NBER WORKING PAPER SERIES

THE IMPACT OF EMPLOYMENT DURING SCHOOL ON COLLEGE STUDENT ACADEMIC PERFORMANCE

Jeffrey S. DeSimone

Working Paper 14006
http://www.nber.org/papers/w14006

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2008

I thank the William T. Grant Foundation for funding, Farasat Bokhari and Cagatay Koc for detailed suggestions, and other participants in a session at the 2006 Southern Economic Association meetings for helpful comments. 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 Impact of Employment during School on College Student Academic Performance
Jeffrey S. DeSimone
NBER Working Paper No. 14006
May 2008
JEL No. I2,J22

ABSTRACT

This paper estimates the effect of paid employment on grades of full-time, four-year students from four nationally representative cross sections of the Harvard College Alcohol Study administered during 1993–2001. The relationship could be causal in either direction and is likely contaminated by unobserved heterogeneity. Two-stage GMM regressions instrument for work hours using paternal schooling and being raised Jewish, which are hypothesized to reflect parental preferences towards education manifested in additional student financial support but not influence achievement conditional on maternal schooling, college and class. Extensive empirical testing supports the identifying assumptions of instrument strength and orthogonality. GMM results show that an additional weekly work hour reduces current year GPA by about 0.011 points, roughly five times more than the OLS coefficient but somewhat less than recent estimates. Effects are stable across specifications, time, gender, class and age, but vary by health status, maternal schooling, religious background and especially race/ethnicity.

Jeffrey S. DeSimone
Department of Economics
University of Texas at Arlington
701 S. West St.
Arlington, TX 76019
and NBER
jdesimone@uta.edu

1. Introduction

Many high school and college students work part-time. Does this affect their school performance? Employment during school could improve grades if working fosters attributes that are complementary with academic success, such as industriousness or time management skills, or instead reduce grades by constraining time and energy available for schoolwork. Alternatively, working might be correlated with academic performance, yet not directly impact it, if unobserved student differences influence both labor supply and grades. Unmotivated students might neither work for pay nor receive good grades because they put little effort into the labor market or school. In contrast, students uninterested in academics might work long hours that would otherwise have been devoted to leisure. Students might underestimate the link between college achievement and future earnings (e.g. Jones and Jackson, 1990; Loury and Garman, 1995), or any associated positive externalities, when making labor supply decisions. If so, obtaining a consistent estimate of how such decisions affect academic performance is prospectively important for policy consideration.

For high school students, much research has been devoted to this question, yielding decidedly mixed evidence. Some studies estimated negative effects of part-time employment on school performance (e.g. Singh, 1998; Eckstein and Wolpin, 1999; Oettinger, 1999), others showed that grades improve with low work hours but fall with long hours (e.g. Schill et al., 1985; Lillydahl, 1990; Quirk et al., 2001), and still others failed to detect a causal relationship (e.g. Schoenhals et al., 1998; Warren et al., 2000; Dustmann et al., 2007).

Three recent studies of high school students, all of which used two-stage least squares (2SLS), illustrate the disparity of conclusions in this literature. Using state child labor laws as instruments in 1992 National Education Longitudinal Study data on high school seniors, Tyler

(2003) found a large negative effect of additional work hours on standardized test scores. Contradicting this, in annual 1991–2004 Monitoring the Future data on high school seniors, DeSimone (2006) specified components of the student unearned income distribution as instruments to uncover an inverse U-shaped relationship in which grades peak at 15 weekly work hours. Meanwhile, for National Longitudinal Survey of Youth 1997 (NLSY97) 10^{th}–12^{th} graders, Rothstein (2007) estimated that current and lagged work hours have small negative grade impacts that weaken when individual fixed effects are included and lose significance when instrumented using local wage and unemployment rates and state child labor laws.

For college students, the topic has received less attention but seems equally relevant. Many students work specifically to pay for tuition and coursework is presumably more difficult. Observed work propensities and intensities are high. In the 2001 Harvard College Alcohol Study (CAS), 62 percent of respondents reported working for pay in the previous month, and employed students work nearly 29 weekly hours on average. Yet as with high school students, previous research has not reached a consensus on how employment affects academic performance.

Four early studies treat work hours as exogenous. Among 836 students in his 1976–1979 introductory macroeconomics classes at Towson State University, Paul (1982) estimated that 10 additional work hours lowered exam scores by two percent. For 1,933 National Longitudinal Survey males who entered four-year colleges in fall 1972, Ehrenberg and Sherman (1987) found little impact of work hours on grades. Gleason (1993) depicted evidence of an inverse U-relationship in 1980s data: compared to unemployed students, grade point averages (GPAs) were 0.25 points higher for those working 1–10 weekly hours but 0.06 points lower for those working 31–40 weekly hours. Hood et al. (1992) similarly found that students working 7–14 hours per week had higher GPAs than those working less or more.

Two recent studies explicitly accounted for the potential endogeneity of hours worked. In 1989–1997 data on 2,372 first semester Berea College students, Stinebrickner and Stinebrickner (2003) use work-study job assignments to instrument for labor supply in a 2SLS model. An additional weekly work hour reduced first-semester GPA by 0.16 points. Kalenkoski and Pabilonia (2008) obtained an analogous negative effect of 0.017, nearly an order of magnitude smaller, using 1997–2004 NLSY97 data on 1,234 full-time, first semester four-year college students. Their three-equation system is estimated with maximum likelihood and specifies parental transfers, a quadratic in the net price of schooling, the state minimum wage, the county unemployment rate and a state work study program indicator as instruments for work hours. Of these, only parental transfers, which itself is endogenously determined, enters the work hours equation significantly. Identification thus occurs predominantly through the idiosyncratic functional form of the model.

This paper estimates the effect of paid employment on college student grades. Like recent studies, it uses an instrumental variable (IV) model to address prospective unobserved heterogeneity in the relationship between labor supply and academic performance. It contributes to the college-level literature by using 1993–2001 data from the CAS, which offers a much larger sample that includes students of all class standings. Compared with Stinebrickner and Stinebrickner (2003), the instruments, though not arising as naturally from a random assignment mechanism, are somewhat stronger. Also, the sample is nationally representative, rather than from a single school with a unique setting, and slightly more recent. Relative to Kalenkoski and Pabilonia (2008), the instruments have considerably more explanatory power for work hours, and the empirical strategy is more directly focused on identifying the impact of working on grades.

Besides the aforementioned data features, the main innovation of this study is its identification strategy. The IV approach used in most previous research exploits geographic differences in factors potentially affecting student work hours, such as child labor laws or unemployment rates. This tactic, which is infeasible here regardless because CAS data lack school location information, has serious limitations in both theory and practice. Theoretically, unobserved factors, such as attitudes or policies, affecting student achievement might vary over localities and be correlated with the instruments, thus threatening the instrument exogeneity assumption. Practically, most college students are too old to be affected by child labor laws, while unemployment rates tend to be weakly related to work hours (Ruhm, 1997; Oettinger, 1999; Rothstein, 2007; Kalenkoski and Pabilonia, 2008).

This study instead specifies as instruments variables representing human capital accumulation and preferences of the respondents' fathers. The maintained identification assumptions, therefore, are that paternal schooling attainment and emphasis are strongly related to student labor supply, yet otherwise unrelated to academic performance or its unobserved determinants. Next these assumptions are discussed in terms of the primary instrument, paternal schooling. Subsequently, reasons why the mechanism through which the secondary instrument, an indicator that the respondent was raised Jewish, affects student work hours and GPA is likely to be similar to that for paternal schooling is explained.

It seems reasonable to expect that paternal schooling has a negative impact on student labor supply. Fathers with higher attainment likely place a greater value on education, and in turn might provide more financial support to their college-enrolled children to allow them to spend less time earning money for tuition and living expenses and more time studying. Moreover, as a component of permanent family income (e.g. Heckman and Carneiro, 2003), paternal schooling should be positively related with student unearned income, which by the standard labor-leisure model negatively affects student labor supply. Empirically, it is easy to verify that these expectations manifest themselves in very large first stage instrument *F*-statistics.

The usefulness of this study, therefore, hinges critically on whether paternal schooling truly is exogenous with respect to student achievement. This assumption is supported by the traditional view of the intergenerational human capital transmission literature, which is that child schooling is much more closely related to maternal schooling than paternal schooling (e.g. Haveman and Wolfe, 1995; Chevalier et al., 2005). Presumably this stems from children spending more time with their mothers than their fathers (Black et al., 2005). Through assortative mating, controlling for maternal schooling might thus adequately capture any correlation between paternal schooling and unobserved student ability or preferences for academics that remains after accounting for endogenous student labor supply.[1]

Nonetheless, recent studies showing significant positive correlations between child schooling and paternal schooling, even holding constant maternal schooling (e.g. Behrman and Rosenzweig, 2002; Plug, 2004; Black et al., 2005; Chevalier et al., 2005; Björklund et al., 2006; Oreopoulos et al., 2006), might cast doubt on the validity of the paternal schooling exclusion restriction. It is important to recognize, though, that this study examines academic performance, not schooling. Paternal schooling might be a poor instrument for the latter because schooling is intergenerationally transmitted, yet have no direct relationship with the former, particularly taking into account its observed strong effect on student labor supply. This is more plausible because the empirical model holds constant not only schooling itself, i.e. years in college, but also maternal schooling, student age and the school attended. Among students within a specific postsecondary institution, of the same attainment and age, and with identical maternal schooling

[1] This suggests that maternal schooling, as a determinant of student achievement, is a poor candidate to instrument for student labor supply. If fathers on average earn higher incomes than mothers, paternal schooling might also be more strongly linked to their children's labor supply than maternal schooling, which is consistent with the CAS data.

and own labor supply, it is conceivable that paternal schooling has no separate relationship with student achievement.

