2006

Adolescent alcohol use and educational outcomes

Wesley A. Austin
University of South Florida

Adolescent Alcohol Use and Educational Outcomes

by

Wesley A. Austin

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Economics
College of Business Administration
University of South Florida

Co-Major Professor: Jeffrey DeSimone, Ph.D.
Co-Major Professor: Gabriel Picone, Ph.D.
Don Bellante, Ph.D.
Richard Smith, Ph.D.

Date of Approval:
August 17, 2006

Keywords: economics, health, drinking, grades, human capital

Table of Contents

List of Tables

Adolescent Alcohol Use and Educational Outcomes

Wesley Austin

ABSTRACT

There is some controversy over whether adolescent alcohol use has deleterious causal effects on educational outcomes. In particular, does drinking reduce academic performance and school enrollment rates and increase truancy, or does the observed negative correlation between drinking and educational outcomes merely reflect common unobservable factors? This dissertation sheds further light on the issue by estimating the causal impacts of alcohol use on various educational outcomes. Specifically, an instrumental variables model is estimated to study the effects of several drinking measures on grades, school enrollment and absenteeism.

Chapter One: Introduction

In many health-related and social science fields, there has long been concern about various harmful effects of alcohol use. One specific potential consequence in which economists have recently shown interest is the reduction of human capital accumulation. This issue is particularly relevant during adolescence and early adulthood, during which decisions regarding high school completion and college attendance are first considered and academic performance realizations that affect longer-term educational and economic outcomes are observed. Excessive drinking is associated with this age group despite its illegality until the age of 21. For instance, the 2003 National Survey on Drug Use and Health (NSDUH) reports that binge drinking, i.e. the consumption of at least five alcoholic beverages in one episode, occurred in the past month among 17 percent of high school students and 35 percent of college students.

There are several ways in which heavy drinking could potentially impair human capital formation. Intoxication could interfere with class attendance and learning, and the time spent in activities where drinking occurs could substitute away from time allocated to studying. This could hurt academic performance in the short term, which might diminish the ability or incentive to continue schooling over the longer term. Risks stemming from intoxication such as injury from accidents or fights, pregnancy and disease from unsafe sex, conflicts with parents or the law, and a tarnished reputation with school authorities might also limit the capability or motivation of a student to remain in school (Cook and Moore 1993). Alternatively, social interactions associated with drinking might improve academic achievement by providing a means of relieving stress (Williams et al. 2003).

Why is the potential impact of alcohol use on educational outcomes relevant for the discipline of economics? Substantively, human capital accumulation bears directly on the fields of labor economics, given that estimating the returns to schooling has been one of the most prominent endeavors in the field; health economics, within which a large literature on the complex relationship between schooling and health has evolved; public economics, since education is the archetypal example of a good that conveys a positive externality; and macroeconomics, because of the importance of human capital to economic growth. Moreover, investigating the various causes and consequences of alcohol use has been a foundational topic in health economics, and understanding the impacts of alcohol policies such as excise taxes, minimum legal drinking ages (MLDAs) and zero-tolerance laws is perennially of great public policy concern.

Perhaps even more importantly, economics is relevant because estimating the effect of drinking on educational outcomes is inherently an empirical matter for which the tools of econometrics can be effectively used. In particular, much evidence, ranging from anecdotal to academic research in disciplines outside of economics, has established a strong negative relationship between the regularity and intensity of drinking and human capital measures such as educational attainment and academic performance. But from a public policy perspective, distinguishing whether this relationship is causal, such that increased alcohol consumption directly reduces completed schooling or lowers grades in school, or merely correlational, with changes in other confounding variables simultaneously leading to drinking and worse educational outcomes, is critical. If frequent or heavy drinking truly compromises academic achievement, programs that decrease such alcohol use should have both the private and social values of the resulting educational gains included as benefits when optimal budget allocations are calculated. This is not true if the association between drinking and poor school performance is merely spurious.

Obtaining an estimate of the magnitude of the causal effect that alcohol use has on educational outcomes should thus be of interest to economists as well as those who analyze or participate in the formation of public policies. The role of econometrics is to identify this causal effect in the context of a broader relationship in which various third factors that are difficult to measure might create an inverse covariance between alcohol use and human capital accumulation, or whether academic performance might have a reverse causal impact on drinking. This task is a natural one to tackle using econometric techniques, because one of the mainstays of econometrics is instrumental variables (IV) regression, a method specifically designed to estimate the causal impact of a variable that does not necessary otherwise vary independently with other unobserved determinants of the outcome being examined.

Yet, only within the past 15 years has the relationship between alcohol use and human capital accumulation been addressed by economists, and even during that time research on the topic had been limited in both quantity and scope. In terms of quantity, fewer than a dozen relevant studies have been published in economics journals, many do not use IV or any other method to specifically account for the possibility that drinking is endogenous in educational outcome equations, and even some that do have used approaches that have since been criticized as unconvincing. Regarding scope, most of the literature has focused on completed schooling, with little attention paid to academic performance, especially among pre-college students, and other aspects of in-school behavior. Furthermore, while much research has found evidence that at least part of the negative relationship between alcohol use and achievement represents a causal impact of drinking, a couple of recent investigations that were quite thorough have disagreed.

This dissertation will estimate the causal impact of various frequencies and intensities of drinking among adolescents and young adults on three sets of human capital accumulation measures. In the results chapter, I begin by examining the impact on recent academic performance, specifically the probability of obtaining an 'A' as well as the probability of receiving a 'C' or lower grade. Grades are an important intermediate outcome that are related to longer-term labor market experiences through their impact on both the quantity and quality of schooling received, but have largely been ignored in the literature. Next, I analyze the effects on school enrollment. This is a commonly examined education outcome among both related studies and broader literatures on human capital accumulation, given that school attendance is easily measured and has a clear marginal impact on future wages that labor economists have long focused on estimating. Finally, I investigate effects on truancy, i.e. classes missed due to "skipping" and illness, which has not been widely studied despite clearly affecting the acquisition of human capital.

A major contribution of my research is the use of data that have not previously been examined in the literature. The aforementioned NSDUH contains extensive annual data on alcohol use and the types of educational outcomes outlined above among large nationally representative samples of U.S. residents aged 12 and above. This survey, which in different permutations dates back to 1979, is a primary source of information on U.S. substance use trends from which data are disseminated widely among both researchers and the media, and have been used by several studies on various aspects of drug use that have been published in economics journals. However, none of the previous research on the relationship between drinking and human capital accumulation has utilized these data.

Data from the NSDUH allow for both breadth and depth of coverage on the topic. Breadth comes from the ability to study aspects of all three types of educational outcomes outlined above using data from an elaborate questionnaire administered to 12–17 year olds on a wide array of youth experiences relating to education and alcohol use, and questions asked of older respondents that pertain to schooling and drinking behaviors. Depth is provided by additional questions on education and drinking intensity, which allow for a more thorough analysis of the main outcome variables this dissertation addresses.

An equally important facet of the NSDUH data is that they are conducive to the use of the IV regression methodology to estimate the causal effect of alcohol use on human capital accumulation. Abundant information is collected on preferences and experiences related to alcohol consumption, including measures of parental disapproval, peer use, and the perceived risks involved, as well as religious sentiment. An assortment of variables are therefore observed that have the potential to serve as instruments for drinking in educational outcome equations, in the sense that they are very likely to be highly correlated with alcohol use but would not have any obvious reason to be otherwise associated with educational outcomes. Current econometric techniques that are straightforward to implement will be applied to empirically test for the appropriateness of these identification restrictions. Moreover, instrumental variables models are estimated in the context of endogenous drinking measures that take the form of binary indicators.

The remainder of this dissertation is structured as follows. Chapter 2 offers an overview of the relevant literature and is divided into three parts, each corresponding to one of the distinct types of educational outcomes under investigation. Chapter 3 explains the research methodology employed to obtain estimates that convincingly represent causal effects of drinking despite its probable endogeneity in the academic achievement equations. Chapter 4 describes the NSDUH data that are analyzed and describes the empirical specification. Chapter 5 offers estimation and specification test results and chapter 6 concludes by highlighting particularly relevant findings and outlining some implications for alcohol policy.

Chapter Two: Background and Literature Review

This chapter begins by discussing the human capital theory upon which my dissertation is based. It then provides details regarding previous studies on the topic of alcohol use and educational outcomes, concentrating on those appearing in the economics literature. First it covers research on completed schooling, the most often investigated human capital consequence of drinking. The chapter then proceeds to discuss academic performance, which has been addressed by only a handful of studies, most of which focus on students in college rather than high school. The review concludes with school-related behaviors, which have largely been ignored despite their clear potential to be impacted by the consumption of alcohol.

Human Capital Theory

Human capital theory asserts that increases in a person's stock of knowledge, i.e. human capital, raises that individual's productivity in the market sector of the economy, where he/she produces money earnings, and in the nonmarket or household sector, where he/she produces commodities that enter the utility function. To obtain gains in productivity, which raises subsequent earnings, a person must invest in formal schooling. Several seminal studies (Becker (1964), Ben-Porath (1967), Grossman (1972) and Mincer (1974) have well developed this theory and its implication. Without loss of generality this section only provides an overview of Grossman (1972) and Mincer (1974).

Grossman (1972) argues that health capital is an important component of an individual utility function. In his model, an individual's stock of knowledge influences her market and nonmarket productivity, while her stock of health determines the total amount of time she can spend producing money earnings and commodities. An individual is assumed to inherit an initial stock of health that depreciates with age and can be prolonged by health investments. Education may increase the health stock if more educated people "produce" more health. Alcohol consumption may impact health by direct negative effects on cognitive functioning and other bodily damage, or by reducing educational acquisition and thus making production of health less efficient.

Theoretically, alcohol use can reduce human capital formation in two respects: 1) the direct negative impact of alcohol consumption on cognitive functioning and other health measures, 2) the time and effort devoted to obtaining and consuming alcohol, which takes time away from activities that augment human capital. Generally, if drinking has negative health consequences, the resulting reduction in human capital lowers productivity and therefore earnings and overall utility.

According to Mincer (1974), the positive relationship between market earnings and human capital investments provide incentives for obtaining higher levels of education. The theoretical specification of the model suggests that it is human capital utilized during working hours that generates earnings. Mincer expresses earnings directly as a function of years of schooling completed and experience.

Accumulation of human capital is subject to optimization, given costs of acquisition and returns to human capital investment. Alcohol use could serve to reduce the rate of return to human capital investments. Thus, the overall optimal level of human capital investment is reduced. One of the implications of the Mincer approach is that the larger the human capital investment, the higher the wage rate tends to be and the more rapid the rate of increase in wages over one's working life. Youth drinking therefore theoretically lowers the wage and its rate of increase over time.

A. Completed Schooling

Several articles published by economists have obtained estimated effects of alcohol use on educational achievement, with measures of drinking and schooling as well as conclusions varying across studies. Comparatively early research produced evidence of a negative relationship, but either made no attempt to econometrically deal with the potential endogeneity of drinking in attainment equations, or did so in a way that has since been criticized as unsatisfactory, so that it is unclear whether this negative correlation indeed represents declines in completed schooling that are caused by drinking. Two more recent and relatively thorough studies found that the causal impact of alcohol use on educational attainment is either small or non-existent. However, another recently completed and equally thorough analysis disagrees, finding evidence of a sizable reduction in the probability of high school completion attributable to previous frequent or heavy alcohol consumption.

The first study to appear in the literature, Cook and Moore (1993), analyzed National Longitudinal Survey of Youth (NLSY) data on the 753 members of the two youngest cohorts (those ages 14 or 15 in 1979) who were enrolled in 12th grade as of the 1982 interview. They estimated IV models in which the effect of current alcohol use on post-secondary attainment was identified by the state excise tax on beer and an indicator for whether the student could legally drink based on the state's MLDA. Results from three separate specifications showed that heavy drinking in 12th grade decreased subsequent schooling, by 0.13 years for each drink consumed in the preceding week, 2.3

years for students drinking on at least two occasions in the previous week, and 2.2 years for respondents who had at least six drinks on at least four occasions in the preceding month. Direct regressions of educational attainment on the alcohol policy measures similarly implied that students from states with higher beer taxes continued further in school and were more likely to graduate from college.

Mullahy and Sindelar (1994), using ordinary least squares (OLS) regressions in data on males from Wave 1 of the New Haven site of the Epidemiological Catchment Area survey, found that the onset of alcoholism symptoms by age 22 was associated with a five percent reduction in completed schooling. The authors emphasize the typically overlooked role that this adverse impact of drinking on educational attainment, if causal, would have in indirectly reducing the incomes of alcoholics.

Yamada et al. (1996) similarly analyzed data on NLSY respondents who were 12th graders during the 1981–1982 academic year using single equation probit models that did not account for the possibility that alcohol use is endogenous. They estimated that the probability of high school graduation was 6.5 percent lower for students who consumed alcohol on at least two occasions in the previous week and 2.0 percent lower for those who drank wine or liquor. In addition, drinking was found to be inversely related to beer taxes, liquor prices, MLDAs and marijuana decriminalization, meaning that each was positively associated with high school graduation rates through its covariance with alcohol use.

Koch and Ribar (2001) examined the relationship between age of drinking onset and educational attainment by age 25 in data on approximately 650 same-sex sibling pairs of each gender from the 1979–1990 waves of the NLSY. Estimates from an IV model that specified sibling onset age as the instrument for respondent onset age imply that delaying alcohol initiation by a year increases subsequent schooling by 0.22 years regardless of gender. However, they argued that this represents an upper bound for the effect size based on the sign of the bias if the assumptions needed for consistency are not met, and indeed OLS and family fixed effects models produce estimates that are three to four times smaller for males, and still smaller and sometimes insignificant for females.

Dee and Evans (2003) called into question the causal effect interpretation of the results from Cook and Moore (1993), arguing that the use of cross-state alcohol policy variation to identify the effects of drinking on other outcomes in an IV framework is potentially problematic because such variation might be correlated with unobservable attributes that affect both alcohol use and the outcome measure, in this case educational attainment. They estimated models that include state fixed effects in order to isolate the effects of within-state policy variation on drinking. In pooled cross sections from the 1977–1992 Monitoring the Future (MTF) surveys, alcohol use declined when MLDAs increased, but did not respond to beer tax changes. Moreover, in 1990 Census data on over a million members of the 1960–1969 birth cohorts, not only did educational attainment fail to rise after MLDAs were increased, but two-sample IV estimates of the effects of drinking on high school completion, college entrance and college completion were all small and insignificant.

The most recent evidence on the subject comes from Chatterji and DeSimone (2005), who estimated the effect of binge and frequent drinking by adolescents on subsequent high school dropout in data from the NLSY Young Adults using an IV model with an indicator of any past month alcohol use as the identifying instrument, while also control for a wide array of potentially confounding variables including maternal characteristics and dropout risk factors. In contrast to the last two studies cited above, the authors found that OLS yielded conservative estimates of the causal impact of heavy drinking on dropping out, such that binge or frequent drinking among 15–16 year old students lowered the probability of having graduated or being enrolled in high school four years later by at least 11 percent. The results of overidentification tests using two measures of maternal youthful alcohol use as additional instruments provided support for their empirical strategy.

B. Grade Point Average

The only previous study that attempts to identify the causal impact of drinking on academic performance among pre-college students is DeSimone and Wolaver (2004), who analyze 2001 and 2003 Youth Risk Behavior Survey (YRBS) data on high school students. They estimated regressions that included proxies for myriad potential sources of unobserved heterogeneity, particularly risk and time preference, mental health, self-esteem, and tastes for substance use. Estimated effects of alcohol use on grades are substantially reduced in magnitude when these additional covariates are added, but typically remain significantly negative. The impact on the extensive margin impact (i.e. whether or not drinking occurred) was over twice as large for binge drinking than for non-binge drinking, and binge drinking also has effects on the intensive margin, in terms of consumption frequency, that non-binge drinking did not. Drinking-related grade reductions were larger among students who are observably more risk averse and future-oriented, and effects on several outcomes with which drinking is likely associated in a non-causal way were insignificant.

Two additional studies estimated that heavy drinking reduced grades among college students surveyed in Harvard's College Alcohol Study (CAS). In the first wave of the CAS, from 1993, Wolaver (2002) used generalized method of moments to estimate a three-equation IV model in which alcohol consumption, study hours and academic performance are simultaneously determined. Instruments included measures of the ease of obtaining alcohol, parents' drinking behaviors, family attitudes about drinking and religiosity (for alcohol use) and peer studying and drinking behavior (for study hours). Results indicated that heavy drinking, in the form of any or frequent binge drinking or drunkenness during the previous month, reduced the probability of an A average by 12 to 18 percentage points, with commensurate increases in the likelihood of receiving Bs and Cs. These mostly represented a direct impact on grades, though indirect effects through a decrease in study hours were also significant. Effects were larger for the underage than for students age 21 and above. However, overidentification tests uniformly indicated that the exclusion restrictions as a whole were invalid, and drinking significantly affected neither study hours nor GPA in models that did not specify religiosity as an instrument.

Williams et al. (2003) used data from the first three CAS waves, from 1993, 1997 and 1999, and used two-stage least squares (2SLS) to estimate a similar three-equation IV model for drinking, GPA and study hours. Variables reflecting the full price of alcohol, including the beer tax, state-level variables related to access and opportunity to use and the costs of drinking and driving, and religiosity, served as instruments for past month alcohol use, which was measured as frequency of consumption, number of drinks consumed per episode, or total number of drinks. Empirically, small positive direct effects of drinking on GPA were outweighed by larger negative indirect effects that operated through reductions in study time. For instance, the estimates implied that an additional drink each day would directly raise GPA by 0.09, but indirectly lower GPA by 0.27 because of a 40 minute fall in daily study time. Again, though, overidentification statistics corresponding to the drinking equations were always significant at the 5 percent level. Also, the use of cross-state alcohol policy variation to identify alcohol use is subject to the aforementioned Dee and Evans (2003) criticism.

