2001

2001-1 Working During School and Academic Performance

Todd R. Stinebrickner

Ralph Stinebrickner

ISSN: 1702-0085
ISBN: 0-7714-2353-5

Working During School and Academic Performance

by

Todd R. Stinebrickner and Ralph Stinebrickner

Working Paper # 2001-1 December 2001

CIBC Working Paper Series

Department of Economics
Social Science Centre
The University of Western Ontario
London, Ontario, N6A 5C2
Canada

This working paper is available as a downloadable pdf file on our website
http://www.ssc.uwo.ca/economics/centres/cibc/

Working during school and academic performance

Todd R. Stinebrickner and Ralph Stinebrickner[1]

please direct correspondence to
Todd R. Stinebrickner
Dept. of Economics
The Social Science Centre
The University of Western Ontario
London Ontario Canada n6a 5c2

trstineb@julian.uwo.ca
phone 519 679-2111 ext. 85293
fax 519 661-3666

Unique new data from a college with a mandatory work-study program are used to examine the relationship between working during school and academic performance. Particular attention is paid to the importance of biases that are potentially present because the number of hours that are worked is endogenously chosen by the individual. The results suggest that, even if results appear reasonable, a researcher should be cautious when drawing policy conclusions about the relationship between hours-worked and a particular outcome of interest unless he/she is confident that potential problems associated with the endogeneity of hours have been adequately addressed.

---

[1]Todd is a CIBC Fellow in Economics at The University of Western Ontario. Ralph is a Professor of Mathematics at Berea College and director of the Osceola Research Institute. We would like to thank Pam Thomas for her invaluable assistance with data preparation and John Bound, Dan Black, and Nicholas Rowe for useful comments. This work was also facilitated by generous support from the Mellon Foundation, SSHRC, Berea College, and the Agnes Cole Dark Fund.

I. Introduction

Important policy decisions have been based on beliefs about the relationship between working during school and a student's current and future academic performance. However, although some previous research examines these matters, currently no consensus exists on the effect that youth employment has on these outcomes. If working during high school has a harmful effect on academic performance, it might be reasonable to strengthen laws that regulate the number of hours that youth can work. Similarly, if working during college is detrimental, individuals who need to work during college in order to pay tuition costs may be at a disadvantage when compared to students from wealthier backgrounds, and work-study based financial aid programs may have certain undesirable side effects.

Difficulty in determining the true impact that work has on academic performance arises largely because the number of hours that an individual works is endogenously chosen. For example, it is sometimes posited that individuals who fare well academically in school tend to be blessed with high levels of "motivation" that may also make them more likely than other students to become involved with non-academic activities such as work.[2] In empirical work, if "motivation" is not fully observed, some of the variation in academic performance that should be attributed to differences in motivation may mistakenly be attributed to differences in work status. Thus, in this scenario, simple econometric models may understate the negative impact that working has on school performance. Indeed, as discussed in the next section, previous studies have sometimes found that academic performance is highest among individuals who are working a moderate number of hours.

Credibly dealing with the endogeneity of work hours in empirical work is typically very difficult. It is true that certain individual characteristics may be able to capture some portion of factors such as

---

[2]This may be a reasonable scenario for students working low to moderate number of hours. However, students who work large numbers of hours may be individuals who are not particularly interested in academics.

motivation. However, given the large amount of variation in academic performance that typically remains unexplained after controlling for available individual characteristics, it is likely that including individual characteristics provides only a relatively low level of protection against this type of endogeneity. Instrumental variable (IV) estimators represent a theoretically appealing way to deal with the endogeneity problem. However, in practice it is often difficult to find instruments that both explain a reasonable portion of the variation in work hours and are unrelated to an individual's academic performance except through their effect on hours worked.[3] Unfortunately, when methods for controlling for endogeneity are not entirely satisfactory, the researcher can be left without any manner in which to determine the extent to which bias may be present in his/her estimates. Even when estimates seem "reasonable," a large amount of bias may be present.

In this paper, we utilize unique new data in an attempt to examine the extent to which the endogeneity of hours may bias estimates of the effect of employment on academic performance. The data are obtained directly from the administrative records of Berea College. Located in central Kentucky, this liberal arts institution operates under a mission of providing an education to those who "have great promise, but limited economic resources." As part of this mission, all students who attend Berea receive full tuition scholarships. Part of the cost of schooling is defrayed through a mandatory work-study program. Although all students must work at least a minimum of ten hours a week, variation in hours worked arises because students can often choose to earn extra income by working additional hours. We wish to note in advance that, given the unique nature of Berea College, it is our belief that our results should be viewed cautiously. Nonetheless, as will be described throughout the paper, the institutional details of the Berea College labor program and the detailed nature of our administrative data

[3]See Bound et al. (1995) for a discussion of the potential problems that can arise in instrumental variables estimation when the correlation between instruments and the endogenous explanatory variable is weak.

provide us with a unique opportunity to obtain some information about the potential importance of the endogeneity problem described above.