The use of a second instrument is vital in allowing for empirical examination of the assumption that paternal schooling is not directly associated with grades. The other instrument used here is an indicator of whether the student was raised Jewish. Botticini & Eckstein (2005, 2007) outline how a religious norm requiring Jewish fathers to educate their sons, which has been operational since around the 3rd century, ultimately spurred entry into skilled occupations by the 9th century. Chiswick (1993) showed that controlling for demographic and skill differentials, including paternal schooling, American Jews in the 1973–1987 General Social Surveys had significantly higher levels of schooling, occupational status and earnings than other whites. Indeed, 3.5% of students in this study's analysis sample were raised Jewish, whereas the National Jewish Population Survey (NJPS; http://www.jewishvirtuallibrary.org/jsource/US-Israel/ujcpop.html) reported a U.S. Jewish population of 5.2 million, or 1.8% of the U.S. population, in 2000.[2] The NJPS further found that, compared to others in the U.S., Jews had higher educational attainment, rates of employment in management, business and professional/technical positions, and household incomes, lower fertility rates and incidence of poverty, and smaller households.

Consequently, the impact of being raised Jewish on work hours is expected to mimic that of paternal schooling, even with paternal schooling held constant. If so, compared with other students, including those with similarly-educated fathers, students raised Jewish will spend fewer hours working for pay in response to greater financial support from their fathers, who have better means of providing such support and also emphasize schooling and the eventual attainment of

[2] Although the religious norm outlined by Botticini & Eckstein (2007) pertained specifically to male offspring, the disproportionate presence of Jewish students in the CAS applies to both genders: 3.6% of sample males and 3.4% of sample females were raised Jewish.

skilled jobs. The logic for assuming that being raised Jewish has no separate correlation with academic achievement also parallels that for paternal schooling. Stronger preferences for schooling among Jewish families would suggest dubious validity of this exclusion restriction. However, it might be the case that Jewish students enroll in more selective colleges and are more likely to attain a specific level of post-secondary schooling by a certain age, but do not perform better than classmates from the same age/grade cohort. The empirical evidence strongly supports this latter hypothesis.

2. Empirical Strategy

To account for unobserved factors that might simultaneously influence GPA and work hours, IV is used to estimate the two-equation model (with variable subscripts for individual students suppressed)

$$(1) \qquad hours = \alpha_0 + \mathbf{z}\boldsymbol{\alpha}_1 + \mathbf{x}\boldsymbol{\alpha}_2 + u$$
$$(2) \qquad gpa = \beta_0 + \beta_1\, hours + \mathbf{x}\boldsymbol{\beta}_2 + v,$$

where *hours* denote time spent in paid labor per week and, as with *gpa*, pertains to the current academic year. Additionally, **x** is a set of factors related to both *hours* and *gpa*, **z** is a set of instruments correlated with *hours* but not otherwise with *gpa*, the α's and β's are coefficients and *u* and *v* are error terms assumed to be uncorrelated with **x** and **z**. Heteroskedasticity-robust standard errors are used in all specifications.

OLS estimation of equation 2 further requires that the error term *v* is uncorrelated with *hours* to produce a consistent estimator of β_1, the causal effect of an additional weekly work hour on GPA. However, unobserved student characteristics that affect school performance, such as preferences for paid or academic work, motivation, risk aversion and time preference, are subsumed into *v*. If these or other omitted determinants of *gpa* are also correlated with *hours*, the resulting nonzero correlation between *v* and *hours* renders the OLS estimator inconsistent, so that the estimated β_1 in part reflects spurious correlation between *hours* and *gpa*.[3]

The IV approach identifies a consistent β_1 if the instrument set $\mathbf{z}$ is highly correlated with hours, but uncorrelated with the error term *v*. Because the model is overidentified, i.e. there are multiple instruments, and standard errors are allowed to be heteroskedastic, generalized method of moments (GMM) is efficient relative to two stage least squares (2SLS) and is therefore used. Defining $\mathbf{Z}$ as the matrix $[\mathbf{z}, \mathbf{x}]$, *n* as the sample size, $\mathbf{\Omega}$ as the Newey-West covariance estimator allowing for arbitrary heteroskedasticity, and $\mathbf{W}^{-1}$ as a weighting matrix, GMM minimizes $J = n^{-1}\mathbf{vZW}^{-1}\mathbf{Z'v}$ using $\mathbf{W} = \mathbf{Z'\Omega Z}$. 2SLS, which yields very similar estimates and identical inferences, is the same estimator with $\mathbf{W} = \mathbf{Z'Z}$ (e.g. Greene, 2003, 201–207; 400–401).

Given the importance, yet theoretical ambiguity, regarding whether paternal attainment and emphasis of schooling is truly exogenous with respect to GPA, formal tests of the overidentifying restrictions are conducted. The test statistic is Hansen's *J* statistic, which is simply the minimized value of the GMM criterion above. Under the null of instrument exogeneity, the *J* statistic is distributed as chi-squared with degrees of freedom equal to one less than the number of instruments (typically more than two because of the way paternal schooling is measured, as explained below). This is analogous to the 2SLS *J* statistic, which is the product of the number of instruments and the *F* statistic for their joint significance in a regression of the second stage GPA equation residual on all the exogenous variables, i.e. $\mathbf{Z} = [\mathbf{z}, \mathbf{x}]$.

Intrinsically, the *J* statistic tests whether each of the two instruments, paternal schooling and being raised Jewish, yields statistically equivalent GMM estimates when used on its own to

[3] The structure of the data might also create selection bias, because dropouts are not observed. If dropout is correlated with both employment and GPA, even IV inconsistently estimates the population effect. However, in high school data, Rothstein (2007) found that correcting for selection on enrollment did not impact her estimates.

identify work hours. A danger is that the *J* statistic could be insignificant if both instruments are endogenous, but coincidentally produce similar results. To ensure this is not the case, models are also estimated in which one instrument is the sole identifier of work hours, while the other instrument is included in both the work hours and GPA equations. This allows for explicit testing of the direct relationship between GPA and each of the instruments in turn. High insignificance of *J* statistics and of each instrument when separately included in the GPA equation would provide strong reassurance that the exclusion restrictions are valid.

3. Data

As already mentioned, this study analyzes data from the CAS, which was administered to a nationally representative set of full-time four-year college students in the spring of 1993, 1997, 1999 and 2001. The main purpose of the CAS was to gather information on college student alcohol use. More relevant for this study, it also collected data on GPA, hours worked for wages, parental schooling and a variety of additional student characteristics.

Wechsler et al. (1994) provides details on sample selection and survey administration. Data are available only from institutions that had a student response rate above a cutoff value of slightly below 60 percent and either participated in the first three surveys or entered for the final survey. Of the 140 schools originally sampled, 128 took part in all three 1993–1999 waves and data are observed for 119 of these. In 2001, data are reported for 119 of the 120 participating schools, six of which are new to the survey. Each school chose a random sample, increasing in size with enrollment, by starting at a random point in the student registry and choosing every *r*th student. In 1993, questionnaires were mailed out in early February; 87 percent of completed surveys were returned by the end of March, with another 10 percent in April and the remainder by June, although this information is reported at the respondent level only in the latter two waves. The student response rate was 69 percent. Logistics were similar for subsequent administrations. Of the 55,169 students interviewed over the four years, 5,404 had missing values for analysis variables, 387 were graduate students, and 5,867 had ages outside the 18–24 year-old range in which exact year of age is observed (all but 115 of whom are older). After further excluding the 1,553 married students (who are added back in for one of the robustness checks), for whom work incentives might differ, the study sample contains 41,958 students.

The CAS records average hours worked per day over the previous month, which is assumed to accurately reflect labor supply for the entire academic year. In 1993, possible responses were integers from 0–7, with a remaining category of "8 or more" selected by seven percent of the sample. In subsequent years, choices ranged only from 0–4, with 22 percent of respondents reporting the residual category of "5 or more." A top code of nine hours is used for 1993 because it yields the closest mean, among values greater than eight rounded to the nearest 0.1, to that imputed by assuming normality in the upper tail. An analogous procedure generated a top code of 6.4 hours for 1997–2001. For consistency with previous research, the analysis uses weekly hours worked, which is simply the daily hours variable multiplied by seven with the resulting value of 44.8 for the top post-1993 category rounded up to 45.[4]

A four-point current academic year GPA variable is constructed from choices ranging in plus and minus increments from A to C–, along with D and "no grade or don't know." Slightly over one percent of students are removed from the sample because they select the final option. Figure 1 shows average GPA and labor supply by survey year. Panels a. and b. reveal that both rose consistently during the period, GPA from 3.09 to 3.22 and weekly hours worked from 16.8

[4] Possible attenuation bias from the measurement error inherent in observing only interval and top coded work hours during the past month further justifies the use of IV in lieu of OLS.

to 17.9. Panels c. and d. suggest that the upward trend in work hours is driven by increased employment, in the face of declining hours among workers, from 1993–1997 and 1999–2001, but by longer hours among workers from 1997–1999. Table 1 gives summary statistics for these four variables. Average GPA is between a B and B+. Three-fifths of respondents worked for pay, and those who do work an average of just over four hours per day.

Naturally, the correlation of 0.97 between the GPA and hours series in panel b. (albeit with just four data points) does not necessarily imply that working improves grades. For instance, grade inflation and the sustained economic expansion over the period might combine to spuriously correlate the two trends. In contrast, the unconditional sample correlation between GPA and work hours is –0.06, so students who work longer hours have slightly worse grades.

