.693) for females who receive the video game treatment than for females who do not receive the treatment. It is not possible to reject the null hypothesis that the effect of the treatment is the same for males as it is for females.

Pooling the male and female observations we estimate a first stage regression of the form

(2) $STUDY_i = \beta_0 TREATMENT_i + \beta_1 X_i + \nu_i$

and show the results in the first column of Table 2. As expected given the random assignment of the treatment, for both males and females the sample means of the variables in X are very similar for students who receive the treatment and those that do not receive the treatment.[7] Thus, the sample means for the males and females provide rough guidance about the estimate of β_0. We find an estimate (std. error) of −.668 (.252) which indicates that the treatment reduces study time by two-thirds of an hour per day. Given that students in the sample study 3.48 hours per day on average, the estimated effect is quantitatively important, and a test of the null hypothesis that the treatment has no effect on study-effort is rejected at all levels of significance greater than .01.

Section 4. D. Does the video game instrument satisfy the exogeneity requirement?

In order for the instrument to be valid, it must be the case that its only influence on a student's grade performance comes through its effect on the student's study-effort. There are two avenues through which this exogeneity requirement could be violated. First, it would be violated if the treatment contains information about a student's unobserved characteristics at the time of college entrance. Second, it would be violated if, in addition to affecting decisions about study-time, the treatment also affects other behaviors that take place during the first semester and influence grade performance. Roommates who bring video games to school may

[7] For example, the null hypothesis that ACT is the same in the population for males (females) who receive treatment and males (females) who do not receive treatment cannot be rejected for any significance levels less than .46 (.37). The null hypotheses that the proportion of students that are BLACK is the same in the population for males (females) that receive treatment and males (females) that do not receive treatment cannot be rejected for any significance levels less than .25 (.37). Similar findings exist for the major and health variables. The proportion of males in the population who receive the treatment is not expected to be the same as the proportion of females in the population who receive the treatment because males and females are not assigned to the same rooms.

be different in observable and unobservable ways than those who do not. As a result, in thinking about these two avenues through which the exogeneity condition could be violated, it is necessary to take into account that the treatment involves both the physical presence of the video game(s) and the presence of whatever type of roommate accompanies the game(s). However, it is important to note at this point that, while it is perhaps tempting a priori to view students who bring video games as types who will tend to encourage a variety of harmful behaviors in their peers, this does not seem to be the case. Specifically, as detailed in the remainder of the paper, we find no evidence that students at Berea who bring video games are of lower observed ability, are less likely to attend class, are more likely to drink alcohol, or have harmful sleep habits.

The first avenue: student characteristics at the time of college entrance

The random assignment of roommates in our sample plays the key role in ensuring that the exogeneity condition is not violated by the first avenue described in the previous paragraph. If students were choosing roommates, they would also (perhaps quite indirectly) be choosing whether roommates bring video games. In this case, the amount that a student intends to study and other factors such as the student's ability could be related to whether his roommate brings a video game. The random assignment of roommates guarantees that, conditional on a student's sex, students in the sample who receive the treatment come from the same population distribution as students in the sample who do not receive the treatment.

It is worth noting that this conclusion assumes that a student's decision about whether to bring a video game is not influenced by whether his randomly assigned roommate is bringing a game. With respect to this assumption, even if some amount of coordination did exist, our estimator would presumably be either unbiased or biased downwards (and, given our results in Section 4.E, still informative) under the assumption that students who bring video games have unobserved characteristics that are similar or less favorable than those of students who do not bring video games.[8] However, it does not seem that we need to rely on such an

[8]Suppose some coordination exists. Then 1). Some students who brought videogames themselves have roommates who decided not to bring video games because of coordination. 2). These students were, in effect, incorrectly put in the "roommate didn't bring video game" group when they should have been put in the "roommate did bring videogame" group. 3). These students tend to be weaker students (under the assumption that students who

argument since the empirical evidence suggests that coordination is not common. Some evidence of this comes from the fact that we do not reject the null hypothesis that there is no relationship between whether a student brings a video game and whether his/her roommate brings a video game. This is what one would expect if, unlike appliances such as refrigerators, bringing a videogame is relatively costless in terms of space and students have a connection to their own video game (on which, e.g., they have built specific human capital). Regardless, in order to provide stronger, direct evidence about this issue we conducted a survey of a more recent cohort of new students at Berea. The response rate on our survey was 85% with 345 out of 405 new students participating. Of these 345 students, 229 (66%) were randomly assigned roommates. We find that 55% of the randomly assigned students had no interaction of any type with their roommates before arriving at Berea. Of the students that did interact before arrival, 7.9% answered both that they had not brought a video game and that the "decision of whether or not to bring a video game was influenced by communication with the roommate before arrival at Berea." Thus, only eight of 229 (3.5%) of students with randomly assigned roommates indicate that their decision to not bring a video game was influenced by interaction with a roommate.[9]

The second avenue: student behaviors during college other than study-effort

With respect to whether the exogeneity condition could be violated through the second avenue described above, there seem to be two general possibilities. One possibility is that, in addition to reducing the

---

bring video games tend to be weaker) because they brought video games themselves. 4). Then, moving these weaker students from the "roommate didn't bring video game" group to the "roommate did bring video game" group would increase the difference in average grades between the groups. As a result, our estimator is conservative under the assumptions.

[9]The conclusion that students in the sample who receive the treatment come from the same population distribution as students in the sample who do not receive the treatment also assumes that, if misreporting of whether a person's roommate brought a video game exists, this misreporting is not systematically related to unobserved factors such as a person's unobserved ability. There does not seem to be any obvious reason that this would be problematic. For example, at least in terms of observable characteristics, we do not find that students with high ability spend more time studying (in which case they might spend less time in their room and perhaps have less opportunity to realize that their roommate has a video game). However, more importantly, as described in a footnote in Section 4.E, our results change very little if, at the cost of creating a smaller sample size, we construct our instrument using a roommate's own report of whether he/she brought a video game. If the instrument is constructed in this way, any possible concerns about reporting error are no longer relevant.

amount of time spent studying, students who receive the treatment also reduce time spent in other activities that influence grade performance directly. Seemingly most important among these other activities is class attendance which is unique in that it directly influences the amount of course material to which a person is exposed. However, also potentially important are other activities that influence how rested or clear-thinking a person is at the time he/she is studying or attending class. The activities that seem most likely to fit this description are sleeping, drinking/partying, and paid employment. In the following paragraphs we examine whether differences in class attendance, sleeping, drinking/partying, and paid employment exist between the treated and untreated groups.[10]

With respect to class attendance, our knowledge of institutional details at Berea suggests that the treatment would have little effect at Berea. Unlike many other schools, class attendance is to a large degree mandatory at Berea and this expectation is made very clear to students. Many faculty members impose strict attendance policies and faculty typically either formally or informally keep track of attendance of individual students. Thus, to a large extent, the decision of whether or not to attend class is not even in the choice set of students at Berea, and, as a result, we expected a prior that attendance would be very high for both students who receive the treatment and those who do not. We can check this empirically. At four times during the first semester, we used Question A in Appendix A to elicit information about the number of times in the previous seven days that a student's classes were scheduled to meet and the number of these classes that the student attended. For each student we compute the proportion of classes that he/she attended across all time-use surveys that he/she completed. In column 1 of Table 3a we regress this proportion, PATTEND, on TREATMENT and SEX. The estimated effect (std. error) of TREATMENT is -.014 (.009). Thus, the estimated effect is not significant at .10 and is quantitatively very small; treated students in the sample have attendance rates that are lower by only 1.4 percentage points or just slightly more than 1.4 percent lower given