In particular, we are able to compare estimates of the effect of working on academic performance obtained using an Ordinary Least Squares (OLS) estimator, a fixed effects (FE) estimator, and an instrumental variables (IV) estimator. The FE estimator takes advantage of semester-by-semester information on hours worked that is available in our administrative data. The IV approach takes advantage of two features of the labor program at Berea that imply that the job to which a person is assigned in the first semester can serve as an instrument for hours worked. First, as will be discussed in detail, it is reasonable to believe that the first-year job assignment process (which takes place before students arrive at Berea) does not create a situation in which the "motivation" level of students is higher in some jobs than in others. Second, the amount of hours that are "available" for a person to work depends on the job to which he/she is assigned. There are intuitive reasons to believe that both the OLS and FE estimators may tend to understate the negative effect that working has on academic performance. Our results suggest that the amount of bias may be substantial. Whereas the estimators both suggest a positive relationship between working and academic performance, the IV estimator indicates that working additional hours has a harmful effect on academic performance.

The remainder of the paper proceeds as follows. In Section II, we briefly examine past research that considers the relationship between employment while in school and academic or labor market outcomes. In Section III, we describe the data from Berea College and we highlight some of the advantages of these data. In Section IV, we provide the details of the results summarized above. In Section V we conclude.

II. Previous Work

Ruhm (1997) provides a very thorough overview of the research from several academic disciplines which examines the effect of working during school. He concludes that for employment in

high school "there is currently no consensus whether student employment improves or worsens school performance." His summary suggests that previous studies have varied substantially in terms of the samples and the methodologies that have been employed. In many cases, the endogeneity of the work decision and the hours decision was not addressed in these studies. Further, the very few cases where attempts were made to deal with the endogeneity problem met with limited success because of difficulty in finding credible instrumental variables.[4] For example, OLS estimates in Turner (1994) indicate that, for students in the High School and Beyond survey, working a moderate number of hours is positively related to high school performance. The effect of hours is found to be statistically insignificant when fixed effects and IV estimators are employed. However, because some of his instruments are potentially related to academic performance as well as the decision to work in high school, it becomes difficult to tell whether working during high school is beneficial, harmful, or unimportant from the standpoint of academic performance.[5]

Ruhm reports that previous studies of the effect of work in high school on future wage outcomes are more conclusive and typically conclude that work in high school is associated with increased future earnings. However, these studies are subject to the same difficulties in adequately controlling for the potential selectivity in hours worked. This problem is recognized by Ruhm who uses geographic characteristics to identify a model which controls for the endogeneity of work. Although his instruments pass the Newey (1985) test for exogeneity based on overidentification restrictions, they provide only weak predictive power for hours. As a result, the effect of hours is measured very imprecisely and the

---

[4]Eckstein and Wolpin (1997) take a very different approach to dealing with the endogeneity problem by formulating and estimating an explicit sequential decision model of high school attendance and work. They find that working while attending high school does reduce academic performance, but the quantitative effects are small.

[5]For example, a student's "access to money from parents" would seem to be correlated with family background/income which has a well-known effect on academic achievement.

correction for selection bias meets with "limited success."[6]

Less research examines the effect of working during college on academic performance. However, a lack of consensus also appears in this context. For example, Paul (1982) finds that working is detrimental to academic performance in college, Hood et al. (1992) find that grade point averages are highest among students with moderate amounts of work, and Ehrenberg and Smith (1987) find positive effects of working in on-campus jobs but negative effects of working in off-campus jobs.

III. Berea College and the Berea College Labor Program

III.1 Overview of Financial Aid and the Labor Program

Consistent with the mission of the school discussed in the introduction, students who are admitted to Berea are typically quite poor; Stinebrickner and Stinebrickner (1998) find that the mean and standard deviation of family income for dependent students who entered Berea between 1989 and 1997 is 22,450 and 13,590. However, each student at Berea receives a full tuition subsidy, and the average student receives more than $3,000 of additional financial aid from private, state, and federal sources.[7] As part of this financial aid package, each student is required to work a minimum of ten hours a week in the school's labor program. In addition to receiving a $2800 labor grant which is part of the financial aid package described above, the student/worker receives a small hourly wage. Before their arrival to Berea, students are assigned to specific "service" type jobs for their first year which pay approximately two dollars an hour. Hourly wages in a particular job increase after the first year with

---

[6]For other recent work which examines the relationship between employment during school and earnings see Light (forthcoming) who finds that high school employment has a "small and relatively short-lived effect on post-school wages" and Hotz et al. (1998) who estimate a dynamic discrete choice model using the NLSY. Light (1998a, 1998b) examines the relationship between in-school work experience and returns to schooling.

[7]This leaves the entering student with an average total room, board, and college fee bill of only approximately $1000. These numbers were taken from the Berea College admissions brochure, 1998. Students graduate from Berea with an average of approximately $1,000 in student loans.