A further issue is that this negative linear correlation masks a nonlinear relationship between labor supply and grades. This is observed in figure 2, which depicts average GPA by each reported work hours value separately for 1993 (panel a.) and the combined subsequent years (panel b.) because the latter has fewer hours categories. In both time periods, GPA is nearly 0.2 points higher for students reporting the lowest hours category (one per day) than for those who do not work. GPA declines roughly linearly with additional hours (at least through seven hours per day in 1993), but does not fall to its level for non-workers until four hours per day, i.e. 28 hours per week. This suggests an inverse U-shaped relationship in which initial work hours increase school performance up to some low but non-trivial level of weekly hours, beyond which further hours hurt school performance. In fact, DeSimone (2006) estimated a relationship of this form, with a break-even point of about 30 weekly hours, for high school students. Nonlinearity is therefore investigated later as part of the robustness checks.

Table 2 lists the exogenous variables, starting with the instruments, along with their means in column 1. For parental schooling, only categorical information is available, and the categories differ between 1997 and the other years. In 1997, all that is observed is whether or not each parent attended a post-secondary institution, whereas obtaining a degree from a four-year college is also reported in the other surveys. Thus, the parental schooling instrument is in fact a set of three binary variables: a post-secondary attendance indicator interacted with a 1997 indicator, along with post-secondary attendance (without graduation) and four-year college graduation indicators each interacted with a "year other than 1997" indicator. In the three years that separately report attendance and graduation, 77.0 percent of fathers attended college and just over two-thirds of these (51.8 percent overall) graduated from four-year schools. In 1997, 68.8 percent of fathers attended college. The bottom of table 2 reports that in comparison, fewer mothers attended (63.1 percent in 1997, 72.7 percent in other years) and completed (42.4 percent) college.[5] Meanwhile, 3.5 percent of respondents were raised Jewish.[6]

Preliminary empirical support for the instrument exogeneity assumption comes from an ordered probit regression of a variable indicating whether the respondent considers academic work to be very important, important, somewhat important, or not at all important on the full set of exogenous factors. If paternal schooling attainment and preferences are correlated with school performance through shared attitudes regarding academic achievement, a correlation with how schoolwork is prioritized by the respondent should emerge. However, the paternal schooling indicators, both by themselves and with the raised Jewish indicator, are jointly insignificant, as is each of the four indicators individually, with all p-values above 0.2. In contrast, the analogous p-

[5] The lower prevalence of post-secondary attendance in 1997 is likely attributable to the phrasing of the question, because in years besides 1997 the relevant choice included "technical schooling beyond high school."
[6] As expected, paternal schooling is highly correlated with both maternal schooling and Judaism. For example, adding the full explanatory variable set to a regression of paternal college attendance on maternal college attendance raises the R-squared only from 0.159 to 0.197. In the latter model, the coefficient of maternal college attendance is 0.337 with a t-statistic of 77.8. That same regression further shows that fathers of Jewish-raised students are 8.2% more likely to have attended college than are fathers of students raised in no religion, with analogous effects of 12.7% for Muslim, 2.2% for Protestant, 1.5% for Catholic and 0.1% for the "other religion" category.

value for the maternal schooling indicators is 0.04. As previewed earlier, the following section offers more direct evidence of instrument validity.

As table 2 also documents, the exogenous variables specified to influence both work hours and grades include binary indicators for survey year, age, gender, year in school, race/ethnicity, health status, being raised Catholic, Muslim, Protestant or in another religion, and maternal schooling indicators defined analogously to the paternal schooling instruments. Regressions also control for college fixed effects, so that the effect of labor supply on academic performance is relative to other students of the same grade, age and gender at the same school.

4. Results

a. Exogenous factors

Column 2 of table 2 reports coefficients and *t*-statistics exogenous variables in the work hours equation, starting with the instruments (further specification tests are postponed to table 3). As predicted, paternal post-secondary schooling and being raised Jewish as opposed to without religion are negatively related to student labor supply. In the years besides 1997, respondents with fathers who attempted any post-secondary schooling work for pay an average of just under one hour less per week, while those with fathers who graduated from a four-year college spend nearly four fewer hours weekly in paid jobs. The latter is about one day's worth of work for the average employed respondent. Similarly, in 1997, students with fathers who attended college work over 3¼ hours less per week than do students with less-educated fathers. Respondents raised Jewish likewise allocate two fewer hours per week to paid employment. These effects are all significant at the one percent level.

Most of the other exogenous variables are also significant determinants of labor supply.

Hours increased over time, particularly by 1999, and with age and grade, even while holding the other constant. Females work more than males. Asians work less than non-Hispanic whites while other races work more, though Hispanic origin is unrelated. Health status is inversely related with paid labor time, except for the very few respondents in poor health. Catholics and those of faiths besides Judaism, Catholicism, Islam and Protestantism work more than do the non-religious, while Muslims work less to almost the same extent as Jews. As hypothesized, maternal schooling reduces labor supply, but by not nearly as much as does paternal schooling.

Column 3 of table 2 displays estimated GPA equation parameters. Grades have improved over time, as already observed, and not surprisingly fall with age holding grade level constant but rise with grade level holding age constant. School performance is highest for females, non-Hispanic whites and Muslims and Protestants and is lowest for blacks, and declines as health status worsens. GPA increases significantly, but only slightly, with maternal schooling.

b. Main results

Table 3 summarizes the main results of the study. Each column pertains to a separate specification, differentiated by whether the instrument set contains both paternal schooling and the Judaism indicator (the baseline model, in column 1), just the former (column 2) or just the latter (column 3). The upper panel reports this information along with instrument strength statistics, while the lower panel gives the OLS and GMM estimates as well as specification tests for the instrument exogeneity assumption.

In the upper panel of column 1, the joint *F*-statistic for the four instruments, each of which was highly significant in table 2, is more than eight times the commonly suggested threshold of 10. This provides assurance that, if the instruments are exogenous, IV bias will be miniscule relative to that from OLS. The partial *R*-squared reveals that the instruments explain a bit less than one percent of additional variation in work hours after controlling for the explanatory variables, including college fixed effects.

The lower panel of column 1 offers the main answer to the question posed by the paper. OLS shows a highly significant negative effect of an additional work hour on grades. In figure 2 this was suggested but not completely clear because of the apparent non-linearity at the employment margin. The magnitude of the OLS coefficient, however, is very small. For example, a 40-hour work week is predicted to reduce GPA by only 0.08 points, i.e. one-fourth of the way from one mark to the next on a plus/minus scale (e.g. from A– to B+). Meanwhile, the GMM estimate is also negative and highly significant, but over five times as large, so that the same 40-hour work week would lower GPA by 0.44, which equates to a full mark (as defined above) plus another one-third. Put differently, a one-standard deviation increase in labor supply of 18 hours per week would decrease GPA by less than 0.04 points according to OLS, but by 0.20 points according to GMM. Further discussion of the GMM effect size is deferred until after checking its robustness in alternative specifications.

The remainder of table 3 provides support for the validity of the exclusion restrictions. First, the last row in column 1 reports an overidentification test statistic *p*-value of above 0.99. This implies that using paternal schooling or the raised Jewish indicator alone to identify work hours would produce identical GMM estimates. The likelihood that the mechanisms through which each instrument works are overlapping, in that paternal schooling outcomes and preferences drive the impact of being Jewish on work hours, makes it possible that both instruments are truly endogenous with respect to school performance yet produce the same GMM coefficient. Yet, with the instruments each having a strong effect on work hours, it seems unlikely that such a process would result in an overidentification statistic of practically zero.

To further investigate the concern from the previous paragraph, column 2 identifies work hours with paternal schooling, while including the raised Jewish indicator in the GPA equation as well as the work hours equation. The upper panel shows that paternal schooling accounts for much of the additional variation explained by the instruments, leading to an even larger *F*-statistic than when the Judaism indicator is also included in the instrument set. In the lower panel, the GMM estimate is unchanged from column 1. In conjunction with this, the subsequent row shows that the raised Jewish indicator is a very highly insignificant correlate of school performance, with a coefficient of effectively zero and an accompanying *t*-statistic of 0.04. This comprises direct evidence that being raised Jewish affects grades only through its influence on labor supply, assuming that paternal schooling is similarly exogenous, which the overidentification test result continues to imply.[7]

Finally, column 3 reverses the roles that paternal schooling and being raised Jewish played in column 2, identifying work hours solely with the latter. As the lower panel depicts, the GMM point estimate remains unchanged, and the three paternal schooling indicators are jointly very highly insignificant in the GPA equation, again with a *p*-value of above 0.99.[8] Taken together, the evidence from table 3 considerably bolsters the case in defense of the study's identification strategy. It is difficult to imagine a scenario in which the instruments' substantial influence on work hours and complete lack of direct or implied correlation with grades, along with the stability of the GMM estimates across specifications, would be consistent with the

[7] In a strict sense, work hours is overidentified in column 2 because paternal schooling is represented by three distinct indicators. The structure of the indicator set, though, suggests the insignificant overidentification test says merely that the GMM estimate is no different in 1997 than in the other years, which table 5 later confirms. Still, temporal stability does provide some further support for the instrument exogeneity assumption.