[10]We do not find it easy to imagine other activities that might be influenced by the presence of a video game and would be expected to have a non-trivial direct effect on grade performance. One possibility is time spent exercising if this activity has a positive (or negative) effect on a person's ability to focus in classes or studying. We could examine this activity using our time-diaries but have not done so at this time.

an overall average attendance rate of approximately .96.[11] We can also provide information about whether the treatment affects class attendance by using information from our time diaries. For each student we construct a CLASSHOURS variable in a manner that is analogous to how the STUDY variable is calculated - by averaging the number of daily hours a person reports being in class over all of the time-use diaries. The regression of CLASSHOURS on TREATMENT and SEX in column 2 of Table 3a indicates that students spend approximately three and one-half hours per day in class and that the estimated effect of the treatment on class attendance is quantitatively small and statistically insignificant.[12]

With respect to the number of hours of sleep, we did not have a strong prior about what to expect. Using our time diaries we construct the variable SLEEP in a way that is directly analogous to the way that the variable STUDY is constructed. The third column of Table 3a shows the results from a regression of SLEEP on TREATMENT and MALE. The estimated effect (std. error) of TREATMENT is .275 (.208). Thus, the effect is not statistically significant and indicates that students in the sample who receive the treatment sleep approximately fifteen minutes more per night than students in the sample who do not receive the treatment. We also use our time-diaries to construct a variable BEDTIME that indicates the time at which a student goes to bed. This variable is created such that positive values indicate the number of hours after midnight and negative values indicate the number of hours before midnight. Column 4 of Table 3a shows a regression of BEDTIME on TREATMENT and MALE. We find that, on average, students go to bed between 12:45 and 1:00, and we find no evidence that the treatment influences BEDTIME.

With respect to drinking/partying, we knew from the many years that we have spent around Berea, that

[11]As mentioned earlier, we also find no difference in class attendance between students who bring video games and those who do not. For example, when we reestimate column 1 of Table 3a after replacing TREATMENT with whether a person brought a videogame himself/herself, we find an estimate (standard error) of -.012 (009).

[12]It seems reasonable to assume that the treated and non-treated students have similar numbers of classes and this assumption is supported by evidence from the first part of Question A in Appendix A. On average, students who receive the treatment report that their classes were scheduled to meet 14.40 hours in the previous seven days. On average, students who do not receive the treatment report that their classes were scheduled to meet 14.10 hours in the previous seven days. A test that the number of scheduled classes is the same in the population for treated and non-treated students cannot be rejected at significance levels less than .44.

the prevalence of drinking is very low relative to other schools.[13] Contributing to this reality is the fact that Berea is a Christian (non-denominational) school and many students come from religious backgrounds in which drinking is not accepted. In addition, the area around Berea is a "dry" area in which alcohol sales are prohibited. Nonetheless, it is worth directly examining this issue. This is possible because our time diaries contain a category "partying." Column 4 of Table 3b shows a regression of the number of hours spent partying on MALE and the TREATMENT. On average, students spend only about ten minutes a day partying, and we find no evidence of a relationship between the number of hours spent partying and whether a person receives the treatment. Approximately 85% of all students do not report any partying on any of the time-use surveys and this percentage also does not vary in a meaningful way with whether a person's roommate brought a video game. While we were certainly not surprised by the low prevalence of weekday drinking, it is at least possible that some students are wary of reporting this information on their time diaries. Nonetheless, our intuition is that, if substantial differences in drinking behavior exist between the treated and non-treated students, these differences would reveal themselves in, for example, the variable BEDTIME. Further, there is no strong reason to believe, a priori, that students who bring video games to school are more likely to drink and there is no evidence in the time diaries that this is the case.[14]

Finally, with respect to paid employment, the institutional details of the school imply that there cannot be substantial differences between treated and untreated students. This is the case because the school has a mandatory work-study program in which all students work ten hours per week in on-campus jobs during the first semester and students are not allowed to hold off-campus jobs.

These results suggest that, while the treatment leads to substantial decreases in study-effort, it has very little effect on other time-use activities that might influence grade outcomes. There is an additional survey question that can help support this conclusion. At the end of the first semester, we asked each student how

---

[13]One co-author has been a faculty member at Berea for over thirty years. The other co-author had many friends who attended Berea, and, as a result, gained a direct knowledge of the social aspects of Berea.

[14]Including a variable which indicates whether a person brought a video game is found to have no effect in column 4 of Table 3b. The proportion of people who bring video games who report drinking on at least one time-use survey, .854, is virtually identical to the proportion for students who do not bring video games, .851.

much time he/she spent playing video games in an average week during the semester. On average, students in the treatment group reported playing 4.06 hours a week and non-treated students reported playing only .79 hours per week. Given that the treatment reduces study time by a little more than one-half of an hour per day, these numbers are remarkably consistent with the notion that the treatment is having little effect on other activities.[15] In addition, this information provides direct evidence that study time is lower for the treatment group because students are playing games. A test that there is no difference in game playing between students who receive the treatment and students who do not receive the treatment yields a t−statistic of 3.54 and is rejected at all traditional significance levels.

The other way that the exogeneity condition could be violated through the second avenue is if, in addition to reducing the amount that a student studies, the treatment also causes a student to study less efficiently. This possibility could be of relevance if the presence of a video game in a room implies that the student may not be able to study in the room when he/she wants to because, for example, the room has become a place where others congregate. We can examine this possibility using question B in Appendix A which asked students about the physical locations where they studied. We find no difference in study locations for those who received the treatment and those who did not. In column 1 of Table 3b we regress the percentage of study time that takes place in the dorm room on TREATMENT and MALE. The estimated effect of TREATMENT is not statistically significant.

A related way that studying might be less efficient for the treated students would be if the video game or the television that often accompanies the video game serves as a distraction while the student is studying - perhaps because the roommate is watching television or playing the game. We can examine this to some extent because question B in Appendix A elicits information about how much time is spent studying with the television on. We do not find any evidence that this is the case in column 2 of Table 3b where we regress the

[15]It seems likely that the fact that increases in video playing seem to come primarily at the expense of studying has something to do with the fact that students at Berea have more required activities than students at other schools. As an example, all students work approximately two hours per day in a mandatory work-study program. As another example, attendance at a series of convocations during the semester is also required.

percentage of time spent studying with the television on TREATMENT and MALE.[16]

It is hard to know for sure whether a person would answer that he/she was "studying with TV on" if his/her roommate was playing a video game on the television. Nonetheless, there is a very natural bound on how much of a student's study time could occur while a video game is being played by his/her roommate. Using the question described above which asked each student how much time he/she spent playing video games in an average week during the fall semester, we find that roommates who bring video games spend 36 minutes per day, on average, playing the video game. Thus, even if we make the extreme assumption that a treated student is studying in the room at all of the times that his/her roommate is playing the video game, only approximately 20% of a student's study time would, on average, take place with the video game on.[17] The fact that this is certainly an extremely conservative bound, when combined with the evidence in column 2 of Table 3b, suggests to us that it is unlikely that treated students are suffering substantially because their studying is taking place while their roommates are playing video games or watching television.[18]

The possibility that students who receive the treatment are studying less efficiently could also be of relevance if treated students have roommates who are less able or less willing to help them directly with their coursework. However, S&S (2006) discuss in depth the avenues through which roommates could transmit peer effects and using unique data on the amount and nature of interactions between roommates conclude that, in the short-run, peer effects are much more likely to be transmitted by good role models influencing the time-use

[16]Similarly, since some video games are played on computers, treated students may be more likely to have a computer in their room and this could represent an academic advantage for treated students. In column 3 of Table 3b we regress the number of hours per week that a student uses a computer for academic reasons on TREATMENT and MALE. Students in the sample whose roommates bring video games report that they use the computer for academic reasons about one extra hour per week than non-treated students in the sample, but the estimated effect of TREATMENT is not statistically significant. Further, even if this possible academic advantage was present for treated students, it would produce a downward bias in our estimator (and, as a result, our findings in Section 4.E, would continue to be informative).