"length of service." In addition, after the first year, students may choose to compete for higher-paying jobs that they find desirable. Students at the upper-end of the pay scale receive approximately four dollars an hour.

III.2 Measurement Benefits of Our Administrative Data

From their administrative records, Berea College made available records for all of the full-time students that matriculated between the fall semester of 1989 and the fall semester of 1997. There are several reasons that these administrative data are useful from the standpoint of measuring the variables of interest for this study. First, because hours will not tend to be constant across the entire school year, it is desirable to link academic performance and hours within time units that are as short as possible. Our data allow us to measure hours and our measure of academic performance, grades, at a semester level rather than at the yearly level which is typically the smallest time unit possible in other studies.[8] Second, the use of administrative data ensures that our semester hours measure is not subject to the types of measurement error that are potentially problematic in longitudinal data which rely on self-reports of hours worked.[9] In other studies, where the outcomes of interest (e.g., grades or test score performance) are measured at yearly intervals, the appropriate hours measure should reflect the total number of hours worked over the entire academic year. However, past evidence that employment hours tend to be misreported in retrospective data suggests that precisely measuring total yearly hours may be difficult. For example, in a validation study of the Panel Study of Income Dynamics, Duncan and Hill (1985) find that respondents overstate the number of hours that they work by between 10% and 12%, and evidence of this problem is also found for the youth in the NLSY by Ruhm (1997). In part because

[8]Although it cannot be examined directly using our data, students at Berea tend to work roughly the same amount each week during a particular semester. Thus, it also seems likely that less inter-week variation in hours will be present for a person at Berea than a person in other data.

[9]Students are not permitted to hold off-campus jobs.

of this problem and in part because some surveys do not contain measures of yearly hours, researchers have often sensibly chosen to characterize hours using responses to questions about the number of hours which were worked by the respondent in a particular reference week. Unfortunately, even if respondents answer this question accurately, using reference week hours to construct yearly work hours will lead to measurement error in the latter if, as indicated by Ruhm (1997), variation exists across weeks in the number of hours that are worked. Finally, because we study a single school, our outcome measure is directly comparable across the different individuals in the sample.[10] Of course, with respect to the latter, it is important that our conclusions take into account the unique nature of our sample.

III.3 A Description of Employment Hours

During the first semester, individuals are not officially "encouraged" to work more than the ten hour a week minimum. In practice, this implies that students are not allowed to add "secondary" jobs to the "primary" jobs that they are assigned for their freshman year. However, variation does appear in work hours during the first semester because some degree of hours flexibility exists within a person's primary job; for our sample, the mean and standard deviation of hours in the first semester is 10.97 and 1.40 respectively. As we will discuss in more detail, some of the variation in first semester hours is explained by the fact that individuals are assigned to different labor jobs.

What is not entirely clear is why students often work more than the minimum number of hours given the low marginal wage rate. One possibility is that, even with full tuition subsidies, the marginal utility of income may be very high for students at Berea who are typically from low income families.[11]

---

[10]This is particularly true in the early semesters when individuals are satisfying general, liberal arts requirements.

[11]In results not shown here, we did not find evidence that lower income individuals within our sample choose to work more hours than higher income individuals. However, even the highest family income earned at Berea is not particularly high. The fact that many students choose to add secondary jobs of a service nature in the second semester may represent some evidence that this explanation is relevant. Average hours are 11.53, 12.34, 12.76, 13.22, 13.55, 14.04, and 14.02 in the second through

A second possibility is that, when particular labor departments face worker "shortages" (in the sense that all necessary hours would not be covered if all assigned workers worked the ten hour minimum), students offer to work additional hours or decide to satisfy requests to work additional hours in order to remain in good standing with their supervisors. Remaining in good standing with a supervisor may make working generally more enjoyable or may lead to positive recommendations from the supervisor in the future. A final possibility is that students find that time spent working is somewhat substitutable with other available leisure activities.

IV. Estimation and Results

In order to estimate the relationship between academic performance, we examine the equation

$$(1) \qquad G_{it} = \alpha X_i + \beta H_{it} + \epsilon_{it}$$

where G_{it} is student i's grade point average for semester t, H_{it} is the number of hours worked per week by person i in semester t, and X_i is the vector of person i's individual characteristics described in the first column of Table 1.[12] Our primary interest is in β. The second and third columns of Table 1 show results when equation (1) is estimated by OLS using data from the first semester (t=1) and the second semester (t=2) respectively. The results indicate that, on average, individuals who work more hours receive statistically **higher** grades, with an additional hour of work per week being associated with

---

eighth semesters respectively. It is likely that this increase is due to a combination of 1) an increase in the wage that is paid as individuals increase their job rank and experience standing, 2) an increase in the benefits of human capital accumulation as students are able to switch from service-type jobs to jobs which are more relevant to their future careers, 3) an increase in the non-pecuniary benefits of working as students become able to choose jobs which are more interesting to them, and 4) an increase in the ease of adding hours as students become familiar with the labor system and learn about potential job opportunities.