[8] In the OLS model, the coefficients of three paternal schooling indicators are positive. Paternal college graduation (in years other than 1997) and attendance (in 1997) each increase GPA by 0.03 points, with *t*-statistics of 3.7 and 2.5, respectively, and the three indicators are jointly significant with a *p*-value of less than 0.0001. Thus, paternal schooling beyond high school does indeed have a small reduced form positive impact on achievement. However, the first stage estimates combined with the complete lack of significance of paternal schooling conditional on endogenous work hours implies that the entire effect occurs through devoting less time to paid labor.

exclusion restrictions being inappropriate because the instruments are not truly exogenous.

The column 3 specification also illustrates the oft-understated importance of instrument strength. Though being raised Jewish is sufficiently related with work hours to produce an *F*-statistic of nearly 19 (the square of the corresponding *t*-statistic in table 2), the GMM effect is insignificant at the 10 percent level. Relatively weak correlation between the instruments and work hours is a potential issue in some of the previously cited studies. For instance, in several of her specifications (e.g. tables 3 and 5 for males, table 7 for females), Rothstein (2007) estimates IV coefficients that are markedly larger than those from OLS, but insignificant despite instrument *F*–statistics between 10 and 15. Also, the instruments in Tyler (2003), with *F*–statistics of around 6, produce significant IV effects in his main specifications, but not in auxiliary models that include state fixed effects even though the estimates themselves are several times larger than those from OLS.

c. Robustness checks

Before discussing the implications of the previous subsection, table 4 shows estimates and test statistics for different specifications using the main analysis sample (rows a.–e.) as well as alternate samples in which the inclusion criteria are modified (rows f.–k.). For convenience, the top row re-displays the results from the baseline specification in column 1 of table 3.

In rows a.–c., models vary according to the way in which the paternal schooling instrument (and correspondingly maternal schooling) is defined, while continuing to use the raised Jewish indicator as an additional instrument. Row a. simply uses an indicator for paternal post-secondary attendance. This parsimonious specification yields a slightly larger GMM coefficient, but produces no other changes. Row b. allows the effect of the paternal college attendance instrument to vary by survey year, leaving results unchanged from row a. Row c.

returns to the baseline instrument set but allows the impacts of both paternal college attendance and completion to vary by year (other than 1997), generating an estimate similar to that of the baseline model. In sum, there seems little reason to either relax the constraint that paternal schooling has the same influence on work hours in different years, or impose a further constraint that only college attendance is used to reflect paternal schooling to be consistent across years.

Row d. alters the baseline model by making the institution fixed effects year-specific, to address the possibility that changes within schools over the survey period alter both student employment and grades. Results are virtually identical to those of the main specification.

To focus on the employment margin, row e. changes the dependent variable to a binary indicator of reporting positive work hours. The identification strategy continues to be effective, as the values of the first stage F-statistic and second stage overidentification test p-value are both large. The estimates parallel those for hours, but the discrepancy between OLS and GMM is considerably larger: while the GMM hours effects are 5–6 times larger than those of OLS, the GMM employment effect exceeds that of OLS by a factor of 30. Correspondingly, the OLS coefficient maintains significance at 5 percent but has a much smaller t-statistic than does the GMM coefficient.

Row f. estimates the effect of work hours only among employed students. A non-trivial component of the way paternal schooling and Judaism reduce labor supply is evidently by allowing students to avoid working altogether, because the first stage F-statistic is much smaller, though still well above 10. Likewise, the instrument exogeneity test p-value is not as large but still insignificant. The OLS and GMM estimates both are about twice as large as in the baseline model, remaining negative and highly significant. This is consistent with figure 2, which shows a steeper GPA/work hours gradient for workers than in the full sample including non-workers.

In fact, the estimates from rows e. and f. can be combined to investigate the nonlinearity apparent in figure 2. This is done in figure 3, which graphs the GMM estimate from the baseline model alongside an alternative estimate that allows for nonlinearity at zero hours. The latter is formed by assuming the row e. GMM coefficient holds at the sample average of 28.7 work hours, and then extrapolating in both directions from average work hours (to the sample minimum, among those employed, and maximum) using the row f. GMM estimate. Thus, the marginal effect on GPA is graphed as .023 x (28.7 – weekly hours) – .403. The main difference from the analogously-constructed OLS nonlinear estimate (not shown) is that the GMM marginal effect in figure 3 becomes negative at just over 11 work hours, whereas the OLS marginal effect does not become negative until 26 hours (similar to figure 2).

Comparing the two marginal effect paths in figure 3, the negative impact of working is overestimated by the linear model for students employed less than half-time and underestimated for students who work more. The linear model appears to very closely approximate the nonlinear model for the 53 percent of workers employed for 14, 21 or 28 weekly hours. Because the impact is close to zero (albeit in different directions) for the 10 percent of workers employed seven hours and few students report other specific hours categories, the only group of concern is the 28 percent of workers reporting the open-ended "five or more hours" per day category in the 1997–2001 surveys. These students are assumed to work 45 hours per week (as previously described), so suffer deleterious GPA impacts of employment equal to 0.50 points from the linear model, but 0.78 points from the nonlinear model. Even this difference, however, is less than one mark in a plus/minus system.

A more common way to model the nonlinearity in figure 2 is by adding a quadratic work hours term. This specification was also estimated, though the results are not reported in table 4.

For the quadratic term, all four instrument coefficients are negative and highly significant, with a joint *F*-statistic of 55.3. Similar to the linear model, the overidentification test is passed easily, with a *p*-value of 0.963. The quadratic hours term does enter positively, but the coefficient is very small and highly insignificant, with a *t*-statistic of 0.13. The linear term coefficient is virtually unchanged at –0.012, but its *t*-statistic falls to 0.92, although the two hours terms are jointly very highly significant. A literal inference is that the hypothesis of a quadratic effect is rejected in favor of a linear effect.[9] Overall, then, given the constraints of the data, the linear model appears to provide a reasonable approximation of the relationship between labor supply and academic performance among CAS students.

The remainder of table 4 shows results from samples that are altered either because an additional variable, which is unobserved for some respondents, is included (row g.), or because of other modifications to the inclusion criteria (rows h.–k.). Row g. controls for indicators corresponding to categories of the aforementioned measure of how important academic work is to the respondent. If these are related to both paternal schooling attainment/preferences and academic performance, their insertion should reduce the effects of both the instruments on labor supply, and of labor supply on grades. However, the instrument *F*-statistic and GMM work hours *t*-statistic decrease little, and the GMM work hours coefficient is unchanged. Rows h.–k. investigate, respectively, whether including ever-married students (along with indicators for being currently or formerly married), or excluding students who have been enrolled for more than four years, are the two oldest years of age or have GPAs below C, impacts the results. Column 2 of table 2 showed that these longer-enrolled and older respondents work considerably more than first-year and the youngest students, which is also true for the lowest-performing and

[9] Though strongly related to both linear and squared hours, the instruments are all binary indicators and reflect only two distinct underlying quantities, and therefore might vary insufficiently to separately identify the two terms.

especially the currently and formerly married. Results change little, though. The GMM estimate diminishes slightly when the worst students are omitted, which might be expected with the grade distribution having a higher lower bound, and increases slightly in the other cases.

d. Interpretation

Tables 3 and 4 conclude that the effect on GPA of an additional paid work hour each week, estimated by GMM, is between –0.010 and –0.012, with the midpoint of –0.011 also being the main estimate from the baseline model. An obvious question is whether the magnitude of this estimate is reasonable. Specifically, is it realistic that a 40-hour work week reduces GPA by 0.44 points? Intuitively, it seems plausible that full-time employment, which constitutes nearly 25 percent of available hours not counting eating, sleeping or attending class, would lower GPA by just over one mark relative to not working at all. Comparatively, this effect is less than one-tenth the size of that obtained by Stinebrickner and Stinebrickner (2003) using 2SLS, –0.16. If the size of their estimate is believable, the GMM estimate in this study is more credible than the corresponding OLS estimate, which is almost two orders of magnitude smaller than their estimate above. The estimate here is fairly comparable with that obtained by Kalenkoski and Pabilonia (2008), although it is smaller by about one-third. Both of these other studies examine only first semester students, though, and table 5 later indicates that the estimate for CAS freshmen is slightly smaller than that from the full sample.

In contrast, the central GMM estimate is 5–6 times larger than the uniform OLS estimate of –0.002. Relative magnitudes of this type between IV and OLS are not uncommon, and furthermore mimic the findings of studies such as Tyler (2003) and Stinebrickner and Stinebrickner (2003). OLS estimates in the former are six to nine times smaller than associated 2SLS estimates, and in the latter are about one-fourth the size and have the opposite sign.

One interpretation is that academically superior students work more, producing a positive spurious correlation between paid labor and school performance that, when using OLS, obscures the negative causal impact of employment on grades. As Stinebrickner and Stinebrickner (2003) outline, more motivated students might try to impress supervisors, who might in response assign such students more hours. Or, better students might choose to work additional hours as a substitute for leisure activities from which they get less stimulation. These hypotheses, however, conflict with results from the first stage regression for the model in row g. of table 4, in which students who rate academic work as important, somewhat important and not important work a statistically significant 1.4, 2.4 and 3.5 more weekly hours, respectively, than those who rate academic work as very important.

Alternatively, at least part of the difference between GMM and OLS could stem from measurement error that biases the latter towards zero. Two observations suggest measurement error is salient. First, the inverse U-shape relationship between work hours and GPA found in DeSimone (2006) has similar turning points for both OLS and IV, with both level and quadratic coefficients much smaller with OLS. Ignoring measurement error, this would suggest that the unobserved heterogeneity biasing OLS operates in different directions for students working low and high amounts (divided by a threshold of about 30 hours), which seems counterintuitive. Second, Stinebrickner and Stinebrickner (2003) estimate positive OLS effects, meaning that OLS bias is attributable to something beyond measurement error, but use administrative labor supply data that are much more precise than survey data. This contrasts with Tyler (2003) and DeSimone (2006), which used information recorded in five hour per week intervals and estimate proportionately much larger differences between OLS and IV. The coding of work hours in daily interval amounts makes the data for this study even more susceptible to measurement error that would attenuate OLS coefficients.