[17]Treated students study approximately three hours per day.

[18]Suppose that the times during the day at which a student studies in the room (1.8 hours per day, on average) are chosen randomly from the available non-sleep hours of the student and that the times during the day at which a student's roommate plays the video game (approximately 36 minutes per day, on average) are chosen randomly from the available non-sleep hours of the roommate. Then only approximately 2% of a treated student's overall study time would take place while his/her roommate is playing a video game. This percentage would be understated to some extent if there are some hours during the day when, for example, both students have classes. However, it would be overstated to some extent if students tend to be somewhat hesitant to play a distracting video game if their roommate is studying and/or if students look for other places to study if a roommate is playing a distracting video game.

decisions of their roommates than by high ability students helping their roommates understand their coursework.[19] Further, in our data we find no relationship between the TREATMENT variable and the amount of time a student spends interacting with his roommate on academic matters, and, at least in terms of college entrance exam scores, we find no evidence that treated students have lower ability roommates than non-treated students.[20] In short, it seems highly unlikely that grade differences between treated and non-treated students are being driven in a non-trivial manner by differences in help with coursework from roommates.

While it is never possible to empirically establish with full certainty that an instrument satisfies the condition of being exogenous, the random assignment of roommates ensures that students in treated and untreated groups are identical in the population at the time of entrance and the unique features of our survey collection efforts allow us to credibly examine the remaining reasons that this condition might be violated. Thus, it seems very plausible to believe that the instrument satisfies the exogeneity condition, and we assume that this is the case in the remainder of the paper. However, it is worth noting that, as discussed in Section 4.F, the reduced form relationship between the instrument and grade performance is able to provide important causal evidence related to grade performance even without requiring that the exogeneity assumption described above be fully satisfied.

Section 4.E. Instrumental variable estimates

As described earlier, the intuition about how the IV estimator achieves identification is straightforward with the binary instrument. The assumption that the instrument is valid implies that, conditional on sex, all factors other than study-effort that influence grade performance are identical for treated and non-treated students in the population. Thus, if studying has no effect on grade performance, grade performance would be identical (conditional on sex) for the treated and untreated groups even though study-effort is different

---

[19]There are many reasons for this conclusion. One issue is that it may be quite costly for students to help each other given that they may not be taking the same classes with the same faculty members (and are often not close friends). We find empirical evidence that, while roommates often spend considerable amounts of time together, they spend little of this time "studying or discussing course material."

[20]When we estimate a linear regression of a person's ACT score on whether a person brought a video game to school on ACT and MALE the estimated effect (std. error) on ACT is .526 (.534). Thus, holding sex constant, students in the sample who bring video games have average ACT scores that are one-half of a point higher than students who do not bring video games.

between the groups. As can be seen in the second row of Table 1, males in the sample who receive the treatment have grades that are .239 lower than males who do not receive the treatment and females in the sample who receive the treatment have grades that are .128 lower than females who do not receive the treatment. The size of the IV estimate takes into account the differences in average study-effort that led to these differences in average grades. So, for example, given that the treatment reduces study-effort by .667 of an hour for males, a Wald estimate of the effect of studying on GPA obtained from the sample of males would be .239/.667=.358. Similarly, a Wald estimate of the effect of studying on GPA obtained from the sample of females would be .128/.467=.274.

Formal IV estimates are shown in column 2 of Table 4. As noted earlier, our small sample makes it difficult to estimate the model separately for males and females. However, the earlier evidence that it is not possible to reject the null hypothesis that the treatment has the same effect on the study-effort of males and females along with the evidence in the previous paragraph that Wald estimates are similar for males and females suggests that pooling males and females is reasonable. The IV estimate indicates that an additional hour of studying per day causes first semester grade point average to increase by .360. Thus, the IV estimate is much larger than the OLS estimate in column 1 of Table 4.

Although, as expected, the effect is estimated with much less precision under IV than under OLS, a test of the null hypothesis that studying has no effect on grade performance produces a t-statistic of 1.963 and the test is rejected at significance levels greater than .051.[21] It is important to keep in mind that the existence of a large amount of sampling variation implies that non-trivial uncertainty exists about the size of the population parameter. Nonetheless, our paper can make a significant contribution even without being able to pin down the exact size of the true effect. Roughly speaking, while previous work that could not deal with the endogeneity problem found that with certainty that the effect of effort is small, our results indicate that the effect of effort is likely to be very important. Further, this message becomes even stronger in Section 5 where

[21]For a subset of 173 observations we observe a roommate's own report of whether he/she brought a video game. Constructing the instrument using the roommate's own report, our estimate for this subset is slightly higher, .402, although, in part because of the smaller sample size, the estimator is less precise and the t-statistic is somewhat lower, 1.8.

we take advantage of additional instruments which substantially increase the precision of our estimator.

To provide additional support that our results are not being driven by differences (between the treated and untreated groups) in behaviors other than study-effort, we also estimated a specification which added as regressors all of the dependent variables in Table 3a and Table 3b. In the interest of space considerations, full results are not shown, but the estimated effect (std. error) in this specification was .377 (.198). We also found that the results changed very little when we added an explanatory variable which indicates whether the student himself/herself brought a video game.[22] While random assignment implies that both specifications with and without "own" analogs to the instruments are valid on theoretical grounds, here and in Section 5 we choose to present full results from the specifications without the own values simply because the effect of interest is more precisely estimated in these specifications (although the point estimates are larger both here and in Section 5 when own values are included).