[12]Family income, which is missing for a subset of our students (e.g., those who report being independent), is not included in X. However, the results in this paper were found to be very similar when family income was included (and estimation utilized the subset of individuals with non-missing family income information).

increased grade point averages of .043 in the first semester and .021 in the second semester.[13] Given the types of service jobs in which individuals work in the first year, it seems unlikely that a given person's academic performance would improve if he decided to add hours above the weekly minimum of ten hours. Instead, it seems likely that the OLS estimator of β is biased by the fact that the number of hours is endogenously chosen. The benefit of our data is that they allow us to take further steps to analyze the extent to which bias might be present.

Endogeneity bias will be present to some extent if there exist unobserved determinants of grades (i.e., elements of ϵ_{it}) which also influence the number of hours that a person works. Perhaps the most commonly discussed scenario in this context is one in which the "problematic" unobserved determinants are person-specific, permanent, individual attributes such as "motivation" or "energy" which might drive certain people to both work more and to study more. In this scenario, β can be estimated consistently using a fixed effects estimator (FE) that takes advantage of grade point average and hours information from multiple semesters. The FE point estimate of β obtained using data from students' first two semesters at Berea is .010 and the standard error of the estimator is .006.[14] The null hypothesis that additional hours have no effect on academic performance (i.e., β=0) can no longer be rejected at standard levels of significance less than .10.

Thus, if the only source of endogeneity in hours arises because of permanent unobservable factors, it would be reasonable to conclude that additional hours of work have little effect on academic performance. However, in this and other applications in which a researcher is interested in the

---

[13]The difference in the first and second semester point estimates appears to be related to the sample selection that takes place as some individuals do not return for the second semester. The first semester point estimate is .024 when equation (1) is estimated using only those individuals who returned for the second semester.

[14]Students who are only observed for one semester do not contribute to the FE estimator of β. Using more than two semesters of data creates complications because comparing grades across semesters becomes more difficult as individuals begin to take classes that are specific to their majors.

relationship between hours and a particular economic outcome, there are also likely to exist sources of hours endogeneity that do not fit the description of being constant across time for a particular person. For example, in this application it seems very likely that individuals find it optimal to increase hours in semesters when classes (or other activities) are not particularly burdensome. In this scenario, the estimate of β obtained using the FE estimator may tend to understate the true negative impact that working has on school performance.

In the remainder of the paper, we examine whether we can learn more about the importance of endogeneity through an instrumental variables approach that uses a person's first semester job assignment as an instrument. We distinguish five specific jobs that employ a relatively large number of first-year students: janitor, worker in the Berea College crafts industry, worker in the campus hotel/restaurant, worker in the library, and worker in the audio-visual department. The remainder of students work in a wide-array of jobs around campus that we group into a category called "Other." The first row of Table 2 shows the number of entering freshmen who are assigned to each of the six jobs at Berea between 1989 and 1997.

In order to use an entering student's job assignment as an instrument, three conditions must be satisfied. First, the job to which an individual is assigned must influence the number of hours that the student works. Second, the student's job must not have a direct effect on academic performance. That is, the only influence that a job has on academic achievement must come through its effect on the number of hours that are worked. Third, the process by which jobs are assigned must not lead to a case in which individuals with an unobserved propensity to fare well academically or an unobserved ability to balance academics and work are grouped together in particular jobs. We first discuss these three assumptions in more detail. We then turn to examining the importance of endogeneity in this

application under the assumption that these three conditions are true.[15]

The second condition requires that jobs do not have a direct effect on academic performance. When considering the validity of this assumption, it is important to keep in mind that the jobs that are assigned for the first semester are virtually all "service" type jobs. Therefore, although specific tasks do vary somewhat by job assignment, the general nature of the various jobs is quite similar. For example, by and large, the service type jobs require an element of physical activity and do not provide time that can be used for studying. Thus, there is no obvious reason that a student in any particular job would be at a disadvantage because he finishes work especially exhausted or because he receives less study time than students in other jobs.[16] Thus, overall it seems quite plausible to believe that jobs do not have a direct effect on academic performance.

As mentioned earlier, students are not allowed to add secondary jobs in their first semester at Berea. Thus, the number of hours that a person can work in the first semester is limited to the hours that can be obtained in his assigned job. In order for our first condition to be satisfied, it must be true that job assignment explains some of the variation in first semester hours worked. In the first column of Table 3, first semester weekly hours are regressed on a set of individual observable characteristics. The R^2 for the regression is only approximately .002, and a test of the null hypothesis that the coefficients are jointly insignificant cannot be rejected at a .10 level of significance.[17] In the second column, dummy variables for the jobs are added. Three jobs (hotel, craft, and janitor) have t statistics with an absolute value greater than 2.76 (and as large as 7.90), another job (library) has a t statistic of -1.89. As a result,

---

[15]The second and third conditions imply the standard condition that the instrument is uncorrelated with the unobservable in the outcome equation.