A third possible explanation is that there is heterogeneity in the effects of both the instruments on working and working on academic performance. If so, the GMM estimate is properly interpreted as the local average treatment effect among students for whom the instruments, at the margin, influence work hours (Imbens and Angrist, 1994). It could be that students for whom parental schooling achievement and attitudes affect labor supply experience relatively large GPA improvements from reducing time in paid work. In that case, the true average causal effect of employment on grades is less than that implied by the GMM estimate. But this possibility seems contradicted by the much smaller size of the GMM estimate in this study compared with those of the two recent studies of college students, as discussed above.

e. Stratified samples

Table 5 lists estimates from stratified samples, using the baseline specification of table 3, column 1. Not counting the auxiliary models in the last three rows of panel a., each panel divides the sample into mutually exclusive and exhaustive groups.

The top four rows of panel a. affirm that the effect of working on grades is quite stable across survey years. The overidentification test result is troublesome for 2001, but the GMM estimate changes little when data from that year are combined with those from 1999 (the previous survey year), when 1993 data are further added (thus pooling the three years in which parental schooling is consistently defined), or when 2001 data are omitted. Thus, although grades and work hours have both risen over time, their relationship appears unaltered.

Estimates are similar across genders (panel b.), year in school (d.), age (e.), and rating of academic work importance (g.). As noted earlier, the effect for first-year students, the group studied by both Stinebrickner and Stinebrickner (2003) and Kalenkoski and Pabilonia (2008), is slightly smaller than that in the full sample, which itself is smaller than what those two studies yield. Panel g. provides further evidence that failure to observe underlying differences in attitudes about schoolwork does not impart bias.

In panel c., the deleterious impact of working is substantially larger for non-white or Hispanic students. At average work hours, in fact, employment would explain roughly four times the unconditional GPA difference between the two groups. The relatively low first stage IV *F*-statistic gives some pause, but this would suggest bias towards OLS, not away from zero.

Panel f. shows that negative employment effects are largest for those who report being in the best health. It could be that this is why the healthiest students work less than others, which is a small part of the reason for their higher grades. Similarly, Jews & Muslims, who in panel h. are most hurt academically by spending time in paid jobs, were documented earlier to be the students who worked the least, holding constant other factors.[10] Or, being the best academic performers might explain why the opportunity cost of work time, in terms of achievement foregone, is higher for the healthiest respondents. This might also contribute to why students with mothers who attended college experience more harm from working than others in panel i., though the small difference by academic importance rating in panel g. belies this logic.

For panel j., the only paternal schooling instrument that is identified is graduating from college in years other than 1997. Thus, the group with more paternal schooling has only this variable and being raised Jewish as instrument, while the group with less paternal schooling must rely on the Jewish indicator as the only instrument. The very low *F*–statistic for the latter category supports the conjecture that the interaction of a Jewish upbringing with employment and grades among college students operates in conjunction with paternal schooling preferences

---

[10] The panel h. models must exclude the raised Jewish indicator as an instrument since it is defined relative to the non-religious, who are stratified into their own group.

and outcomes. In particular, only 1.0% of students with fathers who did not attend college are Jewish, compared to 4.3% of those with college-attending fathers. Thus, the large impact of working on GPA in that group is not statistically meaningful or necessarily reliable in any case.

5. Conclusion

The contribution of this research is twofold. On a basic level, it is the only study to estimate the effect of paid employment on academic performance among a nationally representative set of four-year college students from all years in school using a methodology designed explicitly to deal with the likely endogeneity of labor supply. More specifically, the IV strategy is intuitive and straightforward, yet demonstrably credible and robust to a variety of specifications. The main result is that each additional weekly work hour reduces academic year GPA by 0.011 points. Thus, a 30-hour work week lowers the average grade by one mark, i.e. from A– to B+, compared with not participating in the labor market at all.

These results are consistent with what some college instructors regularly experience: students who blame class tardiness and absence, failure to submit assignments and poor exam performance on their employment obligations. However, the findings of this study suggest that the negative relationship between labor supply and grades is not simply attributable to less academically motivated students working long hours. In that case, the aforementioned hypothetical lackluster students would not necessarily perform better academically if they were prevented from working, which is simply an activity to which bad students devote more time than good students. Instead, students who spend longer hours in paid labor because of preferences or budget constraints related to their fathers' schooling attainment and attitudes ultimately perform worse in school than they otherwise would.

Whether this issue is relevant for government policymakers and academic administrators depends on whether student labor supply choices are myopic. If students undervalue future benefits to society accruing from grade points sacrificed to attain additional current income, actions to limit hours that college students are allowed to spend in paid jobs while enrolled in classes might be warranted. However, it is unclear who would regulate and monitor such restrictions. Furthermore, rising real college costs will seemingly put added pressure on students to earn while they learn. Distributing more tuition revenue as financial aid to students or implementing programs like Georgia's HOPE scholarship would presumably lessen such pressure, but would require some combination of alternative funding sources or cutting other potentially beneficial school or public programs.

References

Behrman, Jere R. and Mark R. Rosenzweig, "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?" *American Economic Review*, March 2002, 92(1), 323–334.

Björklund, Anders, Mikael Lindahl and Eric Plug, "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data," *Quarterly Journal of Economics*, 121(3), August 2006, 999–1028.

Black, Sandra E., Paul J. Devereux and Kjell G. Salvanes, "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital," *American Economic Review*, March 2005, 95(1), 437–449.

Botticini, Maristella and Zvi Eckstein, "From Farmers to Merchants, Conversions and Diaspora: Human Capital and Jewish History," *Journal of the European Economic Association*, 5(5), September 2007, 885–926.

Botticini, Maristella and Zvi Eckstein, "Jewish Occupational Selection: Education, Restrictions or Minorities?" *Journal of Economic History*, 65(4), December 2005, 922–948.

Chevalier, Arnaud, Colm Harmon, Vincent O'Sullivan and Ian Walker, "The Impact of Parental Income and Education on the Schooling of their Children," Institute for Fiscal Studies Working Paper 05/05, January 2005.

Chiswick, Barry R., "The Skills and Economic Status of American Jewry: Trends over the Last Half Century," *Journal of Labor Economics*, 11(1), Part 1, January 1993, 229–242.

DeSimone, Jeff, "Academic Performance and Part-time Employment among High School Seniors," *Topics in Economic Analysis & Policy*, 2006, 6(1), Article 10.

Dustmann, Christian and Arthur van Soest, "Part-time Work, School Success and School Leaving," *Empirical Economics*, 32(2-3), May 2007, 277–299.

Eckstein, Zvi, and Kenneth I. Wolpin, "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities," *Econometrica*, 67(6), November 1999, 1295–1339.

Ehrenberg, Ronald G., and Daniel R. Sherman, "Employment While in College, Academic Achievement, and Postcollege Outcomes: A Summary of Results," *Journal of Human Resources*, 22(1), Winter 1987, 1–23.

Gleason, Philip M., "College Student Employment, Academic Progress, and Postcollege Labor Market Success," *Journal of Student Financial Aid*, 23(2), Spring 1993, 5–14.

Greene, William H., Econometric Analysis, 5e, Prentice Hall, Upper Saddle River, NJ, 2003.

Haveman, Robert and Barbara Wolfe, "The Determinants of Children Attainments: A Review of Methods and Findings," *Journal of Economic Literature*, 33(4), December 1995, 1829–1878.

Heckman, James and Pedro Carneiro, "Human Capital Policy," National Bureau of Economic Research Working Paper 9495, 2003.

Hood, Albert B., Andrew F. Craig and Bruce W. Ferguson, "The Impact of Athletics, Part-time Employment, and Other Academic Activities on Academic Achievement," *Journal of College Student Development*, 33, September 1992, 447–453.

Imbens, Guido W. and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62(2), March 1994, 467–475.

Jones, Ethel and John D. Jackson, "College Grades and Labor Market Rewards," *Journal of Human Resources*, 25, Spring 1990, 253–266.

Kalenkoski, Charlene Marie, and Sabrina Wulff Pabilonia, "Parental Transfers, Student Achievement, and the Labor Supply of College Students," BLS Office of Productivity and Technology Working Paper 416, March 2008.

Lillydahl, Jane H., "Academic Achievement and Part-Time Employment of High School Students," *Journal of Economic Education*, 21(3), Summer 1990, 307–316.

Loury, Linda Datcher and David Garman, "College Selectivity and Earnings," *Journal of Labor Economics*, 13(2), April 1995, 289–308.

Oettinger, Gerald S., "Does High School Employment Affect High School Academic Performance?" *Industrial and Labor Relations Review*, 53(1), October 1999, 136–151.

Oreopoulos, Philip, Marianne E. Page and Ann Huff Stevens, "The Intergenerational Effects of Compulsory Schooling," *Journal of Labor Economics*, October 2006, 24(4), 729–760.

Paul, Harvey, "The Impact of Outside Employment on Student Achievement in Macroeconomic Principles," *Journal of Economic Education*, 13(2), Summer 1982, 51–56.

Plug, Erik, "Estimating the Effect of Mother's Schooling on Children's Schooling using a Sample of Adoptees," *American Economic Review*, 94(1), March 2004, 358–368.