Section 4.F. Causal evidence from the reduced form

The reduced form impact of the TREATMENT in Table 5 shows that, conditional on the other covariates, students in the treatment group receive grades that are .241 lower than students in the untreated group, and a test of the null hypothesis that the treatment has no effect on grades is rejected at significance levels greater than .01. At its most general level, this is important because it indicates that even non-drastic policy changes have the ability to influence grade performance in a non-trivial fashion. In addition, suppose that one believes that the class-attendance, sleeping, and drinking information in the time diaries provides convincing evidence that the treated and non-treated students are equally likely to attend class and are equally rested and clear thinking during class, but is worried that treated students study less efficiently than non-treated students in some unobserved way. In this case, the .241 difference in grades between the treated and non-treated groups would most reasonably be attributed to a combination of the fact that treated students study

[22] In this case, the estimate (std. error) is .363 (.195). In the first stage analog to column 1of Table 2, we find that students who bring video games themselves study .418 less hours per day than students who do not and that this effect is significant at .10. We note that it is not clear on theoretical grounds whether the own effect should be larger or smaller than the effect of the roommate bringing a video game. Students who bring video games may be students who have found they are most able to handle the temptation the games may represent. Perhaps more importantly, video games may be not be dissimilar to other toys in the sense that usage might be particularly intense in the period after first exposure and might decline after that as the initial novelty wears off.

approximately 40 minutes less per day than non-treated students (conditional on observable characteristics) and the possibility that treated students are studying less efficiently (in some unobserved way not captured by our survey). Thus, in this case, the reduced form estimate provides strong evidence that the amount that a person studies and the manner in which the person studies plays an important role in academic performance, even without requiring the exogeneity condition (needed for IV) to be fully satisfied.

Section 5. Instrumental Variable results taking advantage of additional instruments

In this section we examine whether we can increase the precision of our estimator by taking advantage of information about two other potential instruments - how much a randomly assigned roommate reported studying in high school (RSTUDYHS) and how much a randomly assigned roommate expects to study in college (REXSTUDY) - that was collected at the time of college entrance. This information is available for the 176 individuals in our initial sample whose roommates also chose to participate in our survey and provided legitimate information about these variables. The potential promise of these instruments can be seen in S&S (2006) who concluded that peer effects between first semester roommates are most likely to arise through students influencing the time-use of each other. In the first stage regression in column 2 of Table 2 we find direct evidence that a student's time-use can be influenced by his roommate's time-use behavior; RSTUDYHS is statistically significant at significance levels greater than .032 when included in a specification that also includes the TREATMENT instrument and REXSTUDY.

From an exogeneity standpoint, both the appeal and possible concerns about these instruments are essentially identical to those discussed earlier for the video game instrument. With respect to the former, the combination of the random assignment feature and the fact that the instruments characterize aspects of study-effort of the roommate at the time of college entrance implies that students with different values of RSTUDYHS and REXSTUDY are drawn fro the same population at the time of college entrance. With respect to the latter, the instruments would be problematic if, in addition to influencing a student's study-effort, RSTUDYHS and REXSTUDY also influence other behavior that is related to grade performance. As in Section 4.D, we treat this latter concern as an open empirical question which we can examine directly because we designed our survey with the objective of eliciting information about the set of all college behaviors that could influence grades directly. As described in detail in Appendix B, we find no evidence that RSTUDYHS and REXSTUDY have an effect on these other behaviors.

Thus, as with the video game instrument, it seems plausible to believe that RSTUDYHS and REXSTUDY are valid instruments. In column 3 of Table 4 we find that adding these instruments to the specification in column 2 leads to a substantial increase in precision; the standard error decreases by 33% from .183 to .121. The point estimate decreases somewhat to .291 and we now reject the null hypothesis that studying has no effect at all significance levels greater than .017. We found that the results changed very little when we added explanatory variables which indicate whether the student himself/herself brought a video game, how much the student himself/herself studied in high school, and how much the student expected (at the time of entrance) to study in college.[23] Thus, these results strengthen the conclusion that effort likely plays an important role in the grade production function.

Section 6. Understanding the difference between the IV and OLS estimates

In this section we attempt to understand why the IV estimates in Table 4 are much larger than the OLS estimate. We focus on the difference between the OLS estimate for the full sample, .038, and the IV estimate for the full sample, .360, which appear in the first two columns of Table 4. Part of the difference between these estimates, .322, arises because of the errors-in-variables problem from using STUDY instead of STUDY* in equation (1). As discussed in S&S (2004), the OLS estimator would need to be multiplied by a factor of

$$(3)\quad \frac{\mathrm{Var(STUDY)}}{\mathrm{Var(STUDY)}-\dfrac{\sigma_v^2}{N}}.$$

to correct for this problem, where σ_v^2 is the variance of the unobservable in equation (2) and N is the number

[23]In this case, in results not shown in tables, the estimate (std. error) is .342 (.161). In the first stage analog to column 2 of Table 2 (results also not shown), we find an own effect (std. deviation) of how much a student studied in high school (the own analog to RSTUDYHS) of .029 (.013). The own effect of how much a student expects to study in college is insignificant at traditional levels when included with the high school effort level.

of time-use surveys. It is difficult in our case to know exactly what the bias factor is since N is not constant across people. However, using equation (3) we ascertain that the bias factor is between 1.40 and 1.94.[24] Thus, the difference between the IV and OLS estimates that remains after accounting for the errors-in-variables problem is between .286 and .307.

The direction of the bias due to the endogeneity problem is uncertain from a theoretical standpoint. Sufficient for this is that students with high unobserved ability may study more or less than students with low unobserved ability.[25] However, the fact that the IV estimate is much larger than the OLS estimate suggests that there exists a negative correlation between $STUDY^*_i$ and u_i. One possibility is that students that study more tend to be of lower permanent, unobserved ability than other students. However, while the potential importance of unobserved ability makes it difficult to provide conclusive evidence about this possibility, one gets a sense that this might not be the driving influence from examining the results in the first column of Table 2 which reveal no evidence of a relationship between our observable measure of ability (ACT) and study-effort.

This suggests that the difference between the IV and OLS estimates might arise because students adjust their effort in a particular semester in response to the transitory portion of grades in that semester. The presence of a second semester of grade and study-effort information presents us with an opportunity to independently examine whether there is evidence in the data that students study more when the transitory portion of grades is low. For the time being we think about this transitory portion of grades as "luck" which we imagine captures things like the match quality of a student and his professors during a particular semester and whether the student gets sick at an inopportune time during the semester. We design a test that takes

[24] An estimate of σ_v^2 can be constructed by differencing the individual daily study reports for a particular person. Estimates of VAR(STUDY) can be computed conditional on N from the sample. 1.40 is an estimate of the factor by which the OLS estimator would be biased if all students answered four time-use surveys. 1.94 is an estimate of the factor by which the OLS estimator would be biased if all students answered only one time-use survey.

[25] On one hand, high ability students may enjoy studying more than other students. On the other hand, given that high ability students may achieve the maximum grade in a class at lower amounts of studying, an additional hour of studying may lead to higher grade and future benefits for the lower ability student(s), and, in addition, low ability students may be forced to study more just to "stay afloat."

advantage of the fact that, while study-effort in the first semester may be correlated with the transitory component of grades in the first semester, it should be uncorrelated with the transitory component of grades in the second semester under the assumption that the transitory portion of grades is uncorrelated across time. This implies that the grade difference between the second and first semesters, averaged over all people who studied a particular amount in the first semester, will be larger if this group experienced bad luck on average in the first semester.

To be more specific about this test, it is worthwhile to disaggregate the unobservable in equation (1) into a person-specific, permanent component μ_i and a transitory component ε_{ti} that is assumed to be serially uncorrelated

(4) $u_{ti} = \mu_i + \varepsilon_{ti}$.