Kremer and Levy (2003) offered evidence on the topic from a natural experiment in which students at a large state university were randomly assigned roommates through a lottery system. Males assigned to roommates who reported drinking in the year prior to entering college had GPAs that were lower than those assigned to non-drinking roommates by one-quarter of a point on average, one-half of a point at the 10th percentile, and an average of two-thirds of a point among those who drank frequently prior to college. These effects persisted over time and appear even more important in the context of the lack of any effect on GPA of roommates' high school grades, admission test scores or family background. In contrast, prior drinking of roommates had no affect on female GPAs.

One recent study from outside economics that warrants mention is Jeynes (2002), who examined a sample of 18,726 12^{th} graders from the 1992 National Education Longitudinal Survey. He found that two measures of drinking were negatively related to achievement when simultaneously included in regression equations: increases of one standard deviation in ever having binge drank and ever having been drunk at school were associated with reductions in the composite score from standardized tests in reading and math of 0.25 and 0.09 standard deviations, respectively. Although parental socioeconomic status, daily cigarette smoking, and ever having been under the influence of marijuana and cocaine at school were also held constant, race and gender were the only other variables included in the model, which was estimated by OLS. Thus, typical econometric standards for establishing causation were not met.

C. School-related behaviors

Two studies concerning the impact of drinking on the school-related behaviors that have appeared in the economics literature and merit some attention. Roebuck et al. (2004) examined the likelihood of quitting school and truancy among NHSDA respondents, albeit those interviewed in 1997 and 1998, before questions on grades in school became part of the survey. In a sample of 15,168 12–18 year olds who had not yet completed high school, a probit regression showed that those who consumed any alcohol over the previous year were 0.6 percent more likely to not be enrolled in school, representing a semi-elasticity of nearly 0.2 at the mean dropout rate of 3.1 percent, but a negative binomial regression found no relationship between days truant and any past year drinking among those enrolled. Although measures of illegal drug use were included in the models, because the focus of the study was on marijuana rather than alcohol, no attempt was made to account for the potential endogeneity of drinking.

Markowitz (2001) estimated effects of the number of days the respondent drank and binge drank over the prior 30 days on fighting and weapons carrying in the 1991, 1993 and 1995 waves of the YRBS, using a 2SLS procedure in which three state level price measures, the beer tax, the cocaine price and an indicator of whether marijuana is decriminalized, served as instruments. Analysis of these behavior variables has connecting implications for school attendance in that disciplinary sanctions for such actions may include suspension or expulsion from school. Her results showed that the probability of having been in a physical fight during the previous year rose by about six percentage points with each day of drinking and 11 points with each binge drinking day, but neither drinking variable was related to carrying a gun or other type other weapon in the past 30 days. An important caveat is that the IV methodology is subject to the same criticism as that of Cook and Moore (1993), because state fixed effects were not included and thus cross-state price variation contributed to identification. Indeed, when census division indicators were added, both drinking measures became negative and insignificant in the fighting equations, but significantly positive in the gun carrying equation, while the F-statistics for the joint significance of the instruments fell from around four (significant though not particularly large) to below two and insignificant.

Chapter Three: Research Methodology

The purpose of this dissertation is to investigate whether alcohol use among adolescents and young adults causally influences various measures of human capital accumulation. In determining causation, the primary methodological question that must be resolved is whether drinking is properly specified as an exogenous variable with respect to educational outcomes, or should instead be treated as endogenous. To frame the discussion, consider the following equations, in which drinking (D) is a function of exogenous factors while educational attainment or achievement (E) is a function of some (but not all) of the same exogenous determinants as well as D,

(1) $$D = \alpha_0 + \mathbf{Z}\boldsymbol{\alpha}_\mathbf{1} + \mathbf{X}\boldsymbol{\alpha}_\mathbf{2} + \omega,$$

(2) $$E = \beta_0 + \beta_1 D + \mathbf{X}\boldsymbol{\beta}_\mathbf{2} + \varepsilon.$$

In the above equations, which apply to individual NSDUH respondents (with the corresponding observation-level subscript suppressed), vectors are in bold, **X** represents a set of exogenous variables that could affect both drinking and educational outcomes, **Z** represents another set of exogenous variables that could effect drinking but not educational outcomes (Z), ω and ε are error terms that encompass all factors influencing the drinking and educational outcomes, respectively, that are not explicitly controlled for on the right hand side of the equations, and the α's and β's are parameters to be estimated by the regression analysis.

Econometrically, alcohol use is exogenous in equation 2 if it is uncorrelated with the error term ε. This condition holds, by definition, if none of the unobserved educational outcome determinants are related to drinking. If so, there is no need to estimate equation 1; a single equation regression method such as OLS or probit (in the case of a binary educational outcome measure) will produce consistent estimates of the causal effect of drinking, β_1.

However, two sources of endogeneity could possibly lead to a nonzero correlation between alcohol use and the error term in (2). One is unobserved heterogeneity, which would occur if any of the unmeasured educational outcome determinants that are subsumed in the error term ε are correlated with alcohol use. The resulting estimate of β_1 in (2) would suffer from omitted variable bias, which cannot be eliminated directly because the omitted variables are not recorded in the data. Characteristics such as a lack of concern for the future relative to the present, or a disruptive event such as parental divorce, might simultaneously be responsible for greater alcohol consumption and lower attainment or achievement. Because such factors cannot be observed, though, they are not held constant in the regression, and the negative correlation between drinking and educational outcomes that they induce becomes embedded into the alcohol use coefficient, which is thus biased negatively as an estimate of the causal drinking effect. Conversely, unmeasured ability or socioeconomic background could create a positive bias in the estimated drinking effect, if higher ability students are better able to function normally after alcohol consumption or students who have more money to spend on alcohol are also higher achievers.

The other potential source of endogeneity is reverse causation. If alcohol use and educational outcomes are simultaneously determined, in an econometric sense, educational achievement will not only be a function of drinking, as specified in equation

2, but also will be a contributing factor to the decision regarding whether and how much alcohol to consume, a mechanism not accommodated by the above two-equation framework. In terms of equation 2, shocks to the error term ε that, by definition, influence educational outcomes will ultimately extend to drinking through the feedback effect of educational outcomes on alcohol consumption, thus creating a correlation between alcohol use and ε that renders the estimate of the drinking effect β_1 inconsistent. Again, the resulting bias could occur in either direction depending on the source of the reverse causation: it would be positive if academic success is celebrated by drinking or leads to additional income that is spent on alcohol consumption, but negative if alcohol is used to drown academic sorrows or if academic shortcomings reduce the opportunity cost of drinking.

In order to investigate the possibility that alcohol use is endogenous as an explanatory factor for educational outcomes (to which I will refer to as achievement, for parsimony, but without loss of generality, for the remainder of the section) and generate estimated effects of drinking on achievement that can be interpreted as causal, this analysis will use the method of instrumental variables (IV). To use IV, the vector **Z** in equation 1 above must exist, i.e. there must be at least one, and preferably two or more, variables (i.e. instruments or IV) that affect alcohol use but have no direct impact on achievement. In the case of exactly one instrument **Z**, the IV method works by estimating the causal drinking effect β_1 as the ratio of the sample correlation between the instrument and achievement to the sample correlation between the instrument and alcohol use, i.e.

(3) $$\hat{\beta}_1 = \hat{\text{corr}}[\mathbf{Z}, \text{E}] / \hat{\text{corr}}[\mathbf{Z}, \text{D}],$$

where the “^” symbolizes that the quantity is estimated from the data and the correlations are estimated while holding constant the vector **X** of explanatory factors. The idea is that because the instrument is exogenous and related to achievement only through drinking, the sample correlation between the instrument and achievement is purely a product of that between drinking and achievement. Thus, the sample correlation between the instrument and achievement merely needs to be standardized by that between the instrument and drinking in order to be used as an estimate for the causal effect of drinking on achievement. In the case of two or more instruments, $\hat{D}$, the linear projection of **Z** onto D, takes the place of **Z** in equation 3, in order to reduce the dimensionality so that both correlation expressions are scalars. In fact, this is true even in the case of a single instrument, but for expository purposes it is simpler to consider (3) as written above.

Equation 3 makes transparent the two important conditions that the instrument vector **Z** must satisfy in order for IV to produce consistent estimates of the drinking effect β_1: the instruments must be highly correlated with alcohol use but not correlated with the educational outcome under investigation through any other mechanism besides drinking. If the correlation between the instruments and drinking is not statistically significant, the denominator in (3) is statistically equal to zero, thus rendering the expression for β_1 indeterminate. The strength of this correlation is easily judged from the F-statistic for the joint significance of $\boldsymbol{\alpha_1}$ in equation 1, which is equivalent to the t-statistic on the scalar α_1 when there is a single instrument. Minimally, $\boldsymbol{\alpha_1}$ should be significant at the 1 percent level; beyond this, Staiger and Stock (1997) advise the more stringent requirement that the associated F-statistic be at least 10.

Meanwhile, if a direct correlation between the instruments and achievement exists outside of the pathway from the instruments to drinking to achievement, the numerator in (3) includes variation that is not part of the relationship between drinking and achievement, and consequently the expression is no longer a consistent estimate of the effect of drinking on achievement. The reason multiple instruments are preferred is that this overidentifies equation 2, which allows for specification tests to determine the empirical validity of excluding the instrument set **Z** from (2). In particular, under the null hypothesis that the instruments are not separately correlated with achievement, the sample size multiplied by the R^2 from a regression of the residual in (2), $\hat{\varepsilon}$, on all the exogenous variables (i.e. a constant, **X** and **Z**) is distributed as chi-square with degrees of freedom equal to one less than the number of instruments. The logic is that as the extent of any direct correlation between the instruments and achievement increases, the strength of the partial correlation between $\hat{\varepsilon}$ and **Z**, and thus the R^2 from the above auxiliary regression, does as well.

Typically, the estimator represented by equation 3 is generated by a two-stage least squares (2SLS) procedure. The first stage estimates equation 1 above using OLS. From the estimated parameters, predicted values of alcohol use, $\hat{D}$ (which earlier was called the linear projection of **Z** onto D), are constructed for each respondent using their corresponding values of the explanatory variables **X** and instruments **Z**. The second stage estimates equation 2 using the fitted values $\hat{D}$ in place of observed drinking D. The entire process is performed within a pre-programmed routine in econometric estimation software packages (e.g. Stata), which also provides correct estimates of the standard errors (which should be calculated using the actual rather than predicted drinking variable). In both equations, standard errors that are robust to arbitrary forms of heteroskedasticity will be displayed in the output tables and used to calculate relevant test statistics.

2SLS yields consistent estimates even when alcohol use and/or achievement are represented by a binary indicator, which occurs frequently in my data. However, for binary drinking measures, e.g. an indicator of any past month binge drinking, I will utilize an approach, suggested by Wooldridge (2003) to improve efficiency, which is similar to 2SLS with two modifications. First, before running 2SLS, a preliminary probit regression for equation 1 is estimated. Second, the ensuing 2SLS procedure uses the predicted probabilities of drinking from the probit regression as instruments in place of **Z**. The resulting estimates are likely to be similar in magnitude to those that would be generated by the analogous 2SLS regression, but standard errors will be slightly smaller. The necessary conditions for using 2SLS and its desirable properties still hold.

Two other methodological points should be mentioned. First, although IV estimates are consistent if the instrument strength and exogeneity conditions outlined above are satisfied, they are inefficient relative to OLS if it turns out that alcohol use is truly exogenous with respect to achievement, in which case the OLS estimates can be interpreted as causal effects. Even under ideal circumstances, i.e. very strong instruments that empirically have very little correlation with achievement conditional on drinking, standard errors from IV regressions tend to be much larger than those from OLS regressions. Thus, it is desirable to econometrically test the null hypothesis that drinking is exogenous in the achievement equation. This is easily done using a Hausman (1978) test, which examines whether the IV and OLS estimates of β_1 are significantly different from each other. The relevant test statistic is simply a t-statistic in which the numerator is the difference between the IV and OLS estimates of β_1, and the denominator is the square root of the difference in the estimated variances of β_1 under IV and OLS. Rejection of this null implies that OLS estimates are inconsistent and hence conclusions should be based on IV estimates; failure to reject this null means that OLS estimates are preferable because of their smaller standard errors.

Finally, an additional advantage of IV is that it also addresses the issue of errors in the measurement of the drinking variables, which will prospectively be present at least to some degree because data are self-reported and thus, in the case of measures like alcohol use that require respondents to remember whether and how much consumption occurred during certain time periods, subject to recall error. Even if any resulting measurement error is random, e.g. uncorrelated with actual or measured drinking or achievement, OLS estimates will be biased towards zero (the potential for measurement error that varies systematically with drinking or achievement will be discussed in the following section). However, IV estimates are consistent even when alcohol use is measured with random error.

Chapter Four: Data and Empirical Specification

Data Description

This dissertation will analyze data from the 2002 and 2003 waves of the NSDUH. I do not incorporate earlier data for two reasons. First, beginning in 2002 the NSDUH administrators undertook steps to improve the quality of the data gathered by the survey, which included implementation of improved data collection procedures and the payment of $30 to respondents for completed surveys (which raised response rates). Second, before 2002, information utilized in this study regarding variables that serve as instruments do not consistently appear.

The NSDUH, sponsored by the Substance Abuse and Mental Health Services Administration (SAMHSA), is administered annually to approximately 55,000 civilian, non-institutionalized individuals age 12 and over, chosen so that the application of sample weights produces a nationally representative sample, with approximately equal numbers of respondents from the 12–17, 18–25 and 26 and over age groups. Geographically, eight large states contribute roughly 3,600 respondents each and remaining states provide about 900 respondents each. The sample is stratified by state, which are separated into field interviewer regions that are further divided into segments consisting of adjacent census blocks. Interviewers visit selected households, one or two residents of which complete the survey. The interviewer uses a computerized survey to enter some responses, but most are answered privately in a way that precludes interviewer knowledge of the answers supplied.

An important aspect of the survey that partially dictates sample composition is that the Youth Experiences section, which is the only source of information in the NSDUH on academic performance, is administered only to 12–17 year olds. Academic performance is represented by grades in the most recent marking period, coded as a categorical variable with four choices: A+ through A–, B+ through B–, C+ through C–, or D or below. This dissertation uses this information in the form of two binary variables, one indicating whether the grade was 'A' and the other indicating whether it was 'C' or below.

Questions on other human capital variables were asked to all respondents as part of the core demographics section. Current school enrollment is a binary variable indicating whether the respondent is currently enrolled in middle or high school (including those who are home schooled) or a college/university. The sample for this part of the study is restricted to 16-25 year olds. Approximately 99 percent of youth ages 15 and under report attending school, and individuals ages 26 and above who have not graduated from college are particularly likely to have experienced previous gaps in school enrollment, not currently be enrolled and not return to school in the future.

In addition, all enrolled respondents are asked to report the number of days of school over the past month that were missed because of illness or injury and that were missed due to skipping, with a response range from 0 to 30 days. This dissertation will also examine the effect of drinking on absenteeism due to skipping classes and illness or injury. The latter merits attention in that alcohol consumption may be the reason for reporting illness or injury.

Alcohol use is observed for consumption of various types and over different time periods. The three main measures on which the analysis will focus are the number of days alcohol was consumed during the previous year, the number of drinks consumed in the past month and the occurrence of binge drinking in the past month. Although the timing of the number of drinks and binge drinking variables is not an ideal match for some of the educational outcome measures, in the sense that past month consumption cannot literally affect behavior that preceded the past month, my work will follow that of previous studies in assuming that previous month drinking patterns proxy those occurring in the recent period prior to the previous month.

One other additional piece of alcohol consumption information will also be examined: an indicator of whether respondents exhibited symptoms of alcohol abuse or dependence in the past year. This is retrospectively coded by SAMHSA based on responses to questions corresponding to criteria outlined in the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV), the clinical standard for establishing drug abuse and dependence. There are seven criteria utilized, three of which must be met for categorization as dependent: 1) Spent a great deal of time over a period of a month getting, using, or getting over the effects of the substance, 2) Unable to keep set limits on substance use or used more often than intended, 3) Needed to use substance more than before to get desired effects or noticed that using the same amount had less effect than before, 4) Unable to cut down or stop using the substance every time he or she tried or wanted to, 5) Continued to use substance even though it was causing problems with emotions, nerves, mental health, or physical problems, 6) Reduced or gave up participation in important activities due to substance use, and 7) experienced substance specific withdrawal symptoms at one time that lasted for longer than a day after they cut back or stopped using. To be categorized as abusive, the respondent must respond positively to one of more of the following criteria: 1) has serious problems due to substance use at home, work or school, 2) used substance regularly then did something where the substance use might have put them in physical danger, 3) substance use caused actions that repeatedly got them in trouble with the law, 4) has problems caused by substance use with family or friends and continued to use substance even though it was thought to be causing problems with family and friends.