Quirk, Kimberly J., Timothy Z. Keith and Jeffrey T. Quirk, "Employment during High School and Student Achievement: Longitudinal Analysis of National Data," *Journal of Educational Research*, September/October 2001, 95(1), 4–10.

Rothstein, Donna, "High School Employment and Youths' Academic Achievement," *Journal of Human Resources*, 42(1), Winter 2007, 194–213.

Ruhm, Christopher J., "Is High School Employment Consumption or Investment?" *Journal of Labor Economics*, 15(4), October 1997, 735–776.

Schill, William J., Rosemarie McCartin and Katrina Meyer, “Youth Employment: Its Relationship to Academic and Family Variables,” *Journal of Vocational Behavior*, 26(2), April 1985, 155–163.

Schoenhals, Mark, Marta Tienda and Barbara Schneider, “The Educational and Personal Consequences of Adolescent Employment,” *Social Forces*, 77(2), December 1998, 723–761.

Singh, Kusum, “Part-time Employment in High School and Its Effect on Academic Achievement, *Journal of Educational Research,* 91(3), January/February 1998, 131–139.

Stinebrickner, Ralph, and Todd R. Stinebrickner, “Working During School and Academic Performance,” *Journal of Labor Economics*, 21(2), April 2003, 473–491.

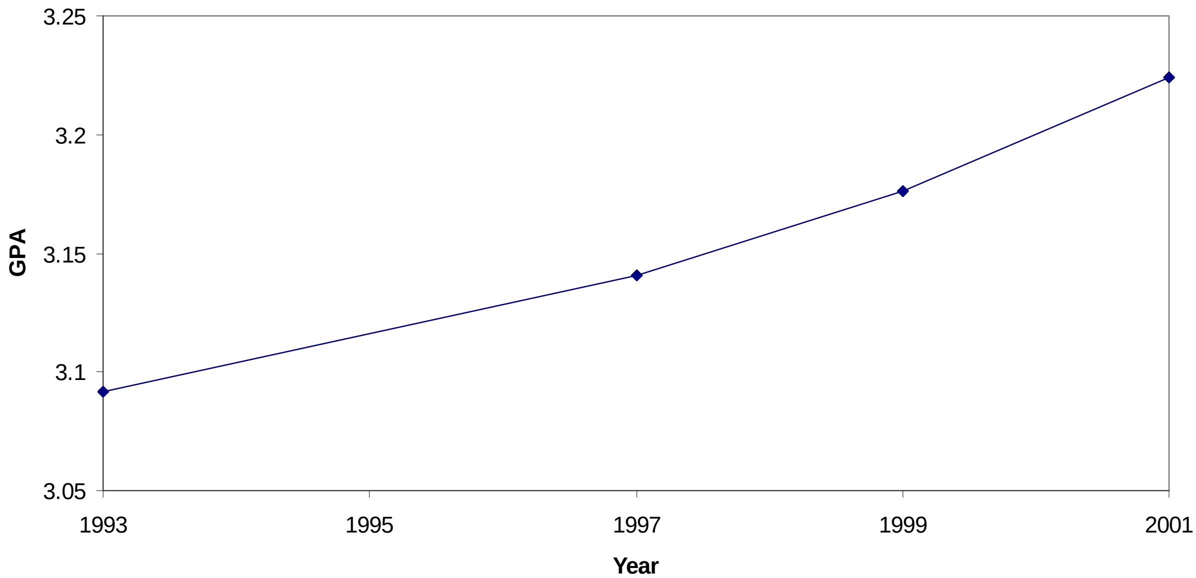
Tyler, John H., “Using State Child Labor Laws to Identify the Effect of School-Year Work on High School Achievement,” *Journal of Labor Economics*, 21(2), April 2003, 381–408.

Warren, John Robert, Paul C. LePore and Robert D. Mare, “Employment during High School: Consequences for Students’ Grades in Academic Courses,” *American Educational Research Journal*, 37(4), Winter 2000, 943–969.

Wechsler, Henry, Andrea Davenport, George Dowdall, Barbara Moeykens and Sonia Castillo, “Health and Behavioral Consequences of Binge Drinking in College,” *JAMA*, 7 December 1994, 272(21), 1672–1677.

Figure 1: GPA and paid work, 1993–2001

a. GPA



b. Weekly hours worked (including non-workers)

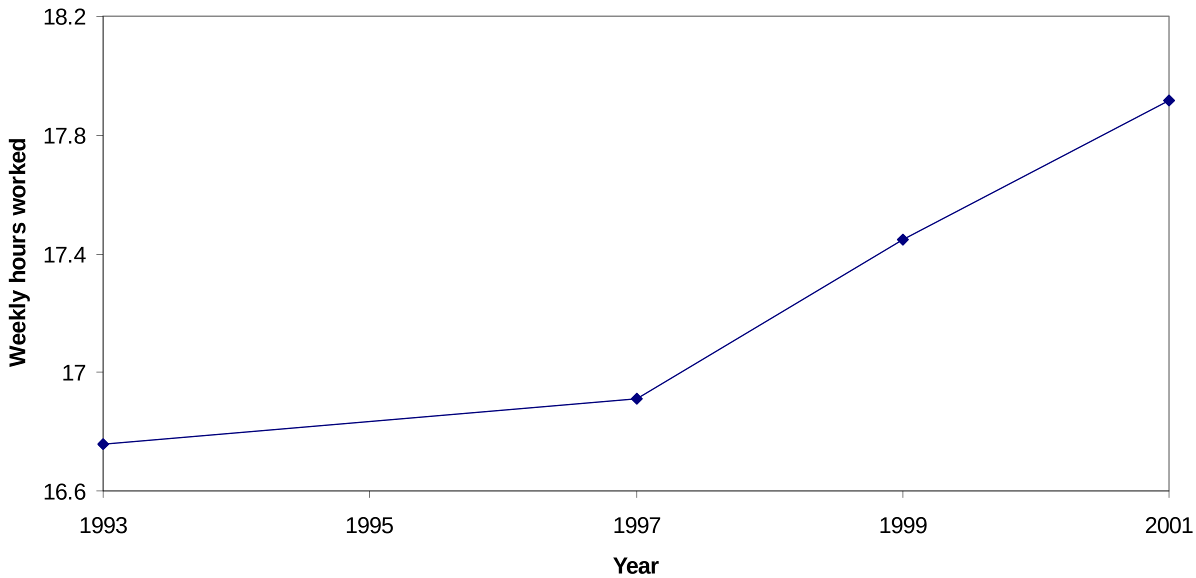
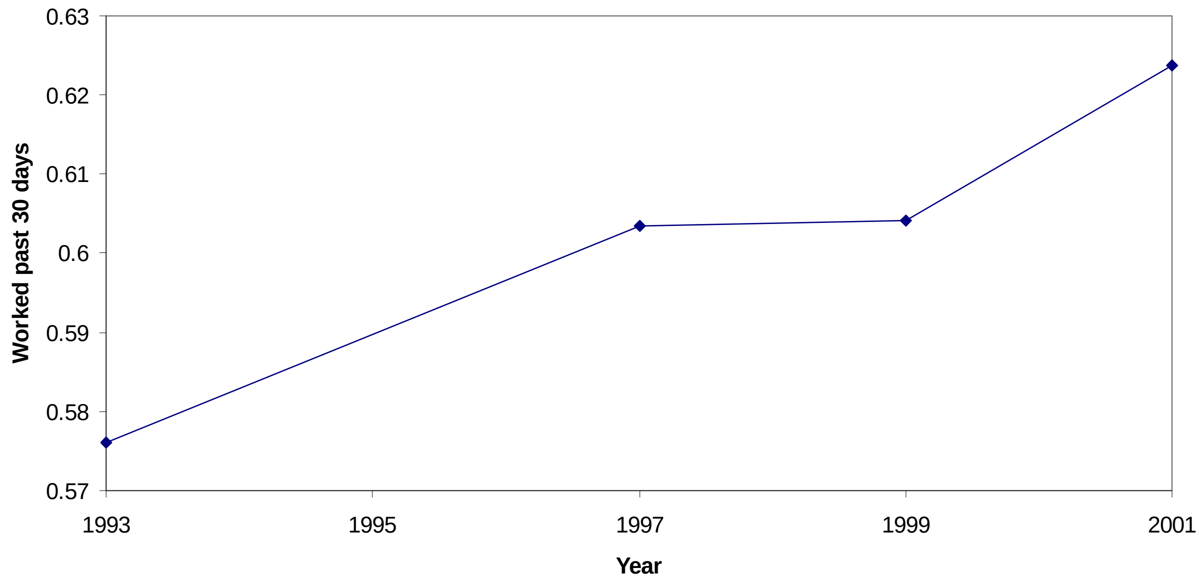