Equation (1) now represents a model in which grades are generated by a study component, $\alpha STUDY^*_i$, a permanent ability component, $\beta X_i + \mu_i$, and a transitory or luck component, ε_{it}. At this point we rename variables slightly to differentiate between the first and second semesters. The grade equation for semesters one and two are given by equations (5) and (6) respectively

(5) $GPA_{1i} = \alpha_0 STUDY_1{}^*{}_i + \alpha_1 X_i + \mu_i + \varepsilon_{1i}$

(6) $GPA_{2i} = \alpha_0 STUDY_2{}^*{}_i + \alpha_1 X_i + \mu_i + \varepsilon_{2i}$

Differencing equation (6) from equation (5) and rearranging yields

(7) $GPA_{1i} - GPA_{2i} - \alpha_0(STUDY_1{}^*{}_i - STUDY_2{}^*{}_i) = \varepsilon_{1i} - \varepsilon_{2i}$.

Thus, the left hand side of equation (7) represents the difference in a person's transitory component or "luck" between the two semesters. For illustrative purposes, consider a case where there are only two study levels in the population: $STUDY_1^*$= high or $STUDY_1^*$=low. Averaging the left hand side of equation (7) over all individuals who have $STUDY_1^*$= high yields $E(\varepsilon_1|STUDY_1^* = high)$ since the assumption that the transitory components are uncorrelated implies that $E(\varepsilon_2 |STUDY_1^*=high)=0$. Similarly, averaging the left hand side of equation (7) over all individuals who have $STUDY_1^*$= low yields $E(\varepsilon_1|STUDY_1^* = low)$. Comparing $E(\varepsilon_1|STUDY_1^* = high)$ to $E(\varepsilon_1|STUDY_1^* = low)$ indicates how the transitory component of grades varies, on average, across the two $STUDY_1^*$ amounts.

This discussion motivates our estimation of an equation of the form

(8) $GPA_{1i} - GPA_{2i} - .360\,(STUDY_{1i} - STUDY_{2i})=\text{constant} + \delta\, STUDY_{1i} + \eta_i$.

We find an OLS estimate (std. error) for δ of -.276 (.040). This implies that students who study an extra hour per day have an average realization of the transitory component ε_1 that is .276 lower than otherwise similar students. Identification for the OLS estimator involves comparing the GPA of students who study an extra hour to the GPA of students who do not study an extra hour. Earlier we found that a difference of between .286 and .307 remains between the IV and OLS estimates remains after accounting for the errors-in-variables problem. The results here indicate that, under the assumption that the transitory component of grades is uncorrelated across semesters, this remaining difference can be attributed to the finding that the average GPA of students who study an extra hour per day would be .276 lower than the average GPA of students who do not study the extra hour under the counterfactual in which both groups study the same amount.

Of course, it is not the case that all variation in the transitory components should necessarily be interpreted literally as "luck." For example, while students at Berea have rather limited flexibility about the classes they take during the first year due to a large number of required "general studies" courses, it is possible that some of the changes in the transitory component across semesters could reflect differences in the difficulty of classes across semesters. To the extent that this is the case, the assumption that the transitory component of grades is uncorrelated across semesters may lose some of its attractiveness. Nonetheless, at the very least, this exercise sounds a cautionary alarm about the use of fixed effects estimators. In this application, a fixed effects estimator would achieve identification using the within person variation in study-effort across the two semesters. However, our results indicate that assuming that this variation is exogenous is most likely problematic. In addition, the evidence that ACT scores are unrelated to study-effort suggests that the variation in study-effort across people, which is discarded by the fixed effect estimator, may be less likely to suffer from problems of endogeneity. As a result, not only is the use of a Fixed Effects estimator unlikely to satisfactorily deal with the endogeneity problems, but the Fixed Effects estimator may perform worse than the OLS estimator. Striking evidence that this is the case is shown in column 4 of Table 4. The estimated effect of studying, -.043, is negative, and a test of the null hypothesis that studying has harmful effect on grades cannot be rejected at levels of significance greater than .10. Thus, the paper suggests that significant caution should be taken when considering the use of Fixed Effects estimator in cases where behavioral responses to information may be present.

Section 7. Conclusion

To the best of our knowledge, this work represents the only evidence about the causal relationship between study-effort and grade production. Many policy decisions depend on the extent to which college outcomes of interest are driven by decisions that take place after students arrive at college rather than by background factors that influence students before they arrive at college.[26] Thus, it is important that both the IV and reduced form estimates suggest that human capital accumulation may be far from predetermined at the time of college entrance. For example, using results from our full sample, an increase in study-effort of one hour per day (an increase of approximately .67 of a standard deviation in our sample) is estimated to have the same effect on grades as a 5.21 point increase in ACT scores (an increase of 1.40 standard deviations in our sample and 1.10 standard deviations among all ACT test takers). In addition, the reduced form effect of being assigned a roommate with a video game is estimated to have the same effect on grades as a 3.88 point decrease in ACT scores (an increase of 1.04 of a standard deviation in our sample and .82 standard deviations among all ACT test takers).

While not the primary focus of this paper, this paper also makes an important contribution to the peer effects literature in general and to the peer effects literature that achieves identification by using college roommates in particular. The goal of the empirical peer effects literature has been to look for empirical evidence which documents that peer effects can matter. This paper provides depth to that literature by not only providing strong evidence that peer effects can matter, but also by providing perhaps the first direct evidence about an avenue (time-use) through which peer effects operate. This paper also makes a contribution to a

[26]Examples include: a) decisions about how to distribute education dollars across student ages; b) decisions about appropriate strategies for counseling students who perform poorly; c) deciding what types of students should be admitted to college (highly motivated or high ability) and its direct importance to merit vs. need based admission decisions.

substantial literature outside of economics by establishing that video games can have a large, causal effect on academic outcomes.

Unlike the OLS results from this work and the results from a small amount of earlier work that did not address the endogeneity problem, our IV results indicate that the effect of studying may be very substantial. Certainly more work in this area is warranted and our findings strongly suggest that other surveys that focus on educational issues should seriously consider collecting information about this very fundamental input in the human capital production process.

References

Altonji, Joseph, Huang, Ching-I, and Taber, Christopher (2004): "Estimating the Cream Skimming Effects of Private School Vouchers on Public School Students," Working paper.

Altonji, Joseph, Elder, T. and Taber Christopher (2005): "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," Journal of Political Economy, 2005.

Betts, 1997, The role of homework in improving school quality, discussion paper 96-16, University of California at San Diego.

Black, Dan and Smith, Jeffrey (2004a): "How Robust is the Evidence on the Effects of College Quality? Evidence form Matching," Journal of Econometrics, August 2004, 121(1-2) 99-124.

Black, Dan and Smith, Jeffrey (2004b): "Estimating the Returns to College Quality with Multiple Proxies for Quality, working paper.

Black, Dan and Jeffrey Smith (2005): "Estimating the Returns to College Quality with Multiple Proxies for Quality," working paper.

Cameron, S and Taber, Christopher (2004): "Borrowing Constraints and the Returns to Schooling, Journal of Political Economy," 112..

Card, David. 1999. "The Causal Effect of Education on Earnings." In Handbook of Labor Economics, volume 3A. Orley Ashenfelter and David Card, eds. New York: Elsevier Science.

Cullen, Julie, Jacob, Brian, and Levitt, Steven (2003): The Effect of School Choice on Student Outcomes: Evidence from Randomized Lotteries, NBER working paper, 10113.

Ehrenberg, Ronald and Sherman, Daniel. "Employment While in College, Academic Achievement and Postcollege Outcomes," Journal of Human Resources, Winter 1987, 22(1), pp. 1-24.