A potentially problematic attribute of the data is non-random measurement error emanating from the self-reported nature of responses. Although IV will eliminate bias from random measurement error, as previously discussed, it cannot salvage data plagued by systematic measurement error, which could affect both educational outcomes and alcohol use. In particular, one might expect respondents to artificially inflate grades, but underreport alcohol consumption. However, studies on the quality of self-reported academic performance and drinking data suggest that such reporting bias should be minimal. Cassady (2001) found that self-reported GPA values are "remarkably similar to official records" and therefore are "highly reliable" and "sufficiently adequate for research use." Similarly, Grant et al. (1988), Midanik (1988) and Reinisch et al. (1991) concluded that youth drinking self-reports are reliable, based on the consistency of responses to alcohol use questions from repeated interviews. Johnston et al. (1988, pg. 20) write that "the considerable amount of inferential evidence that exists strongly suggests that the self-report questions produce largely valid data." And Harrison and Hughes (1997) found that individuals tend to underreport the use of stigmatized substances such as cocaine and heroin much more so than alcohol, whereas survey methods not requiring subjects to verbally answer questions, as in the NSDUH where interviewers are unable to match responses with respondents, increase the accuracy of substance use self-reports.

Empirical Specification

As discussed in chapter three, the empirical strategy employed involves a two equation model. First, a probit regression is conducted and predicted values of drinking are obtained. Second, a standard 2SLS regression is performed. Once again the equation system stated in chapter three is:

(1) $$D = \alpha_0 + \mathbf{Z}\boldsymbol{\alpha}_\mathbf{1} + \mathbf{X}\boldsymbol{\alpha}_\mathbf{2} + \omega,$$

(2) $$E = \beta_0 + \beta_1 D + \mathbf{X}\boldsymbol{\beta}_\mathbf{2} + \varepsilon.$$

In terms of the NSDUH variables, the notation is as follows. *D* represents one of the four drinking measures defined in the preceding section: 1) the number of days the respondent drank in the previous year, 2) the number of drinks the respondent consumed in the previous month, 3) whether the respondent engaged in binge drinking in the last 30 days or, 4) whether the respondent is categorized as alcohol dependent and/or abusive.

E denotes one of the educational outcomes examined: 1) indicators of obtaining an ‘A’ and a ‘C’ or lower average, 2) an indicator of school enrollment and, 3) the number of days the respondent reported absenteeism due to “skipping” or due to illness or injury. **Z** is a set of instrumental variables that influence drinking but not education directly. The instruments are discussed in detail below. **X** represents other plausibly exogeneous educational outcome determinants that are included as explanatory variables in the regression equations. The specific variables utilized in the **Z** and **X** vectors depend upon the outcome, *E*, under investigation and the subsample used: some variables that serve as instruments and explanatory variables are only available only for 12–17 year olds in the Youth Experience section of the NSDUH, while others are available for all age categories.

When *E* is an indicator of obtaining an 'A' average or a 'C' or lower average, the analysis is conducted using high school students ranging in age from 12 to 17 as grades are only available for the Youth Experience sample. The set of instrumental variables for the analysis of grades includes indicators for perceived risk of consuming alcohol, parental disapproval of drinking and peer alcohol use, as defined below. The **X** vector includes measures of student age, race, grade level, family income, family size, whether each of the mother and father live in the household, whether the respondent was born in the U.S., the number of times the family relocated in the previous five years, the extent to which parents help with homework, and two population density categories, also defined below.

When *E* is an indicator of school enrollment or the number of days missed due to "skipping" classes or due to illness or injury, two subsamples are analyzed. One sample includes high school age students ranging in age from 16 to 19 years old and another sample includes college age students ranging from 18 to 25 years old. For the high school age sample, high school graduates are excluded. For the college sample, only high school graduates are included, and college graduates are excluded. For these educational outcomes, the instruments employed are indicators of perceived risk of consuming alcohol, whether religion is important in the respondent's life and whether religious believes influence decisions. The **X** vector includes race, family income, family size, whether the respondent was born in the U.S., the number of times the respondent moved in the previous five years, and two population density categories. Though the same instruments and explanatory variables are used in each case, the samples are stratified in order to allow for the effect that drinking has on enrollment and absenteeism to differ between high school and college students.

The explanatory variables utilized in all the samples are defined as follows. Family income is measured in four categories: $10,000-$19,999; $20,000-$49,999; $50,000-$79,999; and $75,000 or greater, with $10,000-$19,999 as the omitted category. Population density is represented by indicators for two categories: an MSA with one million persons or greater and an MSA of less than one million persons, with non-MSA areas as the omitted category. For race, indicators are specified for African Americans, Native Americans, Asians, and non-white Hispanics, with whites as the omitted category. Family size is measured using two variables, the number of members if the household has one to five members and an indicator for those with over five members, for which the numerical variable is set to zero.

Those explanatory variables that were only available for the youth experience sample are specified as follows: age indicators for 15, 16, or 17 years old (with age 12–14 omitted, because very few high school students are age 12–13), indicators of whether the mother or father reside in the household, indicators for whether parents assisted the student with homework always, sometimes or seldom in the past 12 months, with never as the omitted category, and indicators for whether the student is currently attending the 10th, 11th or 12th grade (with 9th as the omitted grade).

Several NSDUH variables conceivably influence drinking without having direct effects on educational outcomes and are thus candidates to serve as instrumental variables in the analysis. The specific variables utilized in this study are: parental disapproval of alcohol use; peer use of alcohol, perceived risk of bodily harm from alcohol use; whether religion is important in the respondents' life; and whether religious beliefs influence the respondents' decisions. The first two variables are recorded only for 12-17 year olds as part of the Youth Experiences section of the NSDUH. The latter three are available for all age groups and are employed in subsamples of 16-19 and 18-25 year olds.

Information is reported on whether parents would neither approve nor disapprove, somewhat disapprove or strongly disapprove of the respondent having one or two drinks of an alcoholic beverage nearly every day. In this study, a binary variable is created and coded as 0 if the parent is indifferent or somewhat disapproves and 1 if the parent strongly disapproves. Peer use information reflects a question about perceptions of the respondent regarding whether none, a few, most, or all students in the same grade at his or her school consume alcohol. For the peer use variable, a binary measure is defined to designate if the respondent feels that most or all schoolmates consume alcohol. Potential endogeneity of the peer variable, stemming from a possible connection between one's own behavior and perceptions about the behavior of others, should be mitigated by the fact that the relevant questions cover all classmates rather than simply friends, who are presumably chosen by the respondent. Norton et al. (1998), Gaviria and Raphael (2001) and Kremer and Levy (2003) each found evidence that increased rates of drinking among peers raised the propensity to consume alcohol.

With regard to the other three instrumental variables utilized in this study, NSDUH respondents were asked questions regarding religious beliefs and perceived risks involved in drinking. Religiosity encompasses the strength of agreement or disagreement with statements regarding whether religious beliefs are important and influence decisions. The risk variables indicate the extent to which respondents think that people risk harming themselves "physically and in other ways" by having four or five drinks nearly every day.

For the religion factors, binary measures are constructed to indicate if the respondent agrees that religious beliefs are important in his/her life and if the respondent agrees that religious beliefs influence his/her decisions. For risk, the binary measure indicates if the respondent feels there is moderate to great risk of harm, physically or otherwise, from consuming four to five drinks daily.

The instruments are presumed to be correlated with youth drinking yet not exhibit a direct correlation to the educational outcome in question. However, that characteristic of the instruments needs to be empirically tested, as one can imagine other indirect avenues through which the instruments might impact educational outcomes. For instance, parents that disapprove of alcohol use might also strictly discipline their children to succeed in school and channel more family resources to schooling, thereby raising grades and enrollment and lowering truancy. In addition, students who drink might overestimate the accurate amount of drinking by schoolmates and vice versa, and a correlation with education might exist if classmate drinking proxies for school quality. Students that perceive greater risks in drinking may also be more risk averse in general, therefore perceiving greater risks associated with academic failure and experiencing better educational outcomes. Finally, students who state religion is important may also believe education is more important and may be greater discipline to do well in school; more generally, strong religious beliefs might also influence attitudes about education.

Given these caveats, the true exogeneity of the instruments may be comprised. Overidentification tests, and tests of the sensitivity of the IV estimates utilizing differing sets of instruments, are thus conducted to verify the hypothesis of instrument exogeneity.

Theoretically the bias of OLS estimates could be in either direction, meaning that OLS estimates could be either larger or smaller than their IV counterparts. OLS estimates will overestimate the negative effects of drinking if unobserved factors such as the emotional distress or personality characteristics induce a student to drink more while commensurately lowering academic performance, or if reverse causation is such that poor grades lead to drinking. In contrast OLS will underestimate negative effects of drinking if income effects lead to more drinking and better academic performance, or good performance is celebrated with drinking. Also, random measurement error will cause attenuation bias in OLS. It is expected that the major cause of bias is unobserved heterogeneity that inflates the magnitude of OLS estimates, but that income and measurement error effects working in the opposite direction might also be important.

Chapter Five: Empirical Results

This chapter discusses the empirical results. The causal effect youth alcohol use has on these educational outcomes is discussed, as are the effects of other independent variables. The effect of alcohol consumption on high school grades is discussed first. The effect of alcohol use on the probability of obtaining an 'A' average versus lower grades is examined, as is the probability of obtaining a 'C' average or lower versus higher grades. Those results are based solely on the youth experiences sample of 12–17 year olds. Then, the effects of alcohol consumption on the probability of school enrollment are analyzed utilizing the samples of 16–19 year olds who have not graduated from high school and 18–25 year olds who have graduated from high school but not from college. The effects of drinking on absenteeism (the number of school days the student missed due to "skipping" class and illness or injury) are studied as well, utilizing the age 16–19 and 18–25 samples. The youth experience sample utilizes three instrumental variables to identify drinking in the grade equations: parental disapproval of drinking, perceived risk of harm from drinking and peer use of alcohol. The enrollment and absenteeism regressions utilize three instrumental variables: perceived risk of harm from drinking, whether religious beliefs are an important part of the respondents' life, and whether religious beliefs influence how the respondent makes decisions.

In order to assess the impact of instrumental variables, this chapter includes comparisons of coefficient estimates from single equation estimation using OLS with those from instrumental variables models. To determine if the effects of youth drinking on the educational measures are influenced by the choice of instruments, separate analyses are conducted using different sets of instruments. All regressions employ all three instruments in the main specification and the robustness analyses use pairs of two instruments.

Tables 1 and 2 present summary statistics for each of the variables utilized for the 12–17 and 18–25 year olds respectively. For the youth sample, the mean number of days drinks were consumed in the past year is 19.8 while the mean number of drinks consumed in the past month is 8.2. Approximately one-sixth of high school students binge drink while about one-tenth are classified as alcohol dependent. The mean value of reported peer drinking is 0.5. A vast majority of youths, 87 percent, report their parents' discourage drinking. Family income is less than $20,000 for 17 percent of respondents but greater than $75,000 for 26 percent. About 72 percent of respondents live in an MSA, roughly equally split between MSAs with populations greater than and less than one million. Fathers are less likely to be present in the household than are mothers. The proportion of parents that help with homework is also very high. African Americans comprise about 12 percent of the sample while non-white, non-black Hispanics account for about 14 percent.

Table 1. Descriptive Statistics (Youth Experience sample)
(n=18,231)

Variable	Mean	Standard Deviation
Number of days drank-past year	19.780	47.620
Number of drinks in previous month	8.172	38.325
Binge drinking in the past 30 days	0.153	0.360
Abuse/ Dependence on alcohol classification	0.091	0.287
Respondent perceives risk of harm from drinking	0.823	0.382
Respondent perceives schoolmate (peer) drinking	0.496	0.500
Parents disapprove of alcohol	0.867	0.399
Probability of an 'A' grade	0.265	0.441
Probability of a 'C' or lower grade	0.311	0.464
Family income ($20,000-$49,999)	0.365	0.481
Family income ($50,000-$74,999)	0.207	0.405
Family income ($75,000 or more)	0.260	0.438
MSA segement with 1+ million persons	0.361	0.480
MSA segment of less than 1 million	0.358	0.479
Age of student (14 years old)	0.181	0.380
Age of student (15 years old)	0.268	0.440
Age of student (16 years old)	0.274	0.446
Age of student (17 years old)	0.275	0.446
Mother in household	0.906	0.291
Father in household	0.728	0.444
Parents help with homework (always)	0.526	0.499
Parents help with homework (sometimes)	0.235	0.420
Parents help with homework (seldom)	0.125	0.320
Grade in (10th grade)	0.270	0.440
Grade in (11th grade)	0.250	0.430
Grade in (12th grade)	0.160	0.370
Race (African American)	0.120	0.320
Race (Native American)	0.030	0.190
Race (Asian)	0.029	0.169
Race (non-white Hispanic)	0.141	0.348
Number in family	3.267	1.529
Number in family (>5)	0.133	0.340
year 2002	0.485	0.499

For the 18–25 year old sample, the mean number of drinks consumed in the past year increases to 58.2, and the number of drinks consumed in the past month rises to 24.0. The incidence of binging is 0.41 and the mean of risk associated with alcohol use falls from 0.82 in the youth sample to 0.71. Among the 44 percent of respondents reporting current school enrollment, about one day per month of classes is skipped and another nearly three-quarters of class days per month is missed due to illness. The racial composition is similar to that of the youth sample.

Table 2. Descriptive Statistics (18-25 years old)
(n=28,065)

Variable	Mean	Standard Deviation
Number of days drank-past year	58.172	76.960
Number of drinks in previous month	23.982	65.176
Binge drinking in the past 30 days	0.408	0.490
Abuse/ Dependence on alcohol	0.173	0.378
Respondent perceives risk of harm from drinking	0.708	0.454
Respondent states religion is important in life	0.695	0.450
Respondent states religion influences their decisions	0.618	0.485
Current school enrollment	0.437	0.496
Number of skipped school days (past 30 days)	0.924	2.059
Number of school days missed due to illness (past 30 days)	0.722	1.850
Family income ($20,000-$49,999)	0.397	0.341
Family income ($50,000-$74,999)	0.135	0.489
Family income ($75,000 or more)	0.136	0.341
MSA segement with 1+ million persons	0.340	0.473
MSA segment of less than 1 million	0.392	0.488
Age of student (19 years old)	0.147	0.354
Age of student (20 years old)	0.138	0.345
Age of student (21 years old)	0.135	0.342
Age of student (22-23 years old)	0.216	0.411
Age of student (24-25 years old)	0.196	0.397
Last grade completed (Freshman)	0.145	0.352
Last grade completed (Sophomore/ Junior)	0.197	0.398
Race (African American)	0.135	0.330
Race (Native American)	0.013	0.115
Race (Asian)	0.026	0.160
Race (non-white Hispanic)	0.167	0.373
Number in family	2.948	1.327
Number in family (>5)	0.080	0.270
year 2002	0.487	0.499

Table 3 presents summary statistics for the sample of 16–19 year olds who have not graduated from high school. Drinking is greater than in the youth experience sample but not as high as in the 18–25 sample: about 24 percent reports binge drinking in the past 30 days and 12 percent are classified as abusing/dependent on alcohol. Those enrolled skipped school about half a day in the past month, and also missed just over a day per month due to illness. Household income and racial composition roughly mirrors that of the other subsamples.

Table 3. Descriptive Statistics (16-19 years old)
(n=13,526)

Variable	Mean	Standard Deviation
Number of days drank-past year	28.697	56.155
Number of drinks in previous month	12.844	48.536
Binge drinking in the past 30 days	0.239	0.427
Abuse/ Dependence on alcohol classification	0.123	0.328
Respondent perceives risk of harm from drinking	0.738	0.439
Respondent states religion is important in life	0.713	0.451
Respondent states religion influences their decisions	0.599	0.490
Current school enrollment	0.863	0.342
Number of skipped school days (past 30 days)	0.509	1.620
Number of school days missed due to illness (past 30 days)	1.093	2.241
Family income ($20,000-$49,999)	0.363	0.480
Family income ($50,000-$74,999)	0.187	0.390
Family income ($75,000 or more)	0.230	0.421
MSA segment with 1+ million persons	0.347	0.470
MSA segment of less than 1 million	0.366	0.481
Age of student (16 years old)	0.231	0.421
Age of student (17 years old)	0.230	0.421
Age of student (18 years old)	0.188	0.390
Last grade completed (9th)	0.231	0.421
Last grade completed (10th)	0.240	0.427
Last grade completed (11th)	0.229	0.420
Race (African American)	0.134	0.341
Race (Native American)	0.127	0.112
Race (Asian)	0.028	0.165
Race (non-white Hispanic)	0.145	0.352
Number in family	3.181	1.493
Number in family (>5)	0.124	0.330
year 2002	0.482	0.499

A. Drinking and Grades

This section presents results for the effect of youth drinking on grades. Specifically, causal effects that drinking has on the probabilities of obtaining certain grades is estimated using the three instrumental variables listed earlier. The overidentification statistics aid in revealing whether the instrument set is exogenous with respect to academic performance. An analysis is conducted that employs differing pairs of instruments to determine if some instrument sets are more plausibly exogenous. The main results of the IV analysis are also compared with parameter estimates obtained using OLS. While the discussion here focuses mainly on the effects of alcohol consumption on grades, appendix 1 presents the probit estimates from the drinking equations. For the binge drinking measure, appendices 2 and 3 show the IV coefficients and standard errors of all exogenous variables for the for the probability of obtaining an 'A' and a 'C' or lower, respectively. For probit models, tables contain marginal effects at the explanatory variable means.

First Stage Regression Results

Results from the probit and first stage regressions of the drinking measures on the instrumental variables for grade probabilities are shown in table 4. The two left columns are from OLS first stage regressions. The two right columns are probit marginal effects and associated standard errors, except for the last row, which shows the coefficient and standard error for the predicted value of drinking obtained from the drinking probits.