[16]Even library jobs, that intuitively might seem to be the most different of all the jobs listed in Table 2, are similar in nature. For example, a library worker in his first year is likely to be involved in jobs such as shelving books and cleaning and maintenance activities.

[17]The test statistic for the F test is 1.61. The p-value associated with the test is .168.

the null hypothesis that the coefficients on the job indicators are jointly zero is overwhelmingly rejected with an F statistic of 25.840.[18] The inclusion of the job indicators leads to an increase in the R^2 from .002 to .06.

People involved with the labor program at Berea indicated that the variation in hours across jobs has traditionally arisen because some job have had more hours "available" than other jobs. For example, whereas an administrator in the crafts department indicated that only a small amount of hours flexibility is available for employees in this job, it is known that the hotel has traditionally had more than the minimum number of hours available. As discussed earlier, in jobs with more available hours, students may offer to work more hours or may decide that it is worthwhile to satisfy requests from supervisors to work more hours.[19]

The third condition requires that students are not sorted into jobs on the basis of unobservable characteristics such as motivation. The official policy of Berea College is to not sort students into jobs on the basis of academic ability, and, except for a very small number of first-year positions that are perhaps not truly of "service" nature, the college has no incentive to deviate from this policy. Nonetheless, in order to conclude that it is reasonable to believe that the third condition is satisfied, it

---

[18]We note that the quantitative difference in the average hours between jobs is not overwhelming. For example, relative to the Other category, individuals assigned to the campus hotel work an additional 8 extra hours per semester (.6 of an hour more per week) and students assigned to the crafts industry work 7 fewer hours per semester (.5 of an hour less per week). However, it seems very possible that small differences in hours could be important given that all students are attending school full time and are already working at least a minimum of approximately ten hours a week.

[19]A student's decision of whether to work more hours is likely to depend to some extent on his enjoyment of the job. Given the previous argument that jobs are similar in nature, it seems unlikely that differences in the non-pecuniary benefits of particular jobs would explain much of the inter-job variation in hours. Regardless, variation that arises due to differences in job enjoyability is not problematic under the assumption that jobs do not have a direct effect on academic performance. If jobs tend to have very similar physical requirements and do not allow study time, this assumption will only be violated if the enjoyable job produces good "feelings" which somehow improve a person's concentration or desire to study when the person is not at work. Further, even if such a spillover effect does occur for some jobs, the bias which would result would tend to lead to a conservative estimate of the negative effect that working has on grades.

is necessary to examine the job assignment process in more detail. In an effort to give students some choice about their first-year job, students are asked before arriving at Berea to rank their preferences for three of the most common first-year jobs: hotel, janitor, crafts. Because these three jobs are not collectively exhaustive, a subset of the first-year students is randomly selected and assigned to first-year jobs other than hotel, janitor, and crafts. The remaining students are assigned to the hotel, janitor, and crafts jobs. For these students, an effort is made to take into account the students' preference rankings.

Thus, given the policy of the College, the only potential for non-random selection into jobs comes from the fact that students have an opportunity to express their preferences towards the hotel, janitor, and crafts jobs. However, it is important to stress that in order for this to be problematic it would have to be the case that students with high levels of unobserved motivation/ability prefer particular jobs within this group while individuals with low levels of unobserved motivation/ability prefer other jobs. Given the well-advertised service nature of the first-year hotel, janitor, and crafts jobs, there does not seem to be any reason to believe that this type of selection would be problematic. Further, as we will see later, in order for our IV results to be explained by this type of sorting, it would have to be the case that, for example, individuals with high levels of unobserved academic ability or motivation are more likely than other students to prefer the janitor job and less likely to prefer the hotel job. Intuitively, this seems particularly unappealing and the alternative scenario would tend to imply that our IV results are somewhat conservative.[20]

Further, while it is impossible to directly examine whether the third condition is satisfied, it is possible to provide indirect evidence by examining the distribution of observable characteristics across

[20]If one is simply willing to assume that the motivation of hotel workers is at least as high as that of janitors, an accurate or conservative estimate of the effect that working has on academic performance can be obtained (without any other beliefs about the assignment process) by estimating the IV model using only the 700 students in these two groups. We find that the estimated negative effect of an additional hour of work, -.18, is very similar to the IV estimate found later using the full-sample. The t-statistic is -2.36.