Hill, Lester, Jr. (1991): "Effort and Reward in College: A Replication of Some Puzzling Findings." In Replication Research in the Social Science, edited by James W. Neuliep. Newbury Park, CA: Sage, pp. 139-56.

Heckman, J, Lochner, L. and Taber, C (1998): "Tax Policy and Human Capital Formation," American Economic Review, Papers and Proceedings, May, 1998.

Rau, William, and Durand, Ann (2000). "The Academic Ethic and College Grades: Does Hard Work Help

Students to ‘Make the Grade’?” Sociology of Education, *73:19-38.*

Rivkin, S., Hanushek, E. and Kain, J. (2005): “Teachers, School, and Academic Achievement,” Econometrica, 73(2), 417-458.

Ruhm, Chris. “Is High School Employment Consumption or Investment,” Journal of Labor Economics, October 1997, 15(4), 735-776.

Sacerdote, Bruce (2001): “Peer Effects with Random Assignment: Results for Dartmouth Roommates,” Quarterly Journal of Economics, volume 116.

Schuman, Howard, Edward Walsh, Camille Olson, and Barbara Etheridge. (1985): “Effort and Reward: The Assumption that College Grades are Affected by the Quantity of Study.” Social Forces,” 63:945-66.

Schuman, Howard (2000): “Comment: Students’ Efforts and Reward in College Settings,” Sociology of Education, Vol 74, No. 1 (2001), 73-74.

Stinebrickner, Ralph and Todd Stinebrickner, 2003, Working during school and academic performance, Journal of Labor Economics 21(2) (April), 473-491.

Stinebrickner, Todd R. and Ralph Stinebrickner (2004). “Time-Use and College Outcomes,” *Journal of Econometrics*, 121(1-2), 243-269.

Stinebrickner, Todd R. and Ralph Stinebrickner (2006): “What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds,” *Journal of Public Economics,* 90 (2006), 1435-1454.

Zimmerman, David (2003): “Peer Effects in Academic Outcomes: Evidence from a Natural Experiment,” *The Review of Economics and Statistics,* 85(1), 9-23

Table 1
Descriptive Statistics

	All n=210	**Male All** n=95	**Male treatment =0** n=45	**Male treatment =1** n=50	**Female All** n=115	**Female treatment =0** n=88	**Female treatment =1** n=2
STUDY	3.427 (1.631)	3.240 (1.688)	3.591 (1.748)	2.924 (1.583)	3.583 (1.573)	3.693 (1.595)	3.226 (1.473)
GPA - First semester Grade Point Avg	3.004 (.652)	2.853 (.677)	2.979 (.663)	2.740 (.677)	3.129 (.605)	3.159 (.598)	3.031 (.628)
ACT	23.380 (3.709)	22.463 (3.842)	22.155 (3.931)	22.740 (3.779)	24.139 (3.431)	24.205 (3.527)	23.925 (3.149)
BLACK	.171	.189	.200	.180	.157	.159	.148
MAJOR1 - Agriculture	.076	.115	.111	.120	.043	.045	.037
MAJOR2- Business	.176	.168	.133	.200	.182	.204	.111
MAJOR3- Elem. Education	.10	.084	.111	.06	.113	.137	.044
MAJOR4- Humanities	.223	.157	.133	.18	.278	.261	.333
MAJOR5- Science & Math	.209	.252	.222	.28	.173	.156	.235
MAJOR6 - Professional	.119	.094	.133	.06	.139	.147	.111
MAJOR7 - Social Sciences	.071	.084	.088.	.08	.060	.056	.074
Omitted Major Physical Educ.	.024	.042	.066	.02	.008	0.0	.037
HEALTH_BAD fair/poor health	.067	.052	.066	.04	.078	.057	.148
HEALTH_EXC excellent health	.371	.40	.333	.46	.347.	.363	.296

Table 1
Continued

	All	**Male All**	**Male treatment =0**	**Male treatment =1**	**Female All**	**Female treatment =0**	**Female treatment =1**
	n=210	**n=95**	**n=45**	**n=50**	**n=115**	**n=88**	**n=2**
Instrument Section 4 & 5							
TREATMENT - Roommate brought a video game to school	.367	.526			.235		
Instruments Section 5 only n=176							
RSTUDYHS - roommate's hours of study per week in high school	10.278 (10.11)						
REXSTUDY - roommate's expected hours of study per day during college	3.464 (1.826)						

Table 2
First Stage Regressions
The effect of instruments (and other variables) on study hours

Independent Variable	**estimate (std error) n=210**	**estimate (std error) n=176**
INSTRUMENTS		
video game TREATMENT	-.668 (.252)**	-.658 (.268)**
RSTUDYHS		.028 (.013)**
REXSTUDY		.049 (.074)
OTHER VARIABLES		
MALE	-.155 (.244)	-.204 (.263)
BLACK	.417 (.341)	.549 (.350)
ACT	-.019 (.036)	-.016 (.038)
$MAJOR_1$	1.423 (.828)*	1.230 (.816)
$MAJOR_2$	1.421 (.783)*	1.015 (.772)
$MAJOR_3$	1.120 (.811)	.891 (.789)
$MAJOR_4$	1.637 (.784)**	1.410 (.782)*
$MAJOR_5$	1.575 (.776)**	1.375 (.762)*
$MAJOR_6$	1.777 (.806)**	1.604 (.797)**
$MAJOR_7$	2.128 (.836)**	2.006 (.827)**
HEALTH_BAD	.209 (.463)	.221 (.478)
HEALTH_EXC	.095 (.241)	.010 (.258)
	R^2=.092	R^2=.179

Note: The first column uses the entire sample of individuals with randomly assigned roommates. The second column which takes advantage of roommates' reports of how many hours they studied per week in high school (RSTUDYHS) and how many hours they expect to study per day in college (REXSTUDY) uses the subset of these students whose roommates are also members of the sample and are not missing values of RSTUDYHS and REXSTUDY.
*significant at .10
**significant at .05

Table 3a
The effect of video game TREATMENT on other behaviors, n=210

Independent Variable	Dependent Variable PATTEND proportion of classes attended	Dependent Variable CLASSHOURS daily hours in class	Dependent Variable SLEEP daily sleep hours	Dependent Variable BEDTIME time student went to sleep#
	estimate (std. error)	estimate (std. error)	estimate (std. error)	estimate (std. error)
TREATMENT	-.014 (.009)	-.114 (.188)	.275 (.208)	.143 (.199)
MALE	.003 (.009)	.059 (.182)	.209 (.202)	-.276 (.192)
CONSTANT	.962 (.006) **	3.444 (.25)**	7.089 (.138)**	.833 (.130)**
	R^2=.012	R^2=.0016	R^2=.019	R^2=.011

*significant at .10
**significant at .05
#dependent variable is created so that it is zero at 12:00 midnight. Positive numbers represent hours after midnight. Negative numbers represent hours before midnight.