Figure 1 (continued): GPA and paid work, 1993–2001

c. Worked for pay in past 30 days



d. Hours worked per day by students who worked

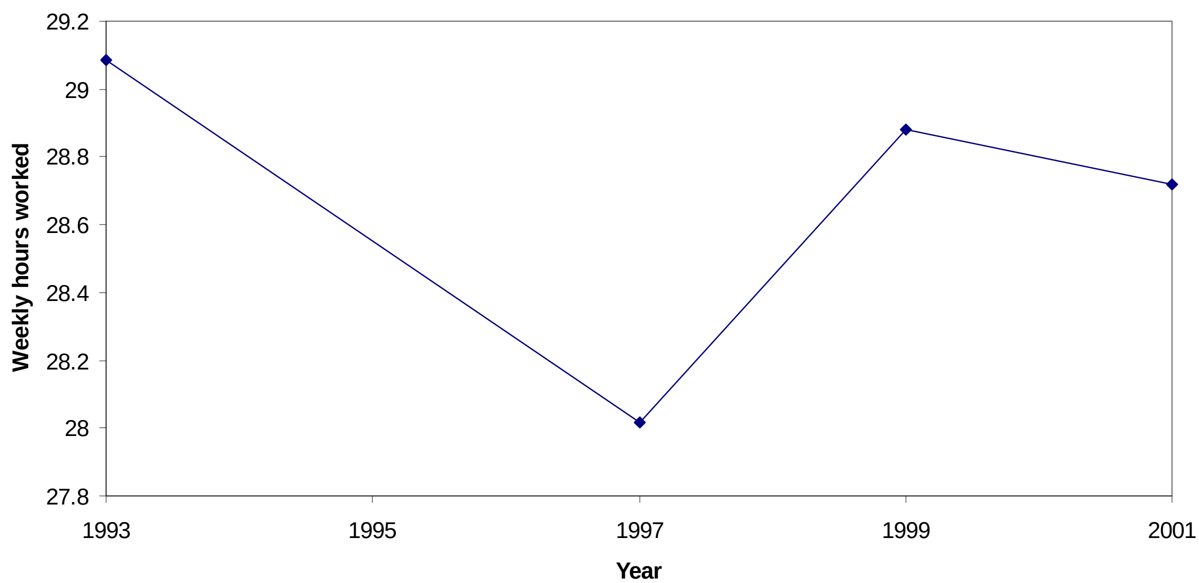
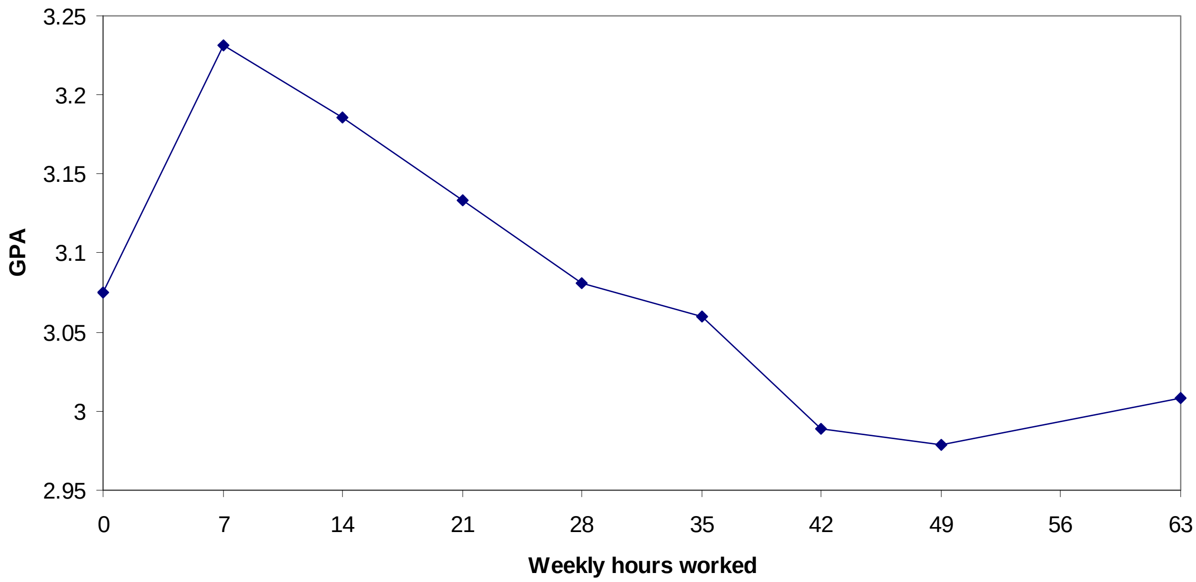


Figure 2: Mean GPA by work hours

a. 1993



b. 1997–2001

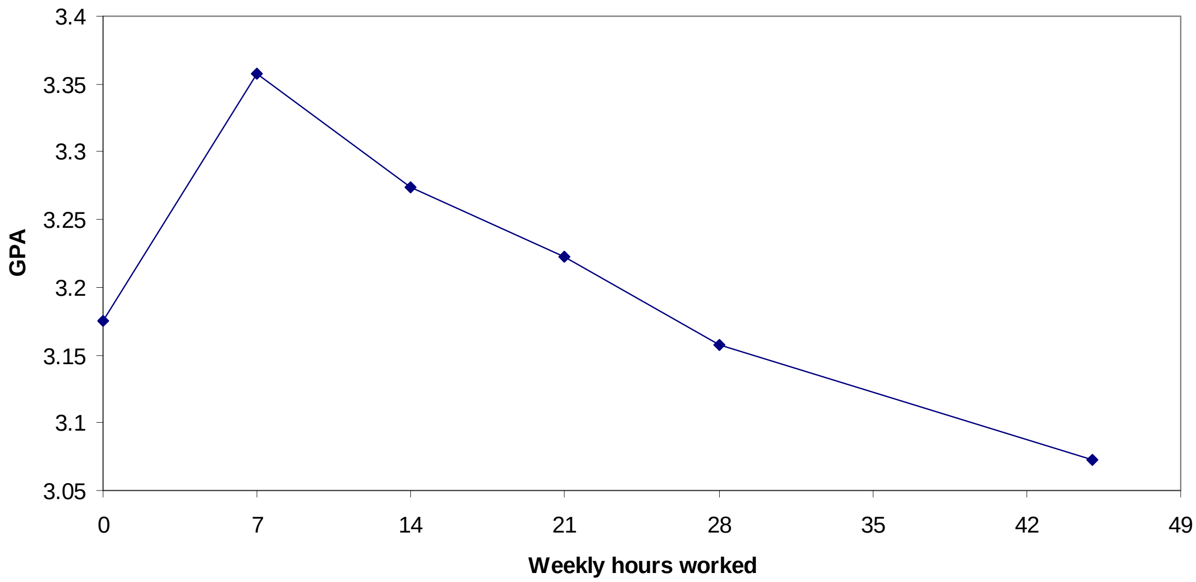
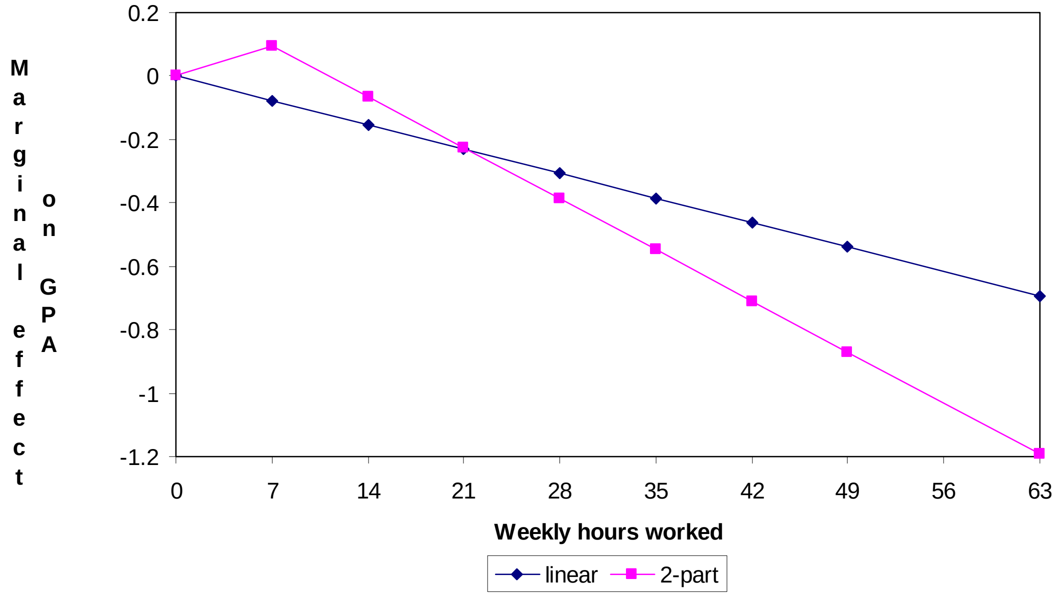


Figure 3: IV linear and two-part model marginal effects



The linear model effect is the GMM estimate from table 3, column 1 (or equivalently the top row of table 4) multiplied by weekly hours worked. For workers, the two-part model effect is calculated from the GMM estimates in rows e. and f. of table 4 as –.403 + (28.7 – weekly hours) (.023). This assumes the marginal effect of working is the effect of the average weekly hours for workers (i.e. 28.7 hours) and adds to that the product of the conditional hours effect and the deviation from average hours to construct the marginal effect at positive observed values of work hours.

Table 1: Means of dependent variables

Variable	Mean
GPA in past year	3.15 (0.58)
Hours worked per week in past 30 days	17.2 (18.1)
Worked in past 30 days	.600
Hours worked per week in past 30 days (if worked)	28.7 (14.8)

The sample size is 25,183 for hours worked among workers and 41,958 for all other variables. Parentheses contain standard deviations for non-binary variables.