Table 4. Probit/ First stage regression estimates for the grade outcomes
(n=18,231)

exogeneous variables	number of days drank in past year	number of drinks in past month	Binge drinking	Abuse/ Dependence on alcohol
Risk of bodily harm from drinking	-15.890	-10.420	-0.125	-0.078
	(0.894)	(0.749)	(0.008)	(0.006)
Peer use of alcohol	11.949	6.241	0.086	0.054
	(0.716)	(0.600)	(0.005)	(0.004)
Parental disapproval of alcohol use	-24.706	-14.201	-0.162	-0.074
	(0.716)	(0.845)	(0.005)	(0.007)
F stat/ chi2-coefficient of joint significance	486.83	246.49	1104.87	612.52
P-value of significance level	(0.0000)	(0.0000)	(0.0000)	(0.0000)
predicted drinking coefficient			0.978	0.934
			(0.024)	(0.032)

The results demonstrate that peer alcohol use has positive effects on drinking while parental disapproval and perceived risk have negative effects. For respondents who agreed that there is moderate to great risk of harm from consuming 4-5 drinks almost every day, the number of days that drinking occurred in the past year is lowered by approximately 16 days, the number of drinks consumed in the past month is reduced by 10 days, the probability of binge drinking in the last 30 days falls by 0.13, and the likelihood of being categorized as abusive/dependent on alcohol falls by 0.08.

For respondents who report most or all their schoolmates use alcohol, the number of days drinking occurred in the past year rises by 12 days, the number of drinks consumed in the past month rises by 6, the probability of binge drinking in the last 30 days rises by 0.09, and the likelihood of being categorized as abusive/dependent on alcohol rises by 0.05.

Parental objection to alcohol use has the strongest effect. For respondents who report their parents strongly disapprove of having one or two drinks per day, the number of days drinking occurred in the past year is lowered by 25 days. The number of drinks consumed in the past month is reduced by 14, while the likelihood of binge drinking in the last 30 days falls by 0.16. The likelihood of being categorized as abusive/ dependent on alcohol falls by 0.07. The F statistics and χ^2 coefficients and associated p-values give strong evidence of joint instrument significance with respect to all the drinking measures. The predicted drinking coefficients in the first stage regressions from the Wooldridge binary endogenous variable method are 0.98 for binge drinking in the past 30 days and 0.93 for abuse/dependence on alcohol.

The Effects of Drinking on the Probability of Obtaining an 'A'

Table 5 presents results for the probability of obtaining an 'A' versus lower grades. Drinking has significant, negative effects on the probability of earning an 'A'.

Table 5. IV estimates of drinking on the probability of an 'A'
All three youth experience instruments
(n=18,231)

Alcohol variables	IV	OLS
number of days drank-past year	-0.003*	-0.0009*
Marginal Effect Standard Error	(0.0002)	(0.00006)
P-value of overidentification test	0.106	
Hausman statistic	-10.008*	
number of drinks in past month	-0.005*	-0.0006*
Marginal Effect Standard Error	(0.0005)	(0.00008)
P-value of overidentification test	0.051	
Hausman statistic	-9.826*	
binge drinking	-0.351*	-0.125*
Marginal Effect Standard Error	(0.030)	(0.008)
P-value of overidentification test	0.006	
Hausman statistic	-8.621*	
abuse/ dependence on alcohol	-0.557*	-0.096*
Marginal Effect Standard Error	(0.055)	(0.010)
P-value of overidentification test	0.007	
Hausman statistic	-8.944*	

*Statistically significant at 1%

An additional day increase in the number of past year drinking days reduces the probability of achieving an ‘A’ by 0.003 percentage points. For instance, if a student reports drinking 52 days in the previous year, the probability of having an ‘A’ average in the current grading period is reduced by 0.156 points compared to not drinking at all.

For each drink increase in the number of drinks consumed in the past month, the probability of obtaining an ‘A’ is reduced by 0.005. If the student consumes, on average, two alcoholic drinks per day in the past 30 days, the probability of the student having an ‘A’ average falls by 0.30 relative to abstaining. For respondents that reported binge drinking in the previous 30 day period, there is an associated reduction in the probability of obtaining an ‘A’ average of 0.35. For those categorized as abusive/ dependent on alcohol, the probability of obtaining an ‘A’ is reduced by 0.56.

The Hausman statistics signify that there are statistically significant differences between the OLS and IV parameter estimates for all the drinking measures. The overidentification tests for binging and abuse/dependence on alcohol have associated p-values that offer little evidence in support of the assumption of exogeneity. For the other two drinking measures, instrument exogeneity is not rejected at the 5 percent level.

There are fairly large negative effects on the probability of acquiring an 'A' for each drinking measure analyzed. For instance, for each additional drink consumed in the previous month, there is a 1.8 percent decline in the probability of obtaining an 'A'. This indicates that alcohol consumption on the part of high school students could be impairing the learning process, which in turn reduces the capability of the student to earn top grades. There is also an opportunity cost involved in drinking, which includes reduced study time and possibly increased devotion of the students' monetary resources to consuming alcohol that detracts from the prospect of receiving an 'A' average. These results imply that those costs could be substantial. While the overidentification tests offer weak support for the assumption of instrument exogeneity, the following section further explores the issue by conducting the analysis utilizing differing pairs of instruments.

Instrument Robustness and the Probability of an 'A' Average

Table 6 shows the results of regressions performed with varying pairs of instruments. This exercise is undertaken to determine if there is any sensitivity in the main results to changes in the instrument set. The instrument that is omitted from the IV combination is utilized as an explanatory variable and its coefficient and standard error is reported in the table.

Table 6. IV estimates of drinking on the probability of a 'A' using IV pairs
(n=18,231)

	parent disapprove and risk	risk and peer use	parent disapprove and peer use
Alcohol variables			
number of days drank-past year	-0.002*	-0.003*	-0.003*
Marginal Effect Standard Error	(0.0002)	(0.0004)	(0.0003)
P-value of overidentification test	0.270	0.386	0.036
Hausman statistic	-7.723*	-7.792*	-7.970*
Coefficient (Standard Error) of omitted IV	-0.015 (0.008)	-0.028 (0.014)	0.004 (0.010)
number of drinks in past month	-0.004*	-0.006*	-0.005*
Marginal Effect Standard Error	(0.0005)	(0.0007)	(0.0006)
P-value of overidentification test	0.731	0.049	0.018
Hausman statistic	-8.047*	-7.479*	-7.953*
Coefficient (Standard Error) of omitted IV	-0.020 (0.008)	-0.021 (0.016)	0.007 (0.012)
binge drinking	-0.295*	-0.327*	-0.315*
Marginal Effect Standard Error	(0.033)	(0.041)	(0.037)
P-value of overidentification test	0.015	0.003	0.004
Hausman statistic	-5.985*	-5.512*	-5.954*
Coefficient (Standard Error) of omitted IV	-0.026 (0.007)	-0.010 (0.012)	0.016 (0.010)
abuse/ dependence on alcohol	-0.459*	-0.468*	-0.496*
Marginal Effect Standard Error	(0.062)	(0.068)	(0.070)
P-value of overidentification test	0.002	0.001	0.001
Hausman statistic	-6.354*	-5.870*	-6.160*
Coefficient (Standard Error) of omitted IV	-0.026 (0.008)	-0.026 (0.011)	0.016 (0.010)

*Statistically significant at 1%

The table 6 estimates are similar to those in the regressions in which all three instruments are employed. For binging and abuse/dependence on alcohol, the overidentification tests continue to reject instrument exogeneity. Instrument exogeneity is not rejected for the past year and past month drinking variables when peer use is entered into the grade equation, and this specification also yields the highest overidentification p-values for binge drinking and abuse/dependence. Peer use is accordingly always significant in the grade equation, as is parental disapproval for past year drinking and abuse/dependence. Hausman tests indicate there are statistically significant differences between IV and OLS estimates in all specifications.

Overall, the results of this sensitivity analysis offer some evidence that the identification strategy produces consistent estimates of the effect of drinking on the probability of an ‘A’ average. At the 5 percent level, instrument exogeneity is not rejected in some cases. In addition, for each measure of alcohol use, drinking coefficients are similar regardless of which instruments are used to identify drinking.

The Effects of Drinking on the Probability of a ‘C ’or Lower Average

Table 7 presents the IV regression estimates for the probability the respondent has a ‘C’ or lower grade versus other grades. There are significant and positive effects on the probability of earning a ‘C’ or lower average for all the drinking variables.

Table 7. IV/ OLS estimates of drinking on the probability of a 'C' or lower
All three youth experience instruments
(n=18,231)

Alcohol variables	IV	OLS
number of days drank-past year	0.004*	0.001*
Marginal Effect Standard Error	(0.0003)	(0.00007)
P-value of overidentification test	0.048	
Hausman statistic	10.286*	
number of drinks in past month	0.003*	0.0005*
Marginal Effect Standard Error	(0.0005)	(0.00008)
P-value of overidentification test	0.041	
Hausman statistic	9.704*	
binge drinking	0.429*	0.123*
Marginal Effect Standard Error	(0.036)	(0.009)
P-value of overidentification test	0.021	
Hausman statistic	9.537*	
abuse/ dependence on alcohol	0.689*	0.111*
Marginal Effect Standard Error	(0.067)	(0.011)
P-value of overidentification test	0.063	
Hausman statistic	9.124*	

*Statistically significant at 1%

Each daily increase in the number of past year drinking days raises the probability of having a ‘C’ or lower average by 0.004, while the probability is raised by 0.003 for each extra drink consumed in the previous month. If the respondent drinks 52 additional days in the past year, the probability of a ‘C’ or lower average rises by 0.21 points. And if 30 more drinks are consumed by the student in the past month, the probability of having a ‘C’ or lower grade is raised by 0.09.

For those engaging in binge or abusive/dependent drinking, there is a significant positive effect on the probability the student has a depressed grade point average in the current period. For binge drinkers there is an associated elevation in the probability of having a ‘C’ average of 0.43. For those categorized as abusive/dependent on alcohol, the probability of obtaining a ‘C’ or lower average is raised by 0.69.

The p-values of the overidentification tests afford little support for the assumption of exogeneity. Only for the abuse/ dependence indicator is instrument exogeneity not rejected at the 5 percent level. The Hausman coefficients, however, show that there are statistically significant differences between IV and OLS estimates.

The estimated effects for binge drinking and abuse/dependence are quite large. There may be large opportunity costs associated with this intense drinking, especially at abuse and dependence levels, which drastically undercut academic achievement. Thus, grades are dramatically lower, and possibly high failure rates may account for some of this.

Instrument Robustness and the Probability of a 'C' or Lower Average

Table 8 shows the results of regressions performed with varying pairs of instruments. The analysis is conducted to determine if there is any sensitivity in the results to changes in the instrument set.

Table 8. IV estimates of drinking on the probability of a 'C' using IV pairs
(n=18,231)

	parent disapprove and risk	risk and peer use	parent disapprove and peer use
Alcohol variables			
number of days drank-past year	0.004*	0.004*	0.003*
Marginal Effect Standard Error	(0.0003)	(0.0004)	(0.0003)
P-value of overidentification test	0.033	0.860	0.055
Hausman statistic	7.912*	8.236*	7.311*
Coefficient (Standard Error) of omitted IV	0.008 (0.009)	-0.038 (0.018)	-0.017 (0.012)
number of drinks in past month			
Marginal Effect Standard Error	0.003*	0.006*	0.005*
Marginal Effect Standard Error	(0.0006)	(0.0007)	(0.0006)
P-value of overidentification test	0.216	0.275	0.028
Hausman statistic	6.905*	8.231*	8.010*
Coefficient (Standard Error) of omitted IV	0.007 (0.008)	-0.031 (0.017)	-0.005 (0.013)
binge drinking	0.386*	0.437*	0.376*
Marginal Effect Standard Error	(0.040)	(0.050)	(0.044)
P-value of overidentification test	0.018	0.009	0.005
Hausman statistic	7.325*	6.981*	6.423*
Coefficient (Standard Error) of omitted IV	0.020 (0.008)	-0.003 (0.015)	-0.024 (0.012)
abuse/ dependence on alcohol	0.619*	0.634*	0.602*
Marginal Effect Standard Error	(0.077)	(0.084)	(0.084)
P-value of overidentification test	0.014	0.004	0.004
Hausman statistic	7.028*	6.589*	6.208*
Coefficient (Standard Error) of omitted IV	0.018 (0.009)	-0.016 (0.015)	-0.022 (0.013)

*Statistically significant at 1%

Again, the IV estimates are similar regardless of instrument choice. For the drinking measures, the overidentification tests reject instrument exogeneity at the 10 percent level, save for past year drinking in the risk and peer use specification, and past month drinking except when risk is included in the grade equation. For binge drinking and abuse/dependence factors, the overidentification test results offer little evidence that the instruments are exogeneous.

Throughout the grade analysis, OLS parameter estimates consistently underestimate the magnitude of the negative effects. This could possibly be attributed to higher ability students drinking more or higher income students having more resources to devote both to drinking and their education. Random measurement error that IV corrects could also play a role.

B. Drinking and School Enrollment

This section presents results for the effect of youth drinking on the probability of school enrollment. This outcome variable is described in chapter four. The causal effect that drinking has on this variable is analyzed using the three instrumental variables also described in that chapter. The enrollment analysis is conducted utilizing a sample of high-school age students (16-19 years old) and college age students (18-25 years old). To more accurately determine the validity of the exclusion restrictions for the instruments, a robustness analysis is conducted to using various combinations of instruments. The analysis also discusses results of comparisons between IV and OLS parameter estimates. The discussion that follows focuses primarily on the effect of drinking on enrollment. Appendices 4 and 5 show the probit estimates for enrollment for the 18-25 year old sample and the 16-19 year old sample respectively. Appendices 6 and 7 show the coefficients and standard errors of all exogenous variables for the binge drinking measure for the probability of enrollment for the 16-19 and 18-25 samples respectively.

First Stage Regression Results

Table 9 presents the results of the probit and first stage regressions of the drinking measures on the instruments for the 18-25 year old age group.

Table 9. Probit/ First stage estimates for enrollment (18-25 years old)
(n=28,065)

exogenous variables	number of days drank in past year	number of drinks in past month	Binge drinking	Abuse/ Dependence on alcohol
Risk of bodily harm from drinking	-34.717	-19.622	-0.242	-0.096
	(0.971)	(0.843)	(0.006)	(0.005)
Respondent states religion is important in life	-3.598	-0.954	-0.010	-0.008
	(1.221)	(1.060)	(0.008)	(0.006)
Respondent states religion influences decisions	-13.276	-8.946	-0.106	-0.041
	(1.156)	(1.044)	(0.008)	(0.007)
F stat/ chi2-coefficient of joint significance	564.78	247.24	1671.970	497.940
P-value of significance level	(0.0000)	(0.0000)	(0.0000)	(0.0000)
predicted drinking coefficient			1.006	0.99
			(0.022)	(0.040)

Of those who perceive that there is moderate to great risk of harm from consuming 4-5 drinks almost every day, the number of days drinking occurred in the past year is lowered by about 34 days. The number of drinks consumed in the past month is reduced by 20, while the likelihood of binge drinking in the last 30 days falls by 0.24. The likelihood of being categorized as abusive/dependent on alcohol falls by 0.10.

Importance of religious beliefs reduces all alcohol use measures. For those that report that religion is important in life, the number of days drinking occurred in the past year is lowered by 3.6 days. The number of drinks consumed in the past month is reduced by 0.95, while the probability of binge drinking in the last 30 days falls by 0.01. The likelihood of being categorized as abusive/dependent on alcohol falls by 0.008.

When religiosity impacts decisions, the effects on the drinking measures are more pronounced. The number of days drinking occurred in the past year is lowered by 13 days. The number of drinks consumed in the past month is reduced by about nine, while the probability of binge drinking in the last 30 days falls by 0.11. The likelihood of being categorized as abusive/dependent on alcohol falls by 0.04. The χ^2 coefficients and associated p-values indicate that the instruments are jointly significant for all the drinking measures. The predicted drinking coefficients in the binary drinking measure first stage regresssions are 1.01 for binge drinking in the past 30 days and 0.99 for abuse/ dependence on alcohol.

Table 10 presents the probit and first stage results for the instruments for the 16-19 year old age group.

Table 10. Probit/ First stage estimates for enrollment (16-19 years old)
(n=13,526)

exogeneous variables	number of days drank in past year	number of drinks in past month	Binge drinking	Abuse/ Dependence on alcohol
Risk of bodily harm from drinking	-20.650	-12.721	-0.170	-0.090
	(1.092)	(0.961)	(0.009)	(0.007)
Respondent states religion is important in life	-1.650	-2.780	-0.013	-0.007
	(1.346)	(1.180)	(0.009)	(0.007)
Respondent states religion influences decisions	-10.841	-4.908	-0.100	-0.049
	(1.260)	(1.109)	(0.009)	(0.007)
F stat/ chi2-coefficient of joint significance	185.76	87.46	664.860	299.180
P-value of significance level	(0.0000)	(0.0000)	(0.0000)	(0.0000)
predicted drinking coefficient			1.006	0.989
			(0.034)	(0.052)

For this age group, if moderate to great risk of harm from consuming 4-5 drinks almost every day is perceived, the number of days in which drinking occurred in the past year is lowered by 21 days. The number of drinks consumed in the past month is reduced by roughly13, while the probability of binge drinking in the last 30 days falls by 0.17. The likelihood of being categorized as abusive/dependent on alcohol also falls by 0.09.