Table 3b
The effect of TREATMENT on additional behaviors, n=210

Independent Variable	Dependent Variable percentage of study time that takes place in dorm room	Dependent Variable percentage of study time that takes place in dorm room with tv on	Dependent Variable hours per week using computer for academic purposes	Dependent Variable daily hours partying
	estimate (std. error)	estimate (std. error)	estimate (std.error)	estimate (std. error)
TREATMENT	-2.111 (4.670)	3.515 (2.933)	.963 (1.069)	.007 (.050)
MALE	-4.677 (4.498)	-3.812 (2.825)	-.254 (1.032)	-.015 (.048)
CONSTANT	61.456 (3.058)**	12.756 (1.921)**	6.820 (.699)**	.125 (.033)**
	R^2=.008	R^2=.008	R^2=.012	R^2=0.011

*significant at .10
**significant at .05

Table 4
Estimates of the effect of studying on grade performance:
Ordinary Least Squares, Instrumental Variables, Fixed Effects

Independent Variable	**OLS** **n=210** **estimate (std. error)**	**IV** **instrument: video game TREATMENT** **n=210** **estimate (std. error)**	**IV** **instruments: video game TREATMENT, RSTUDYHS, REXSTUDY** **n=176** **estimate (std. error)**	**Fixed Effects** **n=210** **estimate (std. error)**
CONSTANT	.719 (.408)*	-.073 (.709)	-.062 (.638)	-.050 (.047)
STUDY	.038 (.025)	.360 (.183)**	.291 (.121)**	-.043 (.027)*
SEX	-.132 (.084)	-.023 (.129)	-.010 (.126)	
BLACK	-.220 (.122)*	-.356 (.183)*	-.334 (.176)*	
ACT	.062 (.013)**	.069 (.018)**	.072 (.018)**	
$MAJOR_1$	.834 (.298)**	.393 (.474)	.576 (.410)	
$MAJOR_2$	.793 (.282)**	.356 (.454)	.475 (.380)	
$MAJOR_3$	.725 (.292)**	.335 (.452)	.467 (.389)	
$MAJOR_4$	.796 (.283)**	.298 (.474)	.411 (.403)	
$MAJOR_5$	.643(.280)**	.174 (.462)	.366 (.389)	
$MAJOR_6$	.664(.292)**	.091 (.510)	.143 (.427)	
$MAJOR_7$	.901 (.304)**	.235 (.555)	.243 (.468)	
HEALTH_BAD	.019(.166)	-.029 (.226)	-.020 (.219)	
HEALTH_EXC	.127 (.086)	.115 (.117)	.158 (.118)	
	R^2=.273			

Note: The first, second, and fourth columns use the entire sample of individuals with randomly assigned roommates. The third, which takes advantage of roommates' reports of how many hours they studied per week in high school (RSTUDYHS) and how many hours they expect to study per day in college (REXSTUDY) uses the subset of these students whose roommates are also members of the sample and are not missing values of RSTUDYHS and REXSTUDY.
*significant at .10
**significant at .05

Table 5
Causal impact in reduced form: The direct effect of treatment on first semester grades

Independent Variable	**Dependent Variable** **GPA** **first semester grades** **estimate (std error)**
CONSTANT	.793 (.398)**
TREATMENT	-.241 (.089)**
MALE	-.079 (.086)
BLACK	-.209 (.120)*
ACT	.062 (.012)**
$MAJOR_1$	.906 (.293)**
$MAJOR_2$	.868 (.277)**
$MAJOR_3$	.739 (.287)**
$MAJOR_4$	.889 (.277)**
$MAJOR_5$	.741 (.274)**
$MAJOR_6$	.731 (.285)**
$MAJOR_7$	1.002 (.295)**
HEALTH_BAD	.045 (.164)
HEALTH_EXC	.149 (.085)*
	R^2=.289

*significant at .10
**significant at .05

Appendix A: Survey questions

Survey Question A.

In the last 7 days (one week), how many times were your classes scheduled to meet?_____
Please count up carefully the number of scheduled class meeting for each one of the seven days and add them together. (If your schedule for a particular day included one math class meeting, one GST class, a biology lab, and a music class you would count 4 for that day. Add together these numbers for each day to get a total for the week.

How many of these classes did you actually attend? ________

Survey Question B.

We are interested in where you studied. For a typical week during the Fall semester, tell us the percentage of your study time that took place in each of the following places.
Note: Numbers on the five lines should add up to 100
In dorm room (or at home if live off campus) with TV on _______
In dorm room (or at home if live off campus) without TV on _______
In library, empty classroom, quiet study lounge, or other quite place _______
In TV lounge, other (non-quiet) lounges _______
Other places _______

Question A on the survey asks that you carefully fill out a time diary which is a list of activities during the past 24 hours. In order to complete the time diary on the actual survey form on **page 3**, do the following:

1) Please put an arrow (-->) next to the time that it is right now. Label this arrow with the words **YESTERDAY** and **START**.

2) Now start with the box next to which you put the arrow (-->). Place in this box the activity you were doing during that time period yesterday.

For example, if it is now 7 p.m., you would put an arrow (-->) next to the box labeled "7:00PM".

Next to 7:00PM, you should write what you were doing from 7:00 to 7:20 **yesterday**.

Next to 7:20PM, you should write what you were doing from 7:20 to 7:40 **yesterday**.

Next to 7:40PM, you should write what you were doing from 7:40 to 8:00, and so forth.

As you proceed, you should work down the column below your arrow (-->) and then move to the top of the other column. Complete this other column and then move back to the top of the column where you started and finish filling in until you reach the arrow(-->).

When you begin to fill in the time period boxes, you will be writing your activities from yesterday until you reach the box labeled 12:00 midnight. From then on, you will be writing about your activities earlier today.

A sample completed time diary

Time Period	What were you doing?	Time Period	What were you doing?
MORNING		EVENING	
6:00 AM	SLEEPING	6:00 PM	
6:20 AM		6:20 PM	EATING
6:40 AM	PERSONAL	6:40 PM	
7:00 AM	EATING	7:00 PM	← START YESTERDAY
7:20 AM		7:20 PM	SHOPPING
7:40 AM		7:40 PM	
8:00 AM	IN	8:00 PM	
8:20 AM	CLASS	8:20 PM	
8:40 AM		8:40 PM	
9:00 AM	WORKING (Labor)	9:00 PM	
9:20 AM		9:20 PM	STUDYING
9:40 AM		9:40 PM	
10:00 AM	IN	10:00 PM	
10:20 AM	CLASS	10:20 PM	RECREATION AND
10:40 AM		10:40 PM	STUDYING
11:00 AM		11:00 PM	
11:20 AM	WORKING (Labor)	11:20 PM	
11:40 AM		11:40 PM	STUDYING
AFTERNOON		NIGHT	
12:00 noon	EATING	12:00 midnight	
12:20 PM		12:20 AM	
12:40 PM		12:40 AM	
1:00 PM	IN	1:00 AM	
1:20 PM	CLASS	1:20 AM	
1:40 PM		1:40 AM	
2:00 PM		2:00 AM	
2:20 PM	EXERCISING	2:20 AM	SLEEPING
2:40 PM		2:40 AM	
3:00 PM		3:00 AM	
3:20 PM	STUDYING	3:20 AM	
3:40 PM		3:40 AM	
4:00 PM	IN	4:00 AM	
4:20 PM		4:20 AM	
4:40 PM	CLASS	4:40 AM	
5:00 PM		5:00 AM	
5:20 PM		5:20 AM	
5:40 PM		5:40 AM	

Note(1): The activities will be chosen from the 13 words in **BOLD** which are listed on **page 3** to the right of the time diary form that you will complete

Note(2): Notice in the example that the brace symbol (}) is used when an activity continues through several time periods.