Table 2: Means & coefficient estimates for exogenous variables

	Mean	Coefficient in model from table 3, column 1	
		Hours (1^{st} stage)	GPA (2^{nd} stage)
	(1)	(2)	(3)
Father attended college, did not graduate x (not 1997)	.187	–0.89 (2.98)	
Father graduated from college x (not 1997)	.383	–3.95 (13.6)	
Father attended college x 1997	.179	–3.29 (8.62)	
Raised Jewish	.035	–2.09 (4.35)	
1997	.260	0.37 (0.89)	0.08 (6.28)
1999	.256	1.01 (4.36)	0.10 (13.0)
2001	.206	1.18 (4.71)	0.18 (20.2)
19 years old	.227	0.12 (0.39)	–0.02 (1.35)
20 years old	.218	1.32 (3.22)	–0.02 (1.50)
21 years old	.212	1.63 (3.41)	–0.08 (4.62)
22 years old	.134	3.12 (5.84)	–0.08 (3.95)
23 years old	.058	4.26 (6.70)	–0.08 (3.70)
24 years old	.028	4.57 (6.33)	–0.06 (2.16)
2^{nd} year	.231	2.77 (9.29)	0.13 (10.6)
3^{rd} year	.243	3.81 (9.63)	0.23 (15.1)
4^{th} year	.213	4.74 (10.3)	0.35 (19.6)
5^{th} year & beyond	.065	5.61 (9.41)	0.31 (13.9)
Female	.605	2.90 (16.7)	0.14 (17.5)
Black	.047	1.57 (3.38)	–0.33 (21.2)
Asian	.073	–3.69 (10.9)	–0.07 (4.97)
Native American	.038	1.15 (2.17)	–0.07 (3.99)
Non-white, black, Asian or Native American	.034	1.14 (1.97)	–0.06 (3.27)
Hispanic	.064	–0.53 (1.11)	–0.05 (3.30)
Very good health	.439	0.53 (2.60)	–0.06 (8.88)
Good health	.253	1.33 (5.67)	–0.16 (19.7)
Fair health	.046	1.40 (3.23)	–0.24 (15.5)
Poor health	.005	–0.74 (0.61)	–0.27 (6.32)
Raised Catholic	.374	0.64 (2.37)	–0.01 (1.60)
Raised Muslim	.008	–1.70 (1.68)	0.08 (2.49)
Raised Protestant	.365	0.31 (1.14)	0.04 (5.01)
Raised in other religion	.087	1.06 (2.82)	–0.01 (0.43)
Mother attended college, did not graduate x (not 1997)	.224	–0.12 (0.43)	–0.00 (0.00)
Mother graduated from college x (not 1997)	.314	–1.40 (5.00)	0.03 (2.75)
Mother attended college x 1997	.164	–0.99 (2.74)	0.02 (1.99)

The sample size is 41,958. Parentheses contain absolute values of heteroskedasticity-robust *t*-statistics. Omitted categories are 1993, 18 years old, male, freshman, white, non-Hispanic, excellent health, no religion, and mother did not attend college. Regressions also control for college fixed effects.

Table 3: Effects of hours worked on GPA

	(1)	(2)	(3)
A. 1st stage OLS equation: Dependent variable = Hours worked			
Instruments:			
Paternal schooling	Yes	Yes	No
Raised Jewish	Yes	No	Yes
F-statistic (instruments)	82.3	101.3	18.9
Partial *R*-squared	.0081	.0076	.0004
B. 2nd stage equation: Dependent variable = GPA			
Hours worked/week – OLS	–.002 (13.9)	–.002 (13.8)	–.002 (13.4)
Hours worked/week – GMM	–.011 (6.01)	–.011 (5.79)	–.011 (1.40)
Raised Jewish		–.001 (0.04)	
χ^2 statistic (paternal schooling)			0.09 [.993]
χ^2 statistic (overidentification)	0.09 [.993]	0.09 [.956]	

Parentheses contain absolute values of heteroskedasticity-robust *t*-statistics; brackets contain *p*-values for overidentification test statistics. Regressions also include indicators for gender, age, class, race, Hispanic ethnicity, health status, religious affiliation, college, year and maternal education.

Table 4: Effects of hours worked on GPA in modified specifications & samples

	1^{st} stage F-stat. (1)	OLS estimate (2)	GMM estimate (3)	Overid. χ^2 stat. (4)
Main specification (table 3, column 1)	82.3	–.002	–.011	0.09
		(13.9)	(6.01)	[.993]
Alternative specifications (n = 41,958):				
a. 2 IV: Jewish, Father attended college	102.7	–.002	–.012	0.00
		(14.2)	(5.10)	[.996]
b. 5 IV: Jewish, Father attended by year	42.5	–.002	–.012	1.46
		(14.2)	(5.04)	[.834]
c. 8 IV: Jewish, Father attended/completed by year	41.8	–.002	–.011	2.91
		(13.9)	(6.08)	[.893]
d. College-by-year effects included	79.6	–.002	–.011	0.10
		(13.6)	(5.84)	[.992]
e. Dependent variable = worked (binary)	82.0	–.013	–.403	1.54
		(2.29)	(5.80)	[.674]
Alternative samples:				
f. Non-workers excluded	18.9	–.005	–.023	3.46
(n = 25,183)		(18.9)	(4.56)	[.326]
g. Academic importance rating included	80.8	–.002	–.011	0.35
(n = 41,703)		(12.3)	(5.98)	[.951]
h. Ever married respondents included	81.5	–.002	–.012	0.17
(n = 43,511)		(14.4)	(6.41)	[.982]
i. 5^{th} year & beyond enrollees excluded	80.0	–.002	–.012	0.46
(n = 39,216)		(13.8)	(6.28)	[.928]
j. 23 & 24 year olds excluded	78.9	–.002	–.012	0.49
(n = 38,345)		(14.0)	(6.30)	[.921]
k. C– & below GPAs excluded	81.0	–.002	–.010	2.08
(n = 41,267)		(13.4)	(5.93)	[.556]

Parentheses contain absolute values of heteroskedasticity-robust t-statistics; brackets contain p-values for over-identification test statistics. Regressions also include indicators for gender, age, class, race, Hispanic ethnicity, health status, religious affiliation, maternal education level, college and year. In a.–c., the raised Jewish indicator remains an instrument, while the paternal education instruments are a single indicator of father attending college (a.), year-specific indicators of father attending college (b.), and the father attended variable for 1997 along with year-specific indicators for attending and completing in the other years (c.). Maternal schooling is defined analogously to paternal schooling in these models.

Table 5: Effects of hours worked on GPA in stratified samples

		Sample size (1)	Mean Hrs. (2)	Mean GPA (3)	*F* for IV (4)	GMM estimate (5)	Overid. χ^2 statistic (6)
a.	1993	11,689	16.8	3.09	23.2	–.012 (2.99)	1.62 [.444]
	1997	10,903	16.9	3.14	43.2	–.011 (3.06)	1.27 [.259]
	1999	10,735	17.4	3.18	30.5	–.011 (3.05)	0.73 [.695]
	2001	8,631	17.9	3.22	25.8	–.011 (2.83)	5.50 [.064]
	1999, 2001	19,366	17.7	3.20	56.3	–.011 (4.12)	4.00 [.136]
	1993, 1999, 2001	31,055	17.3	3.16	80.7	–.011 (5.15)	0.61 [.738]
	1993, 1997, 1999	33,327	17.0	3.13	62.6	–.012 (5.43)	0.71 [.871]
b.	Females	25,368	15.2	3.20	61.5	–.010 (4.77)	1.81 [.612]
	Males	16,590	18.5	3.08	23.8	–.014 (3.80)	0.91 [.823]
c.	White non-Hispanic	33,205	17.0	3.18	81.6	–.007 (3.69)	2.20 [.532]
	Non-white and/or Hispanic	8,753	18.0	3.05	5.6	–.038 (3.78)	1.61 [.657]
d.	First year in school	10,416	12.9	3.06	33.1	–.009 (2.75)	3.86 [.277]
	Second or third year in school	19,881	17.4	3.16	37.1	–.013 (4.79)	2.03 [.566]
	Fourth year or beyond in school	11,661	20.7	3.24	16.8	–.009 (2.54)	5.02 [.170]
e.	Age 18–20	23,855	14.8	3.13	53.9	–.011 (4.59)	1.23 [.746]
	Age 21–24	18,103	20.3	3.19	29.7	–.011 (3.91)	1.54 [.674]
f.	Excellent health	10,747	16.1	3.24	21.4	–.019 (4.88)	1.48 [.687]
	Very good health	18,432	17.0	3.18	41.6	–.007 (2.58)	0.87 [.834]
	Good, fair or poor health	12,779	18.5	3.05	22.9	–.011 (2.98)	2.35 [.503]
g.	Academics very important	31,095	16.6	3.24	64.3	–.011 (5.41)	1.42 [.702]
	Academics not very important	10,863	18.8	2.91	21.5	–.008 (2.27)	3.35 [.340]
h.	Raised nonreligious	5,470	16.5	3.15	18.1	–.006 (1.23)	1.41 [.493]
	Raised Catholic	15,695	18.1	3.12	37.1	–.006 (1.88)	2.16 [.339]
	Raised Protestant	15,335	16.9	3.18	34.4	–.014 (4.19)	0.41 [.817]
	Raised Jewish, Muslim or other	5,458	16.3	3.15	13.1	–.020 (3.47)	1.69 [.429]
i.	Mother attended college	29,462	16.0	3.18	65.4	–.014 (6.34)	1.82 [.611]
	Mother did not attend college	12,496	20.1	3.09	17.5	–.006 (1.45)	5.47 [.141]
j.	Father attended college	31,437	15.8	3.18	74.4	–.012 (4.21)	0.00 [.950]
	Father did not attend college	10,521	21.3	3.08	1.1	–.036 (0.85)	

Parentheses contain absolute values of heteroskedasticity-robust *t*-statistics; brackets contain *p*-values for overidentification test statistics. Regressions also include indicators for gender, age, class, race, Hispanic ethnicity, marital status, health status, religious affiliation, maternal education level, college and year.