Importance of religious beliefs and religiously influenced decisions reduce all alcohol use measures. For those who report that religion is important in life, the number of days in which drinking occurred in the past year is lowered by 1.7 days. The number of drinks consumed in the past month is reduced by 2.8, while the probability of binge drinking in the last 30 days falls by 0.01. The likelihood of being categorized as abusive/ dependent on alcohol falls by 0.007. When religiosity impacts decisions, the number of days in which drinking occurred in the past year is reduced by 11. The number of drinks consumed in the past month is reduced by five, while the probability of binge drinking in the last 30 days falls by 0.10. The likelihood of being categorized as abusive/dependent on alcohol falls by 0.05. The F statistics and χ^2 p-values signify support for the hypothesis of joint instrument significance for all the drinking measures. The predicted drinking values are 1.01 for binge drinking in the past 30 days and 0.99 for abuse/ dependence on alcohol.

The Effects of Drinking on School Enrollment (16-19 sample)

Table 11 presents findings for school enrollment using all three instruments. The analysis is conducted utilizing the subsample of 16–19 year olds who have not graduated from high school.

For each daily increase in reported drinking, the probability of being enrolled is subsequently lowered by 0.001. If, for instance, the respondent reports drinking 52 days in the previous year, the likelihood of enrollment is diminished by approximately 0.052 compared to not drinking at all. For each drink consumed in the prior month the probability of enrollment is lowered by 0.001 percent. If the student reports consuming 30 drinks in the previous month, the probability of enrollment decreases by 0.03 points.

Binge drinking and abuse/dependence on alcohol further reduce the probability of enrollment. Binging reduces the probability of enrollment by 0.08. For students who have engaged in binge drinking, the probability of school enrollment declines by approximately 10 percent compared to not binging.

Table 11. IV/ OLS estimates of drinking on school enrollment (16-19 years old)
(all three instruments)
(n=13,526)

Alcohol variables	IV	OLS
number of days drank-past year	-0.001*	-0.0003*
Marginal Effect Standard Error	(0.0001)	(0.00003)
P-value of overidentification test	0.014	
Hausman statistic	-2.339**	
number of drinks in past month	-0.001*	-0.0002*
Marginal Effect Standard Error	(0.0003)	(0.00004)
P-value of overidentification test	0.011	
Hausman statistic	-3.106*	
binge drinking	-0.083*	-0.034*
Marginal Effect Standard Error	(0.023)	(0.005)
P-value of overidentification test	0.021	
Hausman statistic	-2.259**	
abuse/ dependence on alcohol	-0.178*	-0.018*
Marginal Effect Standard Error	(0.046)	(0.006)
P-value of overidentification test	0.017	
Hausman statistic	-3.545*	

*Statistically significant at 1%
**Statistically significant at 5%

For those classified as abusive/dependent with respect to alcohol, the probability of enrollment decreases by 0.18. Categorization as abusive/dependent reduces the probability of school enrollment by 21 percent.

There is again little evidence to support the hypothesis of instrument exogeneity for the three instrument specification. The p-values associated with the overidentification tests for all drinking measures indicate that instrument exogeneity is rejected even at the 5 percent level. For each drinking measure, the Hausman coefficient signifies that statistically significant differences prevail between IV and OLS estimation.

Overall, in the high school sample, there is a strong indication that drinking, possibly by raising the opportunity cost of high school education and impairing cognitive functioning, reduces enrollment in high school. And, considering the additional resources the student devotes toward drinking if the student binge drinks or is abusive/dependent on alcohol, there is compelling evidence that the probability of high school enrollment is largely and negatively impacted. Though overidentification tests cast doubt on instrument exogeneity, drinking effects are significant and rather sizeable. The following section further investigates exogeneity by conducting an analysis using differing pairs of instruments.

Instrument Robustness and School Enrollment (16-19 sample)

Table 12 shows the results of regressions performed with varying pairs of instruments. This is undertaken to determine if there is any sensitivity in the main results to changes in the instrument set. The excluded instrument is used as an explanatory variable.

Models in the first and third columns are preferred, based on large overidentification test p-values, to those in the 2nd column and in table 11. This means that religion being important should be used as an instrument with either of the other two instruments, but not both of them. The IV estimates vary widely across these two models, however, with coefficients in the third column being three to four times as large as those in the first column. Because the Hausman statistic in the first column, which offers the more conservative estimate, is never significant at even the 10 percent level, but is always larger than magnitude than the OLS estimate, one would have to conclude that the OLS estimate gives the best estimate of the causal effect of drinking on enrollment. Table 11 shows that the OLS estimates are significantly negative, but relatively small.

Table 12. IV estimates of drinking on enrollment using IV pairs (16-19 years old)
(n=13,526)

	religion important and risk	**religious decisions and risk**	**religion important & religious decisions**
Alcohol variables			
number of days drank-past year	-0.0004*	-0.0005*	-0.001*
Marginal Effect Standard Error	(0.0002)	(0.0002)	(0.0005)
P-value of overidentification test	0.599	0.029	0.876
Hausman statistic	-0.199	-1.293	-3.448*
Coefficient (Standard Error) of omitted IV	0.015 (0.005)	0.010 (0.005)	-0.027 (0.011)
number of drinks in past month	-0.0005	-0.001*	-0.002*
Marginal Effect Standard Error	(0.0004)	(0.0003)	(0.0007)
P-value of overidentification test	0.723	0.015	0.514
Hausman statistic	-0.999	-1.921**	-3.550*
Coefficient (Standard Error) of omitted IV	0.015 (0.005)	0.009 (0.005)	-0.029 (0.011)
binge drinking	-0.035	-0.062*	-0.145*
Marginal Effect Standard Error	(0.029)	(0.026)	(0.040)
P-value of overidentification test	0.572	0.033	0.336
Hausman statistic	-0.116	-1.206	-2.836*
Coefficient (Standard Error) of omitted IV	0.016 (0.005)	0.010 (0.005)	-0.017 (0.008)
abuse/ dependence on alcohol	-0.104**	-0.145*	-0.371*
Marginal Effect Standard Error	(0.056)	(0.050)	(0.091)
P-value of overidentification test	0.468	0.028	0.905
Hausman statistic	-1.570	-2.577*	-3.879*
Coefficient (Standard Error) of omitted IV	0.013 (0.005)	-0.009 (0.006)	-0.028 (0.010)

*Statistically significant at 1%
**Statistically significant at 10%

The Effects of Drinking on School Enrollment (18-25 sample)

Table 13 shows the findings for enrollment for the 18-25 age group while employing all three instruments.

For each daily increase in reported past year drinking, the probability of being enrolled is subsequently lowered by 0.0004. While for each additional drink increase in the number of drinks the respondent consumed in the past month, the probability of enrollment is lowered by 0.0008. The probability of enrollment is reduced by 0.06 percentage points for those that report binge drinking in the previous 30 day period. For binge drinkers in this sample, the probability of school enrollment is reduced by 14 percent compared to non-binge drinkers. For those categorized as abusive/dependent on alcohol, the probability of enrollment falls by approximately 0.10 points, i.e. 23 percent compared to those not abusive/dependent.

Table 13. IV/ OLS estimates of drinking on school enrollment (18-25 years old)
(all three instruments)
(n=28,065)

Alcohol variables	IV	OLS
number of days drank-past year	-0.0004*	-0.0001*
Marginal Effect Standard Error	(0.0001)	(0.00003)
P-value of overidentification test	0.000	
Hausman statistic	-2.865*	
number of drinks in past month	-0.0008*	-0.0003
Marginal Effect Standard Error	(0.0002)	(0.00003)
P-value of overidentification test	0.000	
Hausman statistic	-3.609*	
binge drinking	-0.063*	-0.002
Marginal Effect Standard Error	(0.019)	(0.005)
P-value of overidentification test	0.000	
Hausman statistic	-3.522*	
abuse/ dependence on alcohol	-0.099**	-0.003
Marginal Effect Standard Error	(0.043)	(0.006)
P-value of overidentification test	0.000	
Hausman statistic	-2.457**	

*Statistically significant at 1%
**Statistically significant at 5%

The p-values associated with the overidentification tests indicate that the assumption of exogeneity is not at all supported when all three instruments are utilized. Results of the Hausman tests revel that there are statistically significant differences between IV and OLS estimates.

As with the high school sample, there is a strong indication that alcohol consumption, plausibly by raising the opportunity cost of post high school education, causally and negatively impacts college level enrollment. And, considering the resources the student devotes toward drinking, particularly if he/ she is abusive/ dependent on alcohol, the probability of post high school enrollment is also lessened to some degree.

Instrument Robustness and School Enrollment (18-25 sample)

Table 14 shows the results of regressions performed with varying pairs of instruments, which parallel those in table 12.

Table 14. IV estimates of drinking on enrollment using IV pairs (18-25 years old)
(n=28,065)

	religion important and risk	religious decisions and risk	religion important & religious decisions
Alcohol variables			
number of days drank-past year	-0.0001	-0.0003**	-0.001***
Marginal Effect Standard Error	(0.0001)	(0.0001)	(0.0003)
P-value of overidentification test	0.684	0.000	0.220
Hausman statistic	-1.790***	-1.081	-4.924*
Coefficient (Standard Error) of omitted IV	0.025 (0.006)	0.012 (0.005)	0.055 (0.013)
number of drinks in past month	-0.0002	-0.0006**	-0.002**
Marginal Effect Standard Error	(0.0002)	(0.0002)	(0.0005)
P-value of overidentification test	0.710	0.000	0.494
Hausman statistic	0.768	-2.578*	-3.609*
Coefficient (Standard Error) of omitted IV	0.025 (0.005)	0.012 (0.005)	0.051 (0.013)
binge drinking	0.012	-0.046**	-0.211**
Marginal Effect Standard Error	(0.023)	(0.020)	(0.045)
P-value of overidentification test	0.642	0.0002	0.186
Hausman statistic	0.685	-2.458**	-4.711*
Coefficient (Standard Error) of omitted IV	0.025 (0.006)	0.013 (0.006)	0.044 (0.012)
abuse/ dependence on alcohol	0.026	-0.053	-0.202**
Marginal Effect Standard Error	(0.049)	(0.046)	(0.090)
P-value of overidentification test	0.171	0.000	0.000
Hausman statistic	0.403	-1.311	-2.309**
Coefficient (Standard Error) of omitted IV	0.027 (0.006)	0.014 (0.006)	0.014 (0.010)

*Statistically significant at 1%
**Statistically significant at 5%
***Statistically significant at 10%

The first column, which uses religion being important and risk as instruments and includes religious decisions in the grade equation, is the only model for which overidentification tests are always insignificant. It also yields the most conservative IV estimates. These again suggest that OLS is consistent and efficient, which in turn suggests that drinking does not significantly impact college enrollment for high school graduates.

C. Drinking and Absenteeism

This section presents results for the effect of youth drinking on the number of school days the student missed due to "skipping" classes and the days missed due to illness or injury for students currently enrolled in school. This outcome variable is described in more detail in chapter four. The analysis is conducted using a sample of high school age students of 16 to 19 years old and college age students of 18 to 25 years old. The causal effect that drinking has on this variable is analyzed using the three instrumental variables listed above. The main results of the IV analysis are also compared with parameter estimates obtained using OLS methodology. The discussion that follows concentrates on the effects of alcohol consumption. Appendices 8 and 9 present all probit estimates for the 18-25 and 16-19 year old age groups respectively. Appendices 10 and 11 present coefficients and standard errors of all explanatory variables for both absenteeism factors, with binge drinking as the selected alcohol use measure for both the 16-19 and 18-25 samples respectively.

First Stage Regression Results

Table 15 presents the probit and first stage results for the sample of 18-25 year old respondents who are enrolled in college.

Table 15. Probit/ First stage estimates for absenteeism (18-25 years old)
(n=8,817)

exogeneous variables	number of days drank in past year	number of drinks in past month	Binge drinking	Abuse/ Dependence on alcohol
Risk of bodily harm from drinking	-34.537	-22.212	-0.283	-0.128
	(1.523)	(1.398)	(0.011)	(0.009)
Respondent states religion is important in life	-4.510	-0.517	-0.020	-0.006
	(1.983)	(1.821)	(0.015)	(0.011)
Respondent states religion influences decisions	-13.585	-8.587	-0.126	-0.054
	(1.874)	(1.726)	(0.015)	(0.011)
F stat/ chi2-coefficient of joint significance	253.33	104.62	668.240	255.250
P-value of significance level	(0.0000)	(0.0000)	(0.0000)	(0.0000)
predicted drinking coefficient			1.011	0.999
			(0.035)	(0.056)

For this age group, of those who perceive that there is moderate to great risk of harm from consuming 4-5 drinks almost every day, the number of days drinking occurred in the past year is lowered by 34 days. The number of drinks consumed in the past month is reduced by 22, while the likelihood of binge drinking in the last 30 days falls by 0.28. The likelihood of being categorized as abusive/dependent on alcohol falls by 0.13.

Importance of religious beliefs reduces all alcohol consumption measures. For respondents who report that religion is important in life, the number of days drinking occurred in the past year is lowered by 4.5 days. The number of drinks consumed in the past month is reduced by 0.5, while the probability of binge drinking in the last 30 days falls by 0.02. The likelihood of being categorized as abusive/ dependent on alcohol falls by approximately 0.006.

When religiosity impacts decisions, the effects on drinking are more pronounced than the effects when importance of religious beliefs is utilized as an IV. The number of days drinking occurred in the past year is lowered by 13.6 days. The number of drinks consumed in the past month is reduced by 8.6, while the probability of binge drinking in the last 30 days falls by 0.13. The likelihood of being categorized as abusive/dependent on alcohol declines by 0.054. The F statistics and χ^2 coefficients and associated p-values indicate the instruments are jointly significant for all the drinking measures. The predicted drinking coefficients in the binge drinking and abuse/dependence first stage regressions are 1.01 and 1.00, respectively.

Table 16 presents the probit and first stage results for the sample of 16-19 year old high school enrollees.

Table 16. Probit/ First stage estimates for absenteeism (16-19 years old)
(n=10,039)

exogeneous variables	number of days drank in past year	number of drinks in past month	Binge drinking	Abuse/ Dependence on alcohol
Risk of bodily harm from drinking	-19.795	-13.035	-0.171	-0.091
	(1.166)	(1.033)	(0.010)	(0.008)
Respondent states religion is important in life	-9.798	-3.138	-0.020	-0.013
	(1.346)	(1.277)	(0.010)	(0.008)
Respondent states religion influences decisions	-3.020	-4.301	-0.090	-0.049
	(1.442)	(1.192)	(0.010)	(0.008)
chi2-coefficient of joint significance	153.670	77.030	515.080	260.640
P-value of significance level	(0.0000)	(0.0000)	(0.0000)	(0.0000)
predicted drinking coefficient			1.027	1.007
			(0.038)	(0.055)

For respondents who agreed that there is moderate to great risk of harm from consuming 4-5 drinks almost every day, the number of days drinking occurred in the past year is lowered by about 20 days. The number of drinks consumed in the past month is reduced by 13, while the probability of binge drinking in the last 30 days falls by 0.17 percentage points. The likelihood of being categorized as abusive/dependent on alcohol also falls by 0.09 points.

Importance of religious beliefs and religiously influenced decisions reduce all alcohol use measures. For respondents who report that religion is important in life, the number of days drinking occurred in the past year is lowered by about 9.8 days. The number of drinks consumed in the past month is reduced by 3.1, while the number of binge drinking episodes in the last 30 days falls by 0.02. The likelihood of being categorized as abusive/ dependent on alcohol falls by 0.013.

When religiosity impacts decisions, the number of days drinking occurs in the past year is reduced by three days. The number of drinks consumed in the past month is reduced by about 4.3, while the probability of binge drinking in the last 30 days falls by 0.09. The likelihood of being categorized as abusive/dependent on alcohol falls by 0.05. Again, the F statistics and χ^2 p-values indicate support for the hypothesis of joint instrument significance for all the drinking measures. The predicted drinking coefficients in the binary measure first stage models are 1.03 for binge drinking in the past 30 days and 1.01 for abuse/dependence.

The Effects of Drinking on Absenteeism (16-19 sample)

The findings in table 17 show that youth drinking among 16 to 19 year old high school students leads to increases in absenteeism reported in the past 30 days.

The regression results show that an additional day increase in the number of past year drinking days elevates days skipped by 0.013 and days missed because of illness by approximately 0.007, relative to refraining from drinking. For each additional drink increase in the number of drinks the respondent consumed in the past month, days missed because of skipping rise by 0.02 and days missed due to illness increase by 0.01. An additional day of drinking in the prior month leads to a 2.3 percent increase in days skipped and an additional one percent increase in sick days.

Binge drinking and abuse/dependence on alcohol further increase truancy. For students who have engaged in binge drinking, the number of days missed due to skipping is elevated by 1.42 days. For those classified as abusive/dependent with respect to alcohol use, the amount of days missed due to skipping escalates to 2.6 days per month compared to those not abusive/dependent. School days missed due to illness rise by 0.62 days for binge drinkers and 1.2 days for those who are alcohol abusive/dependent.

Table 17. IV/ OLS estimates of drinking on absenteeism (16-19 years old)
(all three instruments)
(n=10,039)

	days missed due to skipping		days missed due to illness	
Alcohol variables	**IV**	**OLS**	**IV**	**OLS**
number of days drank-past year	0.013*	0.004*	0.007*	0.002*
Marginal Effect Standard Error	(0.001)	(0.0003)	(0.002)	(0.0004)
P-value of overidentification test	0.008		0.004	
Hausman statistic	5.102*		2.278**	
number of drinks in past month	0.020*	0.003*	0.010*	0.002*
Marginal Effect Standard Error	(0.003)	(0.0003)	(0.003)	(0.0005)
P-value of overidentification test	0.000		0.001	
Hausman statistic	5.489*		2.548**	
binge drinking	1.418*	0.630*	0.622*	0.332*
Marginal Effect Standard Error	(0.186)	(0.041)	(0.197)	(0.056)
P-value of overidentification test	0.007		0.003	
Hausman statistic	4.681*		1.978**	
abuse/ dependence on alcohol				
Marginal Effect Standard Error	2.616*	0.727*	1.202*	0.441*
Marginal Effect Standard Error	(0.360)	(0.051)	(0.368)	(0.070)
P-value of overidentification test	0.016		0.005	
Hausman statistic	5.412*		2.183**	

*Statistically significant at 1%
**Statistically significant at 5%

The overidentification test results are somewhat weak for this specification. For all drinking measures, instrument exogeneity is rejected at the 5 percent level. The Hausman tests all afford the same general result for all drinking measures and both truancy variables: statistically significant differences are present between IV and OLS estimates.