jobs. Earlier, we suggested that there may be a very small number of positions whose nature differs from that of other first-year jobs. Specifically, we had in mind the small number of positions in the audio-visual department which is the only labor department that we have ever heard may receive a non-random selection of students.[21] Thus, it is not particularly surprising that Table 2 indicates that the average American College Testing (ACT) scores of individuals in the audio-visual job are higher than the average ACT scores of individuals in the other jobs. However, Table 2 provides compelling indirect evidence that this type of sorting (or problematic sorting due to student preferences over the craft, library, and hotel jobs) does not exist for any of the other jobs. The null hypothesis that mean total ACT scores for two jobs are equal cannot be rejected at any level of significance lower than .50 for any of the 10 possible job pairs that do not include the audio-visual job.[22]

Previous work has revealed that it is very difficult to find predictors of hours which do not also potentially have a direct relationship to academic performance. It seems plausible to believe that our job assignment variable satisfies the necessary conditions to be a valid instrument for hours in the first semester, in which case is it is possible to learn something about the importance of endogeneity bias at

---

[21]Workers in the audio-visual department are responsible for handling expensive and technically advanced equipment. Using this as a rationale, the the audio-visual department (during our sample period) petitioned for workers of higher ability. We know of no other case where this took place (or other first year jobs where a similar rationale would be present).

[22]The fact that males and females may express different preferences among the craft, janitor, and hotel jobs implies that we should examine ACT scores between jobs after controlling for sex and possibly race. A regression of combined ACT scores on sex, race, and job dummies (excluding audio) indicates that jobs are not related to ACT scores. The highest t-statistic (in absolute value) associated with the jobs (relative to the Other job category) is only .559 and the null hypothesis that the effect of all jobs is zero has a F statistic of only .15 which implies that the null can only be rejected at levels of significance greater than .962.

The assignment process also implies that the distribution of observable characteristics in the craft, hotel, and janitor group should be the same as the distribution of observable characteristics in the group containing all of the other jobs. The null hypotheses that ACT scores, sex, and race are equal in the two groups is not rejected at standard significance levels.

Berea.[23] We now proceed under the assumption that the instrument is valid and we examine this assumption more formally later. In the following analysis, we include the individuals in the audio-visual positions. This would be justified if, as seems likely, the decision of which students to place in the audio-visual positions is determined on the basis of characteristics such as college entrance exam scores that we also observe. However, we note that the results in the remainder of the paper are not sensitive to whether or not these individuals are included.

We begin by considering the intuition behind the IV estimator of β under the assumption that the instrument is valid. If no systematic differences exist between individuals assigned to the various jobs and job assignment has no direct effect on grades, then, under the hypothesis that hours have no effect on grades, individuals in each job should perform equally well in terms of grades even though the second column of Table 3 indicates that they work different numbers of hours. The third column of Table 3, which shows the results when GPA is regressed on personal characteristics and job indicators, suggests that this is not true. For example, individuals at the hotel receive grades that are .085 lower than individuals in the Other category. Janitors receive grades that are .096 higher. The former is significant at a level of .05. The latter is significant at a level of .10. A test of the null hypothesis that job assignment has no influence on grades is rejected at all levels of significance greater than .02.[24] Further, as can be seen by comparing the second and third columns of Table 3, what emerges is a pattern in which individuals who are assigned to jobs which have higher average hours tend to fare worse academically than individuals who are assigned to jobs that have lower average hours. Thus, this descriptive analysis suggests that, despite the earlier OLS and FE estimates, working may actually have a harmful effect on grade performance.

---

[23]Recall that in the second semester students can add secondary jobs.

[24]The F test of the null hypothesis that the coefficients on the job indicators are all equal to zero yields a test statistic of 2.64.

To provide a more formal treatment of our intuition, an IV estimate of β is obtained from a two-stage least squares procedure.[25] The first stage is the weekly hours equation shown in the second column of Table 3. The results of the second stage are shown in Table 4. We find that an additional hour of work has a large and statistically negative impact on grades. More specifically, increasing the amount that a person works by one hour per week lowers the person's semester grade point average by .162.

To provide some evidence regarding the plausibility of this result, we surveyed 311 current first-year students and found that the median student studies approximately 2.0 hours a day. If each additional hour of work leads to one less hour of study-time, an extra half an hour of work each day would imply that the median student would spend 25% less time studying.[26] It seems plausible to us that this could lead to the predicted grade point decline of approximately half a letter grade.[27]

To understand why the OLS and IV estimators yield very different results, realize that the OLS estimator uses both the between job variation in hours that is assumed by the IV estimator to be exogenous and the within job variation in hours that is assumed by the IV estimator to endogenous. Roughly speaking, the effect of the latter on the OLS estimator can be seen in Table 5 which shows OLS estimates of β for the subsample of students who were assigned to each particular job. The estimates of β are all **positive** and are generally statistically significant and quantitatively large. Given the service nature of the first-year jobs, it seems unlikely that more hours are causing better grades. Rather, it seems likely that these positive correlations arise because individuals who are more motivated to do

---

[25]See Heckman (1997), Heckman and Vytlacil (1998), Imbens and Angrist (1994), and Angrist and Imbens (1995) for recent general discussions of instrumental variables estimation.