Note(3): If you are involved in two activities during the same time period(s), please list both activities and circle the activity you spent more time on.

Studying (outside of class)

includes studying for your classes, preparation for class, studying for an exam, doing take-home exams, **homework**, writing essays and papers, optional study sessions, any other work done outside of class time for your classes.

Survey #5 (Please complete both sides of this sheet) CPO 1971 (3)

Question A.

Reminders: Be sure to put an arrow (–>) next to the time that it is right now. And label this arrow with the words **YESTERDAY** and **START**.

Beginning with the **What were you doing** box next to the arrow, fill in your activities starting 24 hours ago (yesterday) and ending right before you began completing this survey.

Please use the 13 words listed in BOLD on the right of this page to describe your activities.

Time Period	What were you doing?	Time Period	What were you doing?
MORNING		EVENING	
6:00 AM		6:00 PM	
6:20 AM		6:20 PM	
6:40 AM		6:40 PM	
7:00 AM		7:00 PM	
7:20 AM		7:20 PM	
7:40 AM		7:40 PM	
8:00 AM		8:00 PM	
8:20 AM		8:20 PM	
8:40 AM		8:40 PM	
9:00 AM		9:00 PM	
9:20 AM		9:20 PM	
9:40 AM		9:40 PM	
10:00 AM		10:00 PM	
10:20 AM		10:20 PM	
10:40 AM		10:40 PM	
11:00 AM		11:00 PM	
11:20 AM		11:20 PM	
11:40 AM		11:40 PM	
AFTERNOON		NIGHT	
12:00 noon		12:00 midnight	
12:20 PM		12:20 AM	
12:40 PM		12:40 AM	
1:00 PM		1:00 AM	
1:20 PM		1:20 AM	
1:40 PM		1:40 AM	
2:00 PM		2:00 AM	
2:20 PM		2:20 AM	
2:40 PM		2:40 AM	
3:00 PM		3:00 AM	
3:20 PM		3:20 AM	
3:40 PM		3:40 AM	
4:00 PM		4:00 AM	
4:20 PM		4:20 AM	
4:40 PM		4:40 AM	
5:00 PM		5:00 AM	
5:20 PM		5:20 AM	
5:40 PM		5:40 AM	

LIST OF WORDS in bold

In Class
Attending class, attending labs, attending required class sessions

Studying (Outside of class time)
(refer to **pg 2** for more details)

Athletics
(Intercollegiate or Intramural - games or practice)

Clubs

Exercising

Recreation
(reading which is unrelated to courses, listening to music, watching movie, spending time with friends, etc.)

Shopping

Eating

Sleeping

Partying

Personal

Working (in Labor position)

Other
(Please describe on your sheet)

Appendix B. Do the additional instruments (from Section 5) satisfy the exogeneity requirement?

Tables Appendix.1a and Appendix.1b present results analogous to Tables 3a and 3b for the RSTUDYHS variable. Appendix.2a, and Appendix.2b present results analogous to Tables 3a and 3b for the REXSTUDY variable. There is virtually no evidence that behaviors other than study-effort are influenced by the presence of a roommate with particular values of REXSTUDY and RSTUDYHS. The RSTUDYHS variable is not significant at .10 in any of the eight regressions in Appendix.1. The REXSTUDY variable is significant at .10 in only one of the eight regressions in Appendix.2 with students in the sample who have roommates who expected to study one more hour per day in college going to be bed about six minutes later per night.

Table Appendix.1a
The effect of RSTUDYHS on other behaviors, n=176

Independent Variable	**Dependent Variable PATTEND proportion of classes attended** **estimate (std. error)**	**Dependent Variable CLASSHOURS daily hours in class** **estimate (std. error)**	**Dependent Variable SLEEP daily sleep hours** **estimate (std. error)**	**Dependent Variable BEDTIME time student went to sleep#** **estimate (std. error)**
RSTUDYHS	.0001 (.0004)	-.001 (.009)	-.007 (.010)	-.006 (.008)
MALE	.0007 (.009)	.005 (.194)	.307 (.217)	-.125 (.200)
CONSTANT	.956 (.008)**	3.452 (.164)**	7.226 (.184)**	.891 (.130)*
	R^2=.0006	R^2=.0002	R^2=.014	R^2=.011

*significant at .10
**significant at .05
#dependent variable is created so that it is zero at 12:00 midnight. Positive numbers represent hours after midnight. Negative numbers represent hours before midnight.

Table Appendix.1b
The effect of RSTUDYHS on additional behaviors

Independent Variable	**Dependent Variable percentage of study time that takes place in dorm room** **estimate (std. error)**	**Dependent Variable percentage of study time that takes place in dorm room with tv on** **estimate (std. error)**	**Dependent Variable hours per week using computer for academic purposes** **estimate (std.error)**	**Dependent Variable daily hours partying** **estimate (std. error)**
RSTUDYHS	.199 (.226)	.905 (.804)	-.006 (.053)	-.001 (.002)
MALE	-5.823 (4.622)	-3.838 (2.968)	–.120 (1.096)	-.001 (.050)
CONSTANT	59.828 (3.959)**	11.024 (3.550)**	7.104 (.938)**	.126 (.043)**
	R^2=.014	R^2=.018	R^2=.0002	R^2=0.011

*significant at .10
**significant at .05

Table Appendix.2a
The effect of REXSTUDY on other behaviors

Independent Variable	**Dependent Variable PATTEND proportion of classes attended**	**Dependent Variable CLASSHOURS daily hours in class**	**Dependent Variable SLEEP daily sleep hours**	**Dependent Variable BEDTIME time student went to sleep#**
	estimate (std. error)	**estimate (std. error)**	**estimate (std. error)**	**estimate (std. error)**
REXSTUDY	.0009 (.002)	-.001 (.053)	.022 (.059)	.097 (.055)*
MALE	.0008 (.009)	.005 (.195)	.316 (.218)	-.118 (.200)
CONSTANT	.955 (.011)**	3.444 (.232)**	7.071 (.260)**	.503 (.235)**
	R^2=.0007	R^2=.0000	R^2=.012	R^2=.020

*significant at .10
**significant at .05
#dependent variable is created so that it is zero at 12:00 midnight. Positive numbers represent hours after midnight. Negative numbers represent hours before midnight.

Table Appendix.2b
The effect of REXSTUDY on additional behaviors

Independent Variable	**Dependent Variable percentage of study time that takes place in dorm room**	**Dependent Variable percentage of study time that takes place in dorm room with tv on**	**Dependent Variable hours per week using computer for academic purposes**	**Dependent Variable daily hours partying**
	estimate (std. error)	**estimate (std. error)**	**estimate (std.error)**	**estimate (std. error)**
REXSTUDY	.964 (1.258)	.905 (.804)	.299 (.298)	-.020 (.013)
MALE	-5.588 (4.643)	-3.838 (2.968)	-.019 (1.097)	.014 (.050)
CONSTANT	58.441 (5.554)**	11.024 (3.550)**	5.940 (1.31)**	.182 (.060)**
	R^2=.008	R^2=.018	R^2=.006	R^2=0.012

*significant at .10
**significant at .05