Instrumental Variable Robustness and Absenteeism (16-19 sample)

The effects of drinking on absenteeism using differing combinations of instruments are shown in table 18. The estimates in the first column are uniformly, and uniquely, reliable with regard to having overidentification test p-values well above 0.1. This again points to religious decisions being a poor choice for use as an instrument. With religious decisions included in the absenteeism equations, the IV estimates are again conservative with respect to the others in table 18 as well as those in table 17.

For skipping, these estimates are all significantly positive and distinct from the corresponding OLS estimates. This implies that drinking among high school students raises the propensity to skip classes. OLS estimates appear to have a positive bias. Again, it could be that higher ability and income students drink more and skip school less, or that measurement error in the OLS estimates imparts severe downward bias.

For illness, the estimates in the first column are insignificant for all drinking measures. The days and drinks coefficients are identical to those under OLS, so one could still legitimately conclude that additional drinking days or monthly drinks induces school absences due to illness, though not nearly as much as they increase days skipped. In contrast, the IV estimates for the binary drinking variables are smaller in magnitude than are the OLS estimates and statistically insignificant, which suggests that binge drinking and abuse/dependence do not increase illness-induced school absences. A lack of effect of heavy drinking, though, is inconsistent with a significant effect of an additional day of drinking or drink. Thus, the safest inference to make is that drinking does not necessarily causally influence school absences arising from illness.

Table 18. IV estimates of drinking on absenteeism using IV pairs (16-19 years old)
(n=10,039)

	days missed due to skipping		
	religion important and risk	**religious decisions and risk**	**religion important & religious decisions**
Alcohol variables			
number of days drank-past year	0.009*	0.012*	0.019*
Marginal Effect Standard Error	(0.002)	(0.002)	(0.003)
P-value of overidentification test	0.233	0.002	0.092
Hausman statistic	2.692*	4.107*	4.769*
Coefficient (Standard Error) of omitted IV	-0.126 (0.045)	-0.018 (0.047)	0.188 (0.080)
number of drinks in past month	0.014*	0.019*	0.033*
Marginal Effect Standard Error	(0.003)	(0.003)	(0.006)
P-value of overidentification test	0.137	0.001	0.021
Hausman statistic	-3.267*	4.453*	4.322*
Coefficient (Standard Error) of omitted IV	-0.151 (0.044)	-0.026 (0.050)	0.230 (0.111)
binge drinking	1.059*	1.365*	1.862*
Marginal Effect Standard Error	(0.223)	(0.220)	(0.293)
P-value of overidentification test	0.247	0.001	0.007
Hausman statistic	2.249**	3.694*	4.681*
Coefficient (Standard Error) of omitted IV	-0.129 (0.044)	-0.026 (0.045)	0.127 (0.068)
abuse/ dependence on alcohol	2.015*	2.583*	3.484*
Marginal Effect Standard Error	(0.432)	(0.432)	(0.596)
P-value of overidentification test	0.207	0.004	0.082
Hausman statistic	3.090*	4.412*	4.704*
Coefficient (Standard Error) of omitted IV	-0.115 (0.046)	-0.008 (0.048)	0.140 (0.074)
	days missed due to illness		
	religion important and risk	**religious decisions and risk**	**religion important & religious decisions**
Alcohol variables			
number of days drank-past year	0.002	0.006**	0.016**
Marginal Effect Standard Error	(0.002)	(0.002)	(0.004)
P-value of overidentification test	0.306	0.001	0.098
Hausman statistic	0.141	1.556	3.214*
Coefficient (Standard Error) of omitted IV	-0.183 (0.064)	-0.038 (0.061)	0.277 (0.111)
number of drinks in past month	0.003	0.008**	0.027**
Marginal Effect Standard Error	(0.003)	(0.003)	(0.008)
P-value of overidentification test	0.282	0.000	0.032
Hausman statistic	0.288	1.612	3.038*
Coefficient (Standard Error) of omitted IV	-0.190 (0.060)	-0.049 (0.062)	0.304 (0.137)
binge drinking	0.046	0.499**	1.066**
Marginal Effect Standard Error	(0.253)	(0.234)	(0.362)
P-value of overidentification test	0.220	0.001	0.015
Hausman statistic	-0.975	0.933	2.176**
Coefficient (Standard Error) of omitted IV	-0.207 (0.064)	-0.061 (0.059)	0.127 (0.095)
abuse/ dependence on alcohol	0.143	1.003*	2.115**
Marginal Effect Standard Error	(0.484)	(0.445)	(0.679)
P-value of overidentification test	0.240	0.001	0.030
Hausman statistic	-0.563	1.343	2.520**
Coefficient (Standard Error) of omitted IV	-0.202 (0.066)	-0.051 (0.061)	0.147 (0.094)

*Statistically significant at 1%

**Statistically significant at 5%

The Effects of Drinking on Absenteeism (18-25 sample)

The findings in table 19 show that youth drinking among respondents 18 to 25 years old leads to increases in absenteeism. Overall there are positive effects on absenteeism due to "skipping" classes and because of illness or injury. The regression results show that the drinking coefficients are statistically significant for both absenteeism measures.

Table 19. IV/ OLS estimates of drinking on absenteeism (18-25 years old)
(all three instruments)
(n=8,817)

	days missed due to skipping		days missed due to illness	
Alcohol variables	**IV**	**OLS**	**IV**	**OLS**
number of days drank-past year	0.006*	0.003*	0.003*	0.001*
Marginal Effect Standard Error	(0.001)	(0.0003)	(0.001)	(0.0002)
P-value of overidentification test	0.005		0.170	
Hausman statistic	2.501**		1.361	
number of drinks in past month	0.011*	0.003*	0.004**	0.0004
Marginal Effect Standard Error	(0.002)	(0.0003)	(0.001)	(0.0003)
P-value of overidentification test	0.009		0.095	
Hausman statistic	3.657*		2.298**	
binge drinking	0.954*	0.398*	0.352**	0.130*
Marginal Effect Standard Error	(0.179)	(0.046)	(0.150)	(0.042)
P-value of overidentification test	0.011		0.086	
Hausman statistic	3.486*		1.642	
abuse/ dependence on alcohol	1.888*	0.626*	0.653**	0.302*
Marginal Effect Standard Error	(0.348)	(0.056)	(0.270)	(0.051)
P-value of overidentification test	0.010		0.122	
Hausman statistic	3.777*		1.383	

*Statistically significant at 1%
**Statistically significant at 5%

For each extra day of alcohol use in the past year, the number of days skipped increases by 0.006 and days missed due to illness increases by 0.003 days. For each additional drink increase in the number of drinks the respondent consumed in the past month, the days missed because of skipping and illness rise by 0.011 and 0.004

respectively. Respondents that had one drink in the prior month experience an approximate one percent increase in days skipped and an approximate one percent increase in sick days, compared to those who did not drink in the past month.

Binge drinking and abuse/dependence on alcohol further increase skipping and days missed because of illness. For students who have engaged in binge drinking, the number of days missed due to skipping is elevated by approximately one day and days missed due to illness rises by approximately one-third of a day. For those classified as abusive/dependent with respect to alcohol, the number of days missed due to skipping increases by approximately two days and days missed due to illness rises by approximately two-thirds of a day.

The p-values associated with the overidentification tests for the days missed due to skipping model offer little support for the assumption of instrument exogeneity. At the 5 percent level, instrument exogeneity is rejected for all drinking measures. Hausman tests generally show that there are statistically significant differences between IV and OLS estimates.

For days missed due to illness, the reverse is true. There is stronger support for the hypothesis of instrument exogeneity, which is never rejected at the 5 percent level. However, Hausman statistics show an insignificant difference between IV and OLS estimates at the 5 percent level except in the case of past month drinks.

Instrumental Variable Robustness and Absenteeism (18-25 sample)

The effects of drinking on absenteeism using differing combinations of instruments are shown in table 20. For skipping, overidentification test results are generally unconvincing, making it difficult to draw any firm conclusions. For illness, overidentification tests uniformly fail to reject the hypothesis that the instruments are valid, but the conservative first column estimates are not significantly different from OLS. Because they are larger than those from OLS, however, and OLS estimates are significantly positive, we can conclude that drinking raises illness-related absences among college students, but that these effects are relatively small.

Overall, the results offered in this chapter demonstrate that youth alcohol consumption impedes the acquisition of human capital. The probability of achieving an 'A' average diminishes with youth alcohol consumption, while the probability of earning a grade of 'C' or lower is actually elevated by youth drinking.

The effect drinking has on the other education outcomes is also significant. For the probability of enrollment, all drinking measures produce a negative impact for both the high school age and college age samples. The analysis also reveals that elevated alcohol consumption engenders increased "skipping" classes for those that are attending school. And the findings provide some evidence that alcohol consumption increases school days missed due to illness.

The robustness analyses suggest that these described effects are sensitive, to some degree, to the choice of instruments. Identification tests indicate that instrument pairs are more plausibly exogenous in some specifications than in others. Statistically significant differences exist among IV and OLS parameter estimates in some, but not all, specifications.

Table 20. IV estimates of drinking on absenteeism using IV pairs (18-25 years old)

(n=8,817)

	days missed due to skipping		
	religion important and risk	**religious decisions and risk**	**religion important & religious decisions**
Alcohol variables			
number of days drank-past year	0.005*	0.007*	0.009*
Marginal Effect Standard Error	(0.001)	(0.001)	(0.002)
P-value of overidentification test	0.009	0.002	0.003
Hausman statistic	1.254	2.526 **	2.363**
Coefficient (Standard Error) of omitted IV	-0.105 (0.052)	0.052 (0.054)	0.148 (0.109)
number of drinks in past month	0.008*	0.011*	0.023*
Marginal Effect Standard Error	(0.002)	(0.002)	(0.006)
P-value of overidentification test	0.035	0.002	0.082
Hausman statistic	2.507**	3.478*	2.990*
Coefficient (Standard Error) of omitted IV	-0.116 (0.051)	0.012 (0.053)	0.328 (0.168)
binge drinking	0.785*	0.963*	1.920*
Marginal Effect Standard Error	(0.187)	(0.189)	(0.430)
P-value of overidentification test	0.043	0.002	0.106
Hausman statistic	2.371**	3.339*	3.667*
Coefficient (Standard Error) of omitted IV	-0.109 (0.050)	0.010 (0.050)	0.319 (0.127)
abuse/ dependence on alcohol	1.579*	1.977*	3.165*
Marginal Effect Standard Error	(0.369)	(0.378)	(0.793)
P-value of overidentification test	0.019	0.003	0.023
Hausman statistic	2.707*	3.701*	3.254*
Coefficient (Standard Error) of omitted IV	-0.095 (0.051)	0.040 (0.052)	0.228 (0.123)
	days missed due to illness		
	religion important and risk	**religious decisions and risk**	**religion important & religious decisions**
Alcohol variables			
number of days drank-past year	0.002	0.002***	0.007*
Marginal Effect Standard Error	(0.001)	(0.001)	(0.002)
P-value of overidentification test	0.571	0.247	0.890
Hausman statistic	0.447	0.851	2.374**
Coefficient (Standard Error) of omitted IV	-0.086 (0.051)	-0.070 (0.049)	0.187 (0.107)
number of drinks in past month	0.002	0.003***	0.014*
Marginal Effect Standard Error	(0.003)	(0.002)	(0.005)
P-value of overidentification test	0.462	0.246	0.444
Hausman statistic	1.158	1.675***	2.280**
Coefficient (Standard Error) of omitted IV	-0.094 (0.049)	-0.084 (0.046)	0.249 (0.148)
binge drinking	0.199	0.278***	0.770***
Marginal Effect Standard Error	(0.171)	(0.158)	(0.369)
P-value of overidentification test	0.421	0.252	0.180
Hausman statistic	0.537	1.101	1.800***
Coefficient (Standard Error) of omitted IV	-0.099 (0.048)	-0.088 (0.045)	0.137 (0.118)
abuse/ dependence on alcohol	0.326	0.470***	1.051***
Marginal Effect Standard Error	(0.312)	(0.293)	(0.624)
P-value of overidentification test	0.415	0.220	0.144
Hausman statistic	0.128	0.630	1.229
Coefficient (Standard Error) of omitted IV	-0.100 (0.048)	-0.084 (0.046)	0.071 (0.109)

*Statistically significant at 1%

**Statistically significant at 5%

***Statistically significant at 10%

Chapter Six: Summary and Conclusions

The economic implications surrounding human capital have intrigued economists for centuries, even dating back to the classical school of economic thought. More specifically, the fields of health economics and labor economics have both examined how physical and mental health may impact human capital formation. Issues involving the consequences of substance use on human capital have been researched in economics only within the previous fifteen years, and the current research on alcohol use and human capital suffers from two important shortcomings. First, existing research focuses largely on alcohol use among *college* students, leaving the impact of drinking among *high school* students largely unaddressed. The literature on adolescence has provided some evidence that alcohol use begins in the early teen years for many students. Failure to examine the experience of high school students leaves an important gap in understanding the relation between alcohol consumption and educational achievement.

Second, while past research has established a negative link between drinking and educational achievement, many of these studies have not accounted for the possibility that the negative correlation between drinking and educational achievement may be the result of unobserved variables that cause simultaneous increases in drinking and reductions in educational achievement. And, for those studies that have incorporated adjustments for endogeneity, the analyses have been conducted utilizing instrumental variable procedures that have been subject to criticism.

The first shortcoming is addressed by use of data from the National Survey of Drug Use and Health (NSDUH). The survey contains many variables pertaining to the behavior and attitudes of students, especially with respect to alcohol use and educational achievement. There is an entire subset of the data devoted only to surveying 12 to 17 year old students and many data pertaining to older students is compiled as well. The NSDUH data also contain several potential variables that can serve as instruments, which is central to the empirical strategy employed in this dissertation. Despite these advantages, or perhaps because the existing research largely neglects high school students, the dataset has not been widely utilized by other researchers in this topic area.

The second deficiency is addressed in this study by the use of an instrumental variable estimation technique. Specifically, the technique of instrumental variables is a statistical method designed to estimate the causal impact an independent variable has on a dependent variable when omitted variable bias and/ or reverse causation is present. By employing such a technique the researcher can more accurately gauge the causal effect the independent variable has on the dependent variable. Thus, the potential problem of falsely concluding that a causal negative relationship exists between drinking and education is mitigated, as are potential biases in parameter estimates.

This dissertation investigates both of these shortcomings for the purpose of improving and extending empirical knowledge about the consequences of alcohol consumption on educational achievement and to derive more accurate estimation results. The consequences of alcohol consumption for educational performance are illustrated first in the IV regression estimates using the probabilities of earning an ‘A’ or a ‘C’ or lower grade as educational outcomes. Results show a strong negative effect of drinking on the probability of earning an ‘A’ and fairly strong positive effects on the likelihood of obtaining a ‘C’ or below. Overidentification tests offer little evidence to support the hypothesis of exogeneity. Results derived from robustness analyses do provide some support for the assumption of instrument exogeneity when the outcome variable is the probability of earning an ‘A’. Evidence of exogeneity is weaker when the outcome variable is the probability of a ‘C’ or lower average. In addition, p-values of Hausman coefficients consistently show that there are significant differences among IV and OLS estimates, and in general OLS underestimates the impact drinking has on grade measures.

This study also analyzes other determinants of human capital development such as school enrollment and absenteeism due to “skipping” classes and reported illness. The effect of student drinking diminishes the probability the student is attending school. This result holds for both the 16-19 and 18-25 age groups.

Drinking has positive effects on absenteeism for both samples. For both absenteeism measures, overidentification tests for the main specification of the 16-19 year old sample offers weak evidence to support the hypothesis of instrument exogeneity. When IV sensitivity is evaluated however, some of the instrument specifications can afford evidence in support of the exogeneity assumption. For the 18-25 year old sample, the overidentification tests in the main specification indicate support for the hypothesis of instrument exogeneity for the days missed due to illness outcome; the support is weaker for the days missed due to skipping outcome. When instrument sensitivity is investigated, there is virtually no support for the exogeneity hypothesis with respect to days missed due to skipping, though there is stronger evidence of exogeneity with respect to days missed due to illness.

The fundamental results in this dissertation confirm the findings of some previous studies in the literature. For instance, the general results presented here parallel those of Wolaver (2002), Williams, et al. (2003) and DeSimone and Wolaver (2004) in that youth drinking has negative effects on grades. The negative effects on school enrollment outlined in this study confirm, at least in general terms, the Cook and Moore (1993) result that heavy drinking reduced subsequent schooling, and contravene the Dee and Evans (2003) conclusion that drinking had no distinguishable effect on educational measures such as high school and college enrollment. This study does corroborate Roebuck (2004), as one of his results show that alcohol consumed in a previous year reduces the probability of subsequent school enrollment.