[26]Theoretically speaking, study-time could decline by more or less than 1 hour for each additional hour of work. The former could occur if working is tiring or if there are significant time-costs of getting to work or preparing for other activities after work (e.g., taking showers etc.)

[27]The survey questions asked individuals about hours spent studying during the previous twenty-four hour period, on the last weekend day, and during the last week. See Juster and Stafford (1985) for an in-depth discussion of the issues related to eliciting time-use information.

well in school are also working more hours, in which case ignoring this source of variation in hours is desirable.

It seems likely that both the actions of students and the actions of labor supervisors contribute to a situation in which highly motivated students work more hours. With respect to the former, highly motivated students would be willing to work more hours than other students if 1) they believe that signalling their motivation level to current supervisors will be advantageous from the standpoint of obtaining the types of jobs that they hope to hold in the future or 2) they are not able to find stimulating extracurricular activities at Berea and do not receive as much utility from other available leisure activities (e.g., "hanging out", watching television, or drinking) as other students.[28] With respect to the latter, labor supervisors would be more likely to request additional hours from highly motivated students and would be more likely to approve additional work requests from highly motivated students if these students are more dependable and conscientious workers.

Finally, we examine the validity of the instruments using a test based on overidentification restrictions (see Davidson and Mackinnon (1993) or Newey (1985)). The test statistic is .949 which is much smaller than the critical vales for the test under any reasonable signficiance levels, and the joint null hypothesis that the grade equation is correctly specified and the job indicators are valid instruments cannot be rejected.[29]

V. Conclusion:

---

[28]After the first year, students have the option to compete for jobs. Motivated students would seem to be more likely than other students to apply for desirable positions (e.g., positions that lead to significant human capital accumulation) with stiff competition. For these students, positive recommendations from supervisors may be very important. Other less motivated students may decide to remain in service type jobs after the first year. For these students, positive recommendations from supervisors are probably not important.

[29]The critical values are 7.779, 9.487, and 13.276 at significance levels of .10, .05, and .01 respectively.

Our instrumental variables estimator suggests that working has a harmful impact on grade performance during the first semester at Berea. Based on our in-depth knowledge of the institutional details of Berea College, we believe that it is very plausible to think that the necessary conditions for the IV estimator are satisfied and that the results that we have obtained are reasonable. In support of these views, this work has received serious attention in recent policy discussions at Berea.

Previous work has suggested that the endogeneity of hours is likely to be a concern in other youth employment contexts and it seems natural to believe that the reasons for these concerns are generally somewhat similar in nature to those expressed here. As a result, the findings in this paper suggest that researchers should be cautious about drawing policy conclusions in situations where it is difficult to deal with the endogeneity issue in a satisfactory way. However, it is important to note that the effect that working has on academic performance will depend on many factors associated with a person's specific situation. Berea is a unique institution and we do not wish to conclude that negative effects of the same magnitude would necessarily exist in other youth employment contexts. In addition, it seems likely that the bias in OLS estimators will tend to be smaller when researchers are using data that contain more information about a person's "motivation."

Finally, from a policy standpoint, it is important to keep in mind that evidence from Ruhm (1997), Light (forthcoming) and others suggests that youth employment can have a beneficial impact on future income. Thus, the evidence in this paper, that working could potentially be more harmful than previous studies indicate, should not necessarily be taken to suggest that hours should be restricted for youth. Rather, it simply suggests that, along with the potential benefits, there may also be an academic cost to working while young.

References

Angrist, Joshua and Imbens, Guido. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," Journal of the American Statistical Association, June 1995, pp. 431-442.

Bound, John, Jaeger, David and Baker, Regina. "Problems With Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak," Journal of the American Statistical Association, June 1995, 90(430), pp. 443-450.

Davidson Russell and Mackinnon James. Estimation and Inference in Econometrics, New York, Oxford University Press, 1993.

Duncan, Greg and Hill Daniel. "An Investigation of the Extent and Consequences of Measurement Error in Labor-economic Survey Data," Journal of Labor Economics, October 1985, 3, pp. 508-532.

Eckstein, Zvi and Wolpin, Kenneth. "Youth Employment and Academic Performance in High School," 1997, Unpublished Manuscript.

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

Heckman, James. "Instrumental Variables," Journal of Human Resources, Summer 1997, 23(3),

pp. 441-462.

Heckman, James and Vytlacil, Edward. "Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling When the Return is Correlated with Schooling," Journal of Human Resources, Fall 1998, 33 (4), pp. 974-987.

Hood, Albert B., Craig, Andrew and Ferguson Bruce. "The Impact of Athletics Part-time Employment, and Other Academic Activities on Academic Achievement," Journal of College Student Development, September 1992, 33, pp. 447-53.

Hotz, Joseph, Xu Lixin, Tienda Marta and Ahituv Avner. "Are There Returns to the wages of Young Men from Working While in School?" 1998, Unpublished Manuscript.