Limitations

Two primary limitations hamper some of the findings. First, in some of the specifications, the impact drinking has on the educational outcome variable is quite large. This may give some researchers pause with regard to the plausibility of the magnitude of the parameter estimates. Another difficulty exists with respect to instrumental variable exogeneity. For grade probabilities, there is in general only weak support for the assumption of instrument exogeneity. For the enrollment outcome, there is some evidence to support instrument exogeneity for the both 16-19 year old and 18-25 year old subsamples. However, support for the exogeneity assumption is dependent on the variables selected as instruments. Generally, the same can be said for the absenteeism analysis.

Policy Implications

Overall, the results derived from the regressions directly imply that alcohol consumption on the part of teenagers and young adults up to age 25 harms human capital attainment. Therefore, any economic or other gains to individuals and society from accruing education could also be reduced. While other factors that reduce educational achievement are certain to interact with alcohol use, the direct impact of drinking is shown to have large effects.

The results have further implications that are applicable to policymaking. First, the estimated effects of drinking on schooling convey information regarding the external benefits on educational outcomes of policies that effectively inhibit alcohol consumption among youths. Identifying whether alcohol use directly leads to lower achievement or more destructive school-related behavior, or is merely spuriously correlated with worse educational outcomes through unobserved variables that influence both sets of behaviors, have been addressed and there exists evidence that causal effects are present. Policies that reduce heavy drinking also become more attractive because the benefits of such policies include the value to society of the resulting gains in educational performance.

Second, the effects of the variables that are utilized as instruments in the IV procedure provide an indication as to various specific endeavors that might be successful in reducing alcohol consumption among youth. Parental disapproval of drinking is negatively and strongly correlated with youth drinking, therefore programs encouraging parental discouragement of youth drinking could be beneficial. Given the discovered impact of religiosity on drinking behaviors, policies and/or social endeavors aimed at encouraging religious activity would have the added benefit of alcohol use reductions.

The importance of perceived risk in using alcohol suggests that campaigns that raise awareness of the risks inherent in consuming alcohol would be fruitful as a prospective policy tool. Strong relationships with perceived peer use of alcohol signal that the dissemination of information regarding actual drinking prevalence among students might be warranted. The significance of peer drinking measures would imply that any policy able to diminish underage alcohol use could have "social multiplier" effects that enhance the attractiveness of such policies.

Third, the finding of a significant causal effect of drinking on human capital accumulation has implications for several fields of economics such as labor and educational economics. These include, for example, reducing the unexplained variation across individuals in wage and earnings equations.

There are some avenues for future research that emanate from this study. If researchers can obtain data on other educational outcomes not covered in the NSDUH, the methodology presented in this dissertation should prove useful. For example, an analyst could evaluate the relation between teen drinking and performance on standardized tests such as the SAT and ACT. While the IV strategy employed met with limited success, further investigation of the causal effects drinking may have on an economic outcome variable should be accompanied by utilizing potentially more exogenous instruments.

Also, information provided by the relationship between youth drinking behaviors and variables such as peer alcohol use, parental disapproval of teen drinking, and religious influences shown in this study should prove useful in conducting cost/benefit analyses of government programs that aim to curb alcohol use among youths. Use of instrumental variable models similar to those in this dissertation, perhaps utilizing factors such as parental divorce and sibling attitudes and behaviors not available in the NSDUH, would likely provide valuable guidance to policymakers.

In sum, this study provides a plethora of information regarding the educational consequences stemming from alcohol use among teenagers and young adults. The findings presented indicate that youth drinking significantly and negatively effects human capital accumulation. Even in areas where the results are not as strong, there remains substantial information that addresses the determinants of youth drinking. The research questions presented in this dissertation also generate topics for future investigation in health economics and other related subject areas.

References

Becker, G. S. (1964). *Human Capital*: (Columbia University Press for the National Bureau of Economic Research, New York).

Ben-Porath Y.(!967). The production of human capital and the life cycle of earnings. *Journal of Political Economy* 75: 353:-367.

Bryant, A., Schulenberg, J., O'Malley, P., Bachman, J., & Johnston, L. (2003). How academic achievement, attitudes, and behaviors relate to the course of substance use during adolescence: A 6-year, multiwave longitudinal study. *Journal of Research on Adolescence*, 13(3), 361–397.

Cassady, P. (2001). Self-reported GPA and SAT scores: A methodological note. *Practical Assessment, Research & Evaluation*, 7(12).

Chatterji, P. (2003). Illicit drug use and educational attainment. Working paper 10045, National Bureau of Economic Research.

Chatterji, P. and DeSimone, J. (2005). Adolescent drinking and high school dropout. Working paper 11337, National Bureau of Economic Research.

Cook, P. J., & Moore, M. J. (1993). Drinking and schooling. *Journal of Health Economics*, 12(4), 411–429.

Dee, T. S., & Evans, W. N. (2003). Teen drinking and educational attainment: Evidence from two-sample instrumental variables estimates. *Journal of Labor Economics*, 21(1), 178–209.

DeSimone, J., & Wolaver, A. M. (2005). Drinking and academic performance in high school. Working Paper 11035, National Bureau of Economic Research.

Evans, W., Oates, W., & Schwab, R. (1992). Measuring peer group effects: A study of teenage behavior. *Journal of Political Economy*, 100(51), 966–991.

Gaviria, A., & Raphael, S. (2001). School-based peer effects and juvenile behavior. *Review of Economics and Statistics*, 83(2), 257–268.

Grant, B. F., Harford, T. C., & Grigson, M. B. (1988). Stability of alcohol consumption among youth: a national longitudinal survey. *Journal of Studies on Alcohol*, 49(3), 253–260.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy,* 80(2), 223-255.

Harrison, L., & Hughes, A. (1997). The validity of self-reported drug use: Improving the accuracy of survey estimates. *NIDA Research Monograph*, 167, 1–16.

Hausman, J. (1978). Specification tests in econometrics. *Econometrica*, 46(6), 1251-1271.

Hoyt, G., & Chaloupka, F. (1994). Effect of survey conditions on self-reported substance use. *Contemporary Economic Policy*, 12(3), 109–121.

Jeynes, W. (2002). The relationship between the consumption of various drugs by adolescents and their academic achievement. *American Journal of Drug and Alcohol Abuse*, 28(1), 15–35.

Johnston, L. D., O'Malley, M., & Bachman, J. G. (1988). Illicit drug use, smoking and drinking by America's high school students, college Students and young adults. Washington, DC: USGPO.

Koch, S. F., & Ribar, D. C. (2001). A siblings analysis of the effects of alcohol consumption onset on educational attainment. *Contemporary Economic Policy*, 19(2), 162–74.

Kremer, M., & Levy, D. (2003). Peer effects and alcohol use among college students. Working paper 9876, National Bureau of Economic Research.

Markowitz, S. (2001). The role of alcohol and drug consumption in determining physical fights and weapon carrying by teenagers. *Eastern Economic Journal*, 27(4), 409–432.

Midanik, L. (1988). Validity of self-reported alcohol use: a literature review and assessment. *British Journal of Addiction*, 83, 1019–1029.

Mincer, J. (1974). *Schooling, Experience and Earnings.* New York: National Bureau of Economic Research..

Mullahy, J., & Sindelar, J. L. (1994). Alcoholism and income: The role of indirect effects. *The Milbank Quarterly*, 72(2), 359–375.

Norton, E., Lindrooth, R., & Ennett, S. (1998). Controlling for the endogeneity of peer substance use on adolescent alcohol and tobacco use. *Health Economics*, 36(7), 439–453.

Reinisch, E. J., Bell, R. M., & Ellickson, P. (1991). How accurate are adolescent reports of drug use? Santa Monica, CA: Rand Corporation.

Roebuck, C., French, M., & Dennis, M. (2004). Adolescent marijuana use and school attendance. *Economics of Education Review*, 23(2), 133–141.

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3), 557–586.

Williams, J., Powell, L. M., & Wechsler, H. (2003). Does alcohol consumption reduce human capital accumulation? Evidence from the college alcohol study. *Applied Economics*, 35(10), 1227–1239.

Wolaver, A. M. (2002). Effects of heavy drinking in college on study effort, grade point average, and major choice. *Contemporary Economic Policy*, 20(4), 415–428.

Wooldridge, J. M. (2003). Introductory econometrics: a modern approach (2nd edition). South-Western College Publishing.

Yamada, T., Kendix, M., & Yamada, T. (1996). The impact of alcohol consumption and marijuana use on high school graduation. *Health Economics*, 5(1), 77–92.

Appendices

Appendix 1. Probit estimates for the probability of 'A' and 'C'
(n=18,231)

explantory variables	Pseudo $R^2 = 0.11$ number of days drank in past year	Pseudo $R^2 = 0.13$ number of drinks in past month	Pseudo $R^2 = 0.14$ Binge drinking	Pseudo $R^2 = 0.12$ Abuse/ Dependence on alcohol
Risk of bodily harm from drinking	-15.890 (0.894)	-10.420 (0.749)	-0.125 (0.008)	-0.078 (0.006)
Peer use of alcohol	11.949 (0.716)	6.241 (0.600)	0.086 (0.005)	0.054 (0.004)
Parental disapproval of alcohol use	-24.706 (0.716)	-14.201 (0.845)	-0.162 (0.005)	-0.074 (0.007)
Mother in household	0.012 (0.014)	-0.006 (0.011)	-0.009 (0.009)	-0.002 (0.006)
Father in household	-0.051 (0.010)	-0.038 (0.009)	-0.027 (0.007)	-0.016 (0.005)
Female	0.47 (0.007)	0.010 (0.006)	0.014 (0.005)	0.006 (0.003)
Grade in (10th grade)	0.027 (0.013)	0.002 (0.011)	0.002 (0.009)	-0.001 (0.006)
Grade in (11th grade)	0.052 (0.021)	0.028 (0.007)	0.013 (0.009)	-0.008 (0.006)
Grade in (12th grade)	-0.073 (0.017)	0.049 (0.015)	0.028 (0.011)	-0.006 (0.007)
Age of student (15 years old)	0.092 (0.035)	0.054 (0.012)	0.040 (0.011)	0.026 (0.008)
Age of student (16 years old)	0.153 (0.035)	0.117 (0.015)	0.096 (0.013)	0.063 (0.011)
Age of student (17 years old)	0.202 (0.016)	0.158 (0.016)	0.133 (0.014)	0.083 (0.011)
Race (African American)	0.153 (0.124)	0.112 (0.009)	0.096 (0.005)	0.054 (0.004)
Race (Native American)	0.015 (0.019)	-0.015 (0.004)	0.013 (0.020)	0.046 (0.002)
Race (Asian)	-0.227 (0.020)	-0.128 (0.013)	-0.090 (0.009)	-0.041 (0.007)
Race (non-white Hispanic)	-0.007 (0.124)	-0.014 (0.010)	-0.009 (0.007)	-0.004 (0.005)
Number in family	-0.016 (0.005)	-0.010 (0.004)	-0.003 (0.003)	-0.003 (0.002)
Number in family (>5)	-0.102 (0.022)	-0.060 (0.015)	-0.032 (0.013)	-0.011 (0.009)
Family income ($20,000-$49,999)	0.065 (0.012)	0.054 (0.010)	0.031 (0.008)	0.020 (0.006)
Family income ($50,000-$74,999)	0.068 (0.013)	0.064 (0.013)	0.043 (0.017)	0.020 (0.007)
Family income ($75,000 or more)	0.117 (0.014)	0.108 (0.013)	0.056 (0.017)	0.025 (0.007)
MSA segement with 1+ million persons	-0.013 (0.010)	-0.013 (0.008)	-0.015 (0.006)	-0.009 (0.004)
MSA segment of less than 1 million	-0.014 (0.009)	-.011 (0.008)	-0.009 (0.006)	-0.009 (0.004)
Move (number of times in last 5years)	0.027 (0.013)	0.015 (0.002)	0.008 (0.003)	0.009 (0.002)
Year 2002 indicator	-0.059 (0.008)	-0.051 (0.008)	-0.048 (0.005)	-0.030 (0.004)
Parents help with homework (always)	-0.133 (0.013)	-0.088 (0.006)	-0.057 (0.008)	-0.047 (0.005)
Parents help with homework (sometimes)	-0.081 (0.014)	-0.042 (0.010)	-0.014 (0.009)	-0.026 (0.004)
Parents help with homework (seldom)	-0.032 (0.016)	-0.019 (0.011)	-0.014 (0.009)	-0.014 (0.005)

(Standard errors are in parentheses)

Appendix 2. All IV estimates on the probability of an 'A' for binge drinking
(n=18,321)

Explanatory variables	IV coefficient (Marginal Effect SE)
binge drinking	-0.351 (0.030)
Mother in household	0.005 (0.008)
Father in household	0.010 (0.005)
Female	0.031 (0.003)
Grade in (10th grade)	0.045 (0.007)
Grade in (11th grade)	0.085 (0.009)
Grade in (12th grade)	0.109 (0.010)
Age of student (15 years old)	-0.042 (0.006)
Age of student (16 years old)	-0.062 (0.0008)
Age of student (17 years old)	-0.053 (0.009)
Race (African American)	-0.018 (0.007)
Race (Native American)	0.005 (0.0020)
Race (Asian)	0.012 (0.007)
Race (non-white Hispanic)	-0.006 (0.006)
Number in family	-0.003 (0.002)
Number in family (>5)	-0.015 (0.011)
Family income ($20,000-$49,999)	0.013 (0.007)
Family income ($50,000-$74,999)	0.038 (0.007)
Family income ($75,000 or more)	0.054 (0.007)
MSA segment with 1+ million persons	-0.000 (0.004)
MSA segment of less than 1 million	-0.000 (0.004)
Parents help with homework (always)	0.001 (0.007)
Parents help with homework (sometimes)	0.000 (0.008)
Parents help with homework (seldom)	0.000 (0.009)
Year 2002 indicator	-0.003 (0.003)

Standard errors are in parentheses.

Appendix 3. All IV estimates on the probability of an 'C' or lower for binge drinking (n=18,231)

Explanatory variables	IV coefficient (Marginal Effect SE)
binge drinking	0.429 (0.036)
Mother in household	-0.020 (0.013)
Father in household	-0.033 (0.010)
Female	-0.133 (0.009)
Grade in (10th grade)	-0.041 (0.010)
Grade in (11th grade)	-0.107 (0.013)
Grade in (12th grade)	-0.16 (0.015)
Age of student (15 years old)	0.065 (0.011)
Age of student (16 years old)	0.084 (0.014)
Age of student (17 years old)	0.172 (0.016)
Race (African American)	0.141 (0.012)
Race (Native American)	0.035 (0.019)
Race (Asian)	-0.033 (0.006)
Race (non-white Hispanic)	0.055 (0.011)
Number in family	0.014 (0.004)
Number in family (>5)	0.059 (0.020)
Family income ($20,000-$49,999)	-0.037 (0.011)
Family income ($50,000-$74,999)	-0.094 (0.013)
Family income ($75,000 or more)	-0.157 (0.012)
MSA segement with 1+ million persons	0.025 (0.009)
MSA segment of less than 1 million	0.027 (0.008)
Move (number of times in last 5 years)	0.025 (0.005)
Parents help with homework (always)	-0.045 (0.012)
Parents help with homework (sometimes)	0.018 (0.013)
Parents help with homework (seldom)	-0.001 (0.014)
Year 2002 indicator	0.005 (0.006)

Standard errors are in parentheses.

Appendix 4. Probit estimates for the probability of enrollment (18-25 years old)
(n=28,065)

explantory variables	Pseudo R^2 = .08 number of days drank in past year	Pseudo R^2 = .09 number of drinks in past month	Pseudo R^2 = .10 Binge drinking	Pseudo R^2 = .06 Abuse/ Dependence on alcohol
Risk of bodily harm from drinking	-34.717 (0.971)	-19.662 (0.843)	-0.242 (0.006)	-0.096 (0.005)
Respondent states religion is important in life	-3.598 (1.221)	-0.954 (1.060)	-0.010 (0.008)	-0.008 (0.006)
Respondent states religion influences decisions	-13.276 (1.156)	-8.946 (1.004)	-0.106 (0.008)	-0.041 (0.007)
Female	-0.020 (0.004)	-0.089 (0.006)	-0.168 (0.006)	-0.090 (0.005)
Race (African American)	-0.109 (0.007)	-0.130 (0.008)	-0.166 (0.008)	-0.062 (0.005)
Race (Native American)	-0.010 (0.021)	-0.056 (0.003)	-0.006 (0.014)	0.070 (0.021)
Race (Asian)	-0.190 (0.012)	-0.202 (0.001)	-0.179 (0.015)	-0.062 (0.001)
Race (non-white Hispanic)	-0.094 (0.007)	-0.109 (0.009)	-0.072 (0.008)	-0.014 (0.006)
Age of student (19 years old)	-0.023 (0.008)	-0.036 (0.109)	-0.038 (0.008)	-0.022 (0.006)
Age of student (20 years old)	-0.058 (0.007)	-0.070 (0.011)	-0.054 (0.008)	0.000 (0.006)
Age of student (21 years old)	-0.106 (0.007)	-0.174 (0.010)	-0.136 (0.007)	-0.035 (0.007)
Age of student (22-23 years old)	-0.058 (0.007)	-0.142 (0.007)	-0.096 (0.007)	-0.014 (0.007)
Age of student (24-25 years old)	-0.106 (0.007)	-0.125 (0.007)	-0.068 (0.007)	-0.011 (0.007)
Last grade completed (Freshman)	0.053 (0.006)	0.084 (0.008)	0.038 (0.009)	-0.015 (0.005)
Last grade completed (Sophomore/ Junior)	0.071 (0.006)	0.117 (0.009)	0.054 (0.008)	-0.025 (0.006)
Number in family	-0.030 (0.002)	-0.038 (0.003)	-0.024 (0.003)	-0.014 (0.002)
Number in family (>5)	-0.222 (0.015)	-0.240 (0.014)	-0.141 (0.013)	-0.060 (0.008)
Family income ($20,000-$49,999)	0.021 (0.005)	0.006 (0.0010)	-0.036 (0.008)	0.028 (0.005)
Family income ($50,000-$74,999)	0.049 (0.007)	0.037 (0.007)	-0.017 (0.009)	0.020 (0.006)
Family income ($75,000 or more)	0.064 (0.007)	0.072 (0.010)	0.022 (0.009)	0.001 (0.007)
MSA segment with 1+ million persons	0.028 (0.006)	0.043 (0.007)	0.002 (0.008)	-0.006 (0.005)
MSA segment of less than 1 million	0.034 (0.005)	0.045 (0.006)	0.021 (0.007)	0.009 (0.005)
Year 2002 indicator	0.004 (0.004)	0.005 (0.006)	0.001 (0.005)	0.011 (0.004)

Standard errors are in parentheses.

Appendix 5. Probit estimates for the probability of enrollment (16-19 years old)
(n=13,562)

explanatory variables	Pseudo $R^2 = 0.06$ number of days drank in past year	Pseudo $R^2 = 0.08$ number of drinks in past month	Pseudo $R^2 = 0.09$ Binge drinking	Pseudo $R^2 = 0.06$ Abuse/ Dependence on alcohol
Risk of bodily harm from drinking	-20.650 (1.092)	-12.721 (0.961)	-0.170 (0.009)	-0.090 (0.007)
Respondent states religion is important in life	-1.650 (1.346)	-2.780 (1.180)	-0.013 (0.009)	-0.007 (0.007)
Respondent states religion influences decisions	-10.841 (1.260)	-4.908 (1.109)	-0.100 (0.009)	-0.049 (0.007)
Female	-0.050 (0.007)	-0.002 (0.008)	-0.039 (0.007)	-0.009 (0.005)
Race (African American)	-0.127 (0.011)	-0.112 (0.012)	0.119 (0.008)	-0.059 (0.006)
Race (Native American)	-0.010 (0.040)	-0.003 (0.036)	0.048 (0.033)	-0.114 (0.002)
Race (Asian)	-0.224 (0.022)	-0.154 (0.018)	-0.130 (0.015)	-0.060 (0.010)
Race (non-white Hispanic)	-0.019 (0.011)	-0.010 (0.012)	-0.008 (0.010)	0.005 (0.014)
Age of student (17 years old)	0.143 (0.001)	-0.144 (0.010)	-0.149 (0.014)	-0.065 (0.008)
Age of student (18 years old)	0.083 (0.013)	-0.101 (0.010)	-0.102 (0.013)	0.036 (0.009)
Age of student (19 years old)	0.059 (0.013)	-0.065 (0.010)	-0.061 (0.015)	0.027 (0.006)
Last grade completed (9th grade)	0.012 (0.010)	0.007 (0.008)	0.018 (0.009)	-0.001 (0.005)
Last grade completed (10th grade)	-0.036 (0.012)	-0.16 (0.013)	-0.004 (0.011)	-0.007 (0.008)
Last grade completed (11th grade)	-0.045 (0.012)	-0.048 (0.012)	-0.017 (0.012)	-0.012 (0.012)
Number in family	-0.024 (0.005)	-0.022 (0.004)	-0.012 (0.007)	-0.091 (0.005)
Number in family (>5)	-0.155 (0.023)	-0.115 (0.016)	-0.073 (0.014)	-0.036 (0.010)
Family income ($20,000-$49,999)	0.055 (0.012)	0.049 (0.001)	-0.028 (0.010)	-0.017 (0.005)
Family income ($50,000-$74,999)	0.046 (0.007)	0.041 (0.012)	-0.039 (0.013)	-0.008 (0.006)
Family income ($75,000 or more)	0.071 (0.014)	0.084 (0.012)	0.051 (0.013)	0.015 (0.008)
MSA segement with 1+ million persons	0.000 (0.011)	-0.016 (0.009)	-0.017 (0.008)	-0.012 (0.006)
MSA segment of less than 1 million	0.000 (0.011)	-0.016 (0.009)	-0.013 (0.008)	-0.007 (0.005)
Year 2002 indicator	0.021 (0.008)	0.019 (0.008)	0.008 (0.007)	0.001 (0.005)

Standard errors are in parentheses.

Appendix 6. All IV estimates on the probability of enrollment for binge drinking
(n=13,526)

16-19 sample	
Explanatory variables	**IV coefficient (Marginal Effect SE)**
binge drinking	-0.083 (0.023)
Female	-0.014 (0.004)
Race (African American)	0.309 (0.007)
Race (Native American)	0.001 (0.027)
Race (Asian)	-0.033 (0.009)
Race (non-white Hispanic)	-0.001 (0.014)
Age of student (17 years old)	0.681 (0.016)
Age of student (18 years old)	0.548 (0.016)
Age of student (19 years old)	03661 (0.018)
Last grade completed (9th grade)	0.120 (0.019)
Last grade completed (10th grade)	0.186 (0.019)
Last grade completed (11th grade)	0.334 (0.019)
Number in family	-0.005 (0.002)
Number in family (>5)	-0.001 (0.012)
Family income ($20,000-$49,999)	0.036 (0.007)
Family income ($50,000-$74,999)	-0.081 (0.008)
Family income ($75,000 or more)	-0.093 (0.007)
MSA segement with 1+ million persons	-0.020 (0.002)
MSA segment of less than 1 million	-0.017 (0.005)
Year 2002 indicator	-0.002 (0.004)

Appendix 7. All IV estimates on the probability of enrollment for binge drinking
(n=28,065)

	18-25 sample
Explanatory variables	**IV coefficient (Marginal Effect SE)**
binge drinking	-0.063 (0.019)
Female	-0.004 (0.006)
Race (African American)	-0.008 (0.008)
Race (Native American)	-0.017 (0.022)
Race (Asian)	0.140 (0.015)
Race (non-white Hispanic)	-0.058 (0.007)
Age of student (19 years old)	-0.243 (0.009)
Age of student (20 years old)	-0.420 (0.009)
Age of student (21 years old)	-0.501 (0.009)
Age of student (22-23 years old)	-0.620 (0.008)
Age of student (24-25 years old)	-0.698 (0.008)
Last grade completed (Freshman)	0.358 (0.007)
Last grade completed (Sophomore/ Junior)	0.508 (0.006)
Number in family	-0.003 (0.002)
Number in family (>5)	-0.055 (0.012)
Family income ($20,000-$49,999)	0.108 (0.005)
Family income ($50,000-$74,999)	0.045 (0.008)
Family income ($75,000 or more)	0.010 (0.008)
MSA segement with 1+ million persons	0.047 (0.006)
MSA segment of less than 1 million	0.034 (0.006)
Year 2002 indicator	-0.005 (0.004)

Standard errors are in parentheses.

Appenidx 8. Probit estimates for absenteeism (18-25 years old)
(n=8,817)

explanatory variables	Pseudo $R^2 = 0.08$ number of days drank in past year	Pseudo $R^2 = 0.10$ number of drinks in past month	Pseudo $R^2 = 0.13$ Binge drinking	Pseudo $R^2 = 0.06$ Abuse/ Dependence on alcohol
Risk of bodily harm from drinking	-34.537 (1.523)	-22.212 (1.398)	-0.283 (0.011)	-0.128 (0.009)
Respondent states religion is important in life	-4.510 (1.983)	-0.517 (1.821)	-0.020 (0.015)	-0.006 (0.011)
Respondent states religion influences decisions	-13.585 (1.874)	-8.587 (1.726)	-0.126 (0.015)	-0.054 (0.011)
Female	0.047 (0.009)	-0.000 (0.006)	-0.088 (0.011)	-0.046 (0.008)
Race (African American)	-0.120 (0.015)	-0.195 (0.017)	-0.238 (0.013)	-0.086 (0.010)
Race (Native American)	-0.005 (0.045)	-0.099 (0.023)	-0.059 (0.014)	0.021 (0.042)
Race (Asian)	-0.191 (0.026)	-0.239 (0.020)	-0.207 (0.020)	-0.058 (0.016)
Race (non-white Hispanic)	-0.059 (0.016)	-0.079 (0.019)	-0.073 (0.017)	-0.039 (0.012)
Age of student (19 years old)	0.017 (0.007)	0.043 (0.016)	0.033 (0.017)	0.026 (0.013)
Age of student (20 years old)	0.060 (0.013)	0.086 (0.019)	0.055 (0.013)	0.016 (0.016)
Age of student (21 years old)	0.129 (0.014	0.230 (0.014	0.192 (0.023)	0.056 (0.014)
Age of student (22-23 years old)	0.084 (0.014)	0.140 (0.020)	0.093 (0.023)	0.046 (0.019)
Age of student (24-25 years old)	0.081 (0.016)	0.176 (0.022)	0.063 (0.028)	0.040 (0.023)
Last grade completed (Freshman)	0.040 (0.012)	0.069 (0.016)	0.043 (0.012)	0.010 (0.012)
Last grade completed (Sophomore/ Junior)	0.056 (0.014)	0.103 (0.017)	0.032 (0.014)	0.018 (0.013)
Number in family	-0.025 (0.002)	-0.039 (0.006)	-0.020 (0.003)	-0.013 (0.004)
Number in family (>5)	-0.239 (0.015)	-0.243 (0.014)	-0.146 (0.013)	-0.052 (0.018)
Family income ($20,000-$49,999)	0.016 (0.005)	0.064 (0.001)	-0.093 (0.008)	0.042 (0.009)
Family income ($50,000-$74,999)	0.005 (0.007)	0.044 (0.007)	-0.074 (0.009)	0.034 (0.006)
Family income ($75,000 or more)	0.017 (0.007)	0.007 (0.010)	0.051 (0.011)	0.023 (0.011)
MSA segement with 1+ million persons	0.016 (0.012)	0.014 (0.015)	0.020 (0.015)	-0.018 (0.010)
MSA segment of less than 1 million	0.016 (0.011)	0.030 (0.014)	0.007 (0.013)	0.004 (0.010)
Year 2002 indicator	0.002 (0.008)	0.006 (0.011)	0.013 (0.015)	0.014 (0.008)

Standard errors are in parentheses.

Appendix 9. Probit estimates for absenteeism (16-19 years old)
(n=10,039)

explanatory variables	Pseudo $R^2 = 0.06$ number of days drank in past year	Pseudo $R^2 = 0.07$ number of drinks in past month	Pseudo $R^2 = 0.09$ Binge drinking	Pseudo $R^2 = 0.06$ Abuse/ Dependence on alcohol
Risk of bodily harm from drinking	-19.795 (1.166)	-13.035 (1.033)	-0.171 (0.010)	-0.091 (0.008)
Respondent states religion is important in life	-9.798 (1.346)	-3.138 (1.277)	-0.020 (0.010)	-0.013 (0.008)
Respondent states religion influences decisions	-3.020 (1.442)	-4.301 (1.192)	-0.090 (0.010)	-0.049 (0.008)
Female	0.076 (0.010)	-0.009 (0.009)	-0.028 (0.008)	-0.001 (0.005)
Race (African American)	-0.120 (0.016)	-0.110 (0.013)	0.116 (0.008)	-0.007 (0.006)
Race (Native American)	-0.003 (0.048)	-0.026 (0.043)	0.021 (0.038)	-0.097 (0.038)
Race (Asian)	-0.215 (0.029)	-0.146 (0.022)	-0.124 (0.015)	-0.063 (0.010)
Race (non-white Hispanic)	-0.008 (0.011)	-0.007 (0.014)	-0.007 (0.009)	0.003 (0.009)
Age of student (17 years old)	0.081 (0.039)	-0.126 (0.010)	-0.111 (0.010)	-0.035 (0.008)
Age of student (18 years old)	0.043 (0.038)	-0.100 (0.032)	-0.084(0.013)	0.017 (0.009)
Age of student (19 years old)	0.015 (0.013)	-0.066 (0.032)	-0.053 (0.015)	0.010 (0.006)
Last grade completed (9th grade)	0.023 (0.034)	0.023 (0.008)	0.022 (0.009)	-0.009 (0.005)
Last grade completed (10th grade)	0.050 (0.012)	0.049 (0.013)	0.044 (0.011)	0.001 (0.020)
Last grade completed (11th grade)	0.091 (0.035)	0.091 (0.035)	0.076 (0.031)	0.002 (0.021)
Number in family	-0.025 (0.006)	-0.018 (0.004)	-0.011 (0.007)	-0.008 (0.005)
Number in family (>5)	-0.158 (0.027)	-0.098 (0.022)	-0.065 (0.014)	-0.025 (0.010)
Family income ($20,000-$49,999)	0.071 (0.015)	0.059 (0.015)	0.030 (0.008)	-0.015 (0.005)
Family income ($50,000-$74,999)	0.055 (0.017)	0.046 (0.017)	0.036 (0.010)	-0.010 (0.011)
Family income ($75,000 or more)	0.097 (0.016)	0.099 (0.017)	0.058 (0.009)	0.019 (0.008)
MSA segment with 1+ million persons	0.002 (0.010)	-0.023 (0.009)	-0.022 (0.008)	-0.012 (0.006)
MSA segment of less than 1 million	0.004 (0.013)	-0.023 (0.009)	-0.019 (0.008)	-0.008 (0.005)
Year 2002 indicator	0.020 (0.014)	0.024 (0.009)	0.008 (0.007)	0.003 (0.006)

Standard errors are in parentheses.

Appendix 10. All IV estimates on absenteeism for binge drinking
(n=10,039)

16-19 sample	days missed due to skipping	days missed due to illness
Explanatory variables	**IV coefficient (Marginal Effect SE)**	**IV coefficient (Marginal Effect SE)**
Binge drinking	1.418 (0.186)	0.622 (0.197)
Female	0.045 (0.035)	0.298 (0.047)
Race (African American)	0.361 (0.062)	0.125 (0.086)
Race (Native American)	-0.021 (0.127)	0.483 (0.231)
Race (Asian)	0.014 (0.069)	-0.070 (0.125)
Race (non-white Hispanic)	0.090 (0.161)	0.041 (0.071)
Age of student (17 years old)	-0.297 (0.157)	0.054 (0.148)
Age of student (18 years old)	-0.149 (0.153)	0.316 (0.162)
Age of student (19 years old)	-0.090(0.161)	-0.482 (0.172)
Last grade completed (9th grade)	-0.190 (0.126)	-0.098 (0.171)
Last grade completed (10th grade)	-0.313 (0.121)	-0.316 (0.163)
Last grade completed (11th grade)	-0.404 (0.128)	-0.482 (0.173)
Number in family	-0.009 (0.020)	-0.028 (0.029)
Number in family (>5)	0.031 (0.093)	-0.006 (0.134)
Family income ($20,000-$49,999)	-0.175 (0.061)	-0.424 (0.082)
Family income ($50,000-$74,999)	-0.268 (0.064)	-0.582 (0.088)
Family income ($75,000 or more)	-0.398 (0.063)	-0.701 (0.086)
MSA segement with 1+ million persons	0.156 (0.045)	0.108 (0.059)
MSA segment of less than 1 million	0.054 (0.040)	0.061 (0.057)
Year 2002 indicator	-0.019 (0.033)	-0.116 (0.044)

Standard errors are in parentheses.

Appendix 11. All IV estimates for absenteeism because of binge drinking
(n=8,817)

18-25 sample

	days missed due to skipping	days missed due to illness
Explanatory variables	**IV coefficient (Marginal Effect SE)**	**IV coefficient (Marginal Effect SE)**
Binge drinking	0.954 (0.179)	0.352 (0.150)
Female	0.062 (0.048)	0.228 (0.045)
Race (African American)	0.449 (0.083)	0.283 (0.080)
Race (Native American)	0.221 (0.334)	-0.069 (0.158)
Race (Asian)	0.339 (0.141)	-0.015 (0.171)
Race (non-white Hispanic)	0.165 (0.090)	0.166 (0.072)
Age of student (19 years old)	-0.081 (0.073)	-0.190 (0.065)
Age of student (20 years old)	-0.001 (0.101)	-0.218 (0.081)
Age of student (21 years old)	-0.045 (0.117)	-0.128 (0.087)
Age of student (22-23 years old)	-0.088 (0.103)	-0.202 (0.085)
Age of student (24-25 years old)	-0.014 (0.119)	-0.021 (0.118)
Last grade completed (Freshman)	0.172 (0.079)	-0.154 (0.068)
Last grade completed (Sophomore/ Junior)	-0.084 (0.077)	-0.281 (0.081)
Number in family	0.023 (0.026)	-0.047 (0.022)
Number in family (>5)	-0.049 (0.120)	-0.209 (0.113)
Family income ($20,000-$49,999)	0.103 (0.065)	0.027 (0.056)
Family income ($50,000-$74,999)	0.153 (0.077)	-0.087(0.065)
Family income ($75,000 or more)	0.229 (0.065)	-0.109 (0.064)
MSA segement with 1+ million persons	0.064 (0.062)	0.048 (0.052)
MSA segment of less than 1 million	0.023 (0.054)	0.068 (0.045)
Year 2002 indicator	-0.020 (0.043)	-0.039 (0.039)

Standard errors are in parentheses.

About the Author

Wesley (Wes) Austin, is a Ph.D. candidate in the Department of Economics at the University of South Florida. He also holds an M.A. in economics and a bachelor degree in finance. While he is interested in several fields of economic research, his primary focus is in the areas of health economics, education economics and labor economics. Prior to entering the Ph.D. program in economics, he held positions as a financial analyst and market researcher and he taught economics on an adjunct basis. Among his varied hobbies are sports (especially football and golf) and aviation.