Imbens, Guido and Angrist, Joshua. "Identification and Estimation of Local Average Treatment Effects," Econometrica, 1994, pp. 467-475.

Juster, Thomas and Stafford, Frank. Time, Goods, and Well-Being, Survey Research Center, Institute for Social Research, The University of Michigan, 1985.

Light, Audrey. "High School Employment, High School Curriculum, and Post-School Wages," forthcoming Economics of Education Review.

Light Audrey. "Estimating Returns to Schooling: When Does the Career Begin?" Economics of Education Review, 1998a, 17(1),pp. 31-45.

Light Audrey. "In-School Work Experience and the Returns to Schooling," 1998b, Unpublished Manuscript.

Newey W. "Maximum Likelihood Specification Testing and Conditional Moments Tests," 1985, Econometrica, 53, 1047-70.

Paul, Harvey. "The Impact of Outside Employment on Student Achievement in Macroeconomic Principles," Journal of Economic Education, Summer 1982, pp. 51-56.

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

Stinebrickner, Todd and Stinebrickner, Ralph. "The Underlying Relationship Between Family Income and Educational Attainment: Evidence from a Liberal Arts College with a Full Tuition Subsidy," Unpublished Manuscript, The University of Western Ontario, 1998.

Turner, Mark D. "The Effects of Part-Time Work on High School Students Academic Achievement," 1994, Unpublished Manuscript.

Table 1
Descriptive statistics and OLS regression for 1st and 2nd semester GPA
first semester and second semester separately

	data	1st semester GPA	2nd semester GPA
	mean std dev	estimate (std error)	estimate (std error)
weekly work hours		.043 (.010)**	.021 (.006)**
black	.10	-.183 (.053)**	-.185 (.051)**
male	.46	-.154 (.031)**	-.117 (.030)**
math ACT	22.35 4.36	.038 (.004)**	.036 (.004)**
verbal ACT	20.51 3.92	.036 (.004)**	.039 (.004)**
constant		.614 (.153)**	.890 (.111)**
Estimated Variance Error		.547	.451
		R^2=.145	R^2=.167
	n=2372	n=2372	n=2078

* indicates t statistics greater than 1.65
** indicates t statistics greater than 2.0

Table 2
Descriptive Statistics by Job Assignment

	janitor	hotel	craft	library	other	audio
n	188	512	246	144	1180	102
combined Math and Verbal ACT	42.51 (.52)	42.61 (.29)	42.72 (.45)	43.02 (.59)	42.74 (.21)	46.38 (.55)

The first row shows the number of students assigned to each job in first semester.
The second row show mean total ACT exam scores for the students assigned to each job. In each box, the number in parenthesis is the estimated standard error of the sample mean.

Table 3
OLS Regressions for First Semester Hours and First Semester Grades

	1st semester hours	1st semester hours	1st semester GPA
	estimate (se)	estimate (se)	estimate (se)
Constant	10.748 (.189)**	10.722 (.187)**	1.066 (.101)**
Black	-.094 (.100)	-.157 (.098)	-.177 (.053)**
Male	-.088 (.060)	-.065 (.059)	-.157 (.031)**
Math ACT	.004 (.007)	.005 (.008)	.038 (.004)**
Verbal ACT	.008 (.008)	.005 (.008)	.037 (.004)**
craft		-.465 (.096)**	.067 (.052)
janitor		-.297 (.107)**	.096 (.058)*
library		-.229 (.120)*	.055 (.065)
other		0.00 (omitted)	0.00 (omitted)
audio		.147 (.141)	-.008 (.076)
hotel		.573 (.072)**	-.085 (.039)**
Estimated Var of Error	1.970	1.871	.548
	R^2=.002	R^2=.060	R^2=.145
	n=2372	n=2372	n=2372

Notes: omitted job is "Other." F test that job indicators are jointly zero in second column yields a test statistic of 21.98.
* t statistic greater than 1.65
** t statistic greater than 2.0

Table 4
Instrumental Variables Estimates for GPA

	first semester GPA
	estimate (se)
weekly hours	-.162 (.050)**
black	-.202 (.057)**
male	-.172 (.034)**
math ACT	.039 (.004)**
verbal ACT	.038 (.004)**
constant	2.82 (.558)**
	n=2372

Estimates from two-stage procedure using job assignment as instruments.
* t statistic greater than 1.65
** t statistic greater than 2.0

Table 5
OLS Estimates of β from Equation (1)
Separately by Job

Job	Estimate of β
craft	.130 (.041)** n=246
janitor	.158 (.055)** n=188
library	.022 (.060) n=144
other	.050 (.016)** n=1180
audio	.137 (.078)* n=102
hotel	.038 (.018)** n=512

Each row shows the estimate of β obtained by estimating equation (1) using the subset of individuals assigned to the particular job.
* t statistic greater than 1.65
** t statistic greater than 